

*UNIVERSITE MOHAMMED V - RABAT*  
*FACULTE DE MEDECINE ET DE PHARMACIE - RABAT-*

*ANNEE: 2017*

*THESE N°: 434*

**LE SEQUENÇAGE DE L'ADN  
ET MEDECINE PERSONNALISEE**

**THÈSE**

*Présentée et soutenue publiquement le :.....*

**PAR**

**Mr. Mohamed EN-NOUALI**  
*Né le 12 Novembre 1992 à Marrakech*

**Pour l'Obtention du Doctorat en Médecine**

**MOTS CLES** : Génomique – Séquençage de nouvelle génération – Médecine individualisée –  
Médecine Personnalisée.

**JURY**

**Mr. T. DAKKA**

Professeur de Physiologie

**PRESIDENT**

**Mr. O. CHOKAIRI**

Professeur d'Histologie Embryologie

**RAPPORTEUR**

**Mr. F. ZOUAIDIA**

Professeur d'Anatomie Pathologie

**Mr. Y. RAHALI**

Professeur de Pharmacie Galénique

**JUGES**

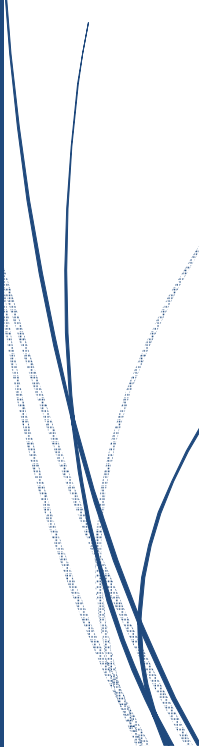


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

سبحانك لا علم لنا إلا ما علمتنا

إنك أنت العليم الحكيم

سورة البقرة: الآية: 31



صَلَّى  
الْعَظِيمِ



**UNIVERSITE MOHAMMED V DE RABAT**  
**FACULTE DE MEDECINE ET DE PHARMACIE - RABAT**

**DOYENS HONORAIRES :**

1962 – 1969 : Professeur Abdelmalek FARAJ  
1969 – 1974 : Professeur Abdellatif BERBICH  
1974 – 1981 : Professeur Bachir LAZRAK  
1981 – 1989 : Professeur Taieb CHKILI  
1989 – 1997 : Professeur Mohamed Tahar ALAOUI  
1997 – 2003 : Professeur Abdelmajid BELMAHI  
2003 – 2013 : Professeur Najia HAJJAJ - HASSOUNI



**ADMINISTRATION :**

**Doyen** : Professeur Mohamed ADNAOUI  
**Vice Doyen chargé des Affaires Académiques et étudiantes**  
Professeur Mohammed AHALLAT  
**Vice Doyen chargé de la Recherche et de la Coopération**  
Professeur Taoufiq DAKKA  
**Vice Doyen chargé des Affaires Spécifiques à la Pharmacie**  
Professeur Jamal TAOUFIK  
**Secrétaire Général** : Mr. Mohamed KARRA

**1- ENSEIGNANTS-CHERCHEURS MEDECINS**

**ET  
PHARMACIENS**

**PROFESSEURS :**

**Décembre 1984**

Pr. MAAOUNI Abdelaziz  
Pr. MAAZOUZI Ahmed Wajdi  
Pr. SETTAF Abdellatif

Médecine Interne – **Clinique Royale**  
Anesthésie -Réanimation  
pathologie Chirurgicale

**Novembre et Décembre 1985**

Pr. BENSAID Younes

Pathologie Chirurgicale

**Janvier, Février et Décembre 1987**

Pr. CHAHED OUZZANI Houria  
Pr. LACHKAR Hassan  
Pr. YAHYAOUY Mohamed

Gastro-Entérologie  
Médecine Interne  
Neurologie

**Décembre 1988**

Pr. BENHAMAMOUCH Mohamed Najib  
Pr. DAFIRI Rachida

Chirurgie Pédiatrique  
Radiologie

### **Décembre 1989**

Pr. ADNAOUI Mohamed  
Pr. CHAD Bouziane  
Pr. OUZZANI Taïbi Mohamed Réda

### **Janvier et Novembre 1990**

Pr. CHKOFF Rachid  
Pr. HACHIM Mohammed\*  
Pr. KHARBACH Aïcha  
Pr. MANSOURI Fatima  
Pr. TAZI Saoud Anas

### **Février Avril Juillet et Décembre 1991**

Pr. AL HAMANY Zaïtounia  
Pr. AZZOUZI Abderrahim  
Pr. BAYAHIA Rabéa  
Pr. BELKOUCHI Abdelkader  
Pr. BENCHEKROUN Belabbes Abdellatif  
Pr. BENSOUDA Yahia  
Pr. BERRAHO Amina  
Pr. BEZZAD Rachid  
Pr. CHABRAOUI Layachi  
Pr. CHERRAH Yahia  
Pr. CHOKAIRI Omar  
Pr. KHATTAB Mohamed  
Pr. SOULAYMANI Rachida  
Pr. TAOUFIK Jamal

### **Décembre 1992**

Pr. AHALLAT Mohamed  
Pr. BENSOUDA Adil  
Pr. BOUJIDA Mohamed Najib  
Pr. CHAHED OUZZANI Laaziza  
Pr. CHRAIBI Chafiq  
Pr. DEHAYNI Mohamed\*  
Pr. EL OUAHABI Abdessamad  
Pr. FELLAT Rokaya  
Pr. GHAFIR Driss\*  
Pr. JIDDANE Mohamed  
Pr. TAGHY Ahmed  
Pr. ZOUHDI Mimoun

### **Mars 1994**

Pr. BENJAFFAR Noureddine  
Pr. BEN RAIS Nozha  
Pr. CAOUI Malika  
Pr. CHRAIBI Abdelmjid  
  
Pr. EL AMRANI Sabah

Médecine Interne – **Doyen de la FMPR**  
Pathologie Chirurgicale  
Neurologie

Pathologie Chirurgicale  
Médecine-Interne  
Gynécologie -Obstétrique  
Anatomie-Pathologique  
Anesthésie Réanimation

Anatomie-Pathologique  
Anesthésie Réanimation – **Doyen de la FMPO**  
Néphrologie  
Chirurgie Générale  
Chirurgie Générale  
Pharmacie galénique  
Ophtalmologie  
Gynécologie Obstétrique  
Biochimie et Chimie  
Pharmacologie  
Histologie Embryologie  
Pédiatrie  
Pharmacologie – **Dir. du Centre National PV**  
Chimie thérapeutique **V.D à la pharmacie+Dir du CEDOC**

Chirurgie Générale V.D Aff. Acad. et Estud  
Anesthésie Réanimation  
Radiologie  
Gastro-Entérologie  
Gynécologie Obstétrique  
Gynécologie Obstétrique  
Neurochirurgie  
Cardiologie  
Médecine Interne  
Anatomie  
Chirurgie Générale  
Microbiologie



Radiothérapie  
Biophysique  
Biophysique  
Endocrinologie et Maladies Métaboliques **Doyen de la FMPA**  
Gynécologie Obstétrique

Pr. EL BARDOUNI Ahmed  
 Pr. EL HASSANI My Rachid  
 Pr. ERROUGANI Abdelkader  
 Pr. ESSAKALI Malika  
 Pr. ETTAYEBI Fouad  
 Pr. HADRI Larbi\*  
 Pr. HASSAM Badredine  
 Pr. IFRINE Lahssan  
 Pr. JELTHI Ahmed  
 Pr. MAHFOUD Mustapha  
 Pr. RHRAB Brahim  
 Pr. SENOUCI Karima

### **Mars 1994**

Pr. ABBAR Mohamed\*  
 Pr. ABDELHAK M'barek  
 Pr. BELAIDI Halima  
 Pr. BENTAHILA Abdelali  
 Pr. BENYAHIA Mohammed Ali  
 Pr. BERRADA Mohamed Saleh  
 Pr. CHAMI Ilham  
 Pr. CHERKAOUI Lalla Ouafae  
 Pr. JALIL Abdelouahed  
 Pr. LAKHDAR Amina  
 Pr. MOUANE Nezha

### **Mars 1995**

Pr. ABOUQUAL Redouane  
 Pr. AMRAOUI Mohamed  
 Pr. BAIDADA Abdelaziz  
 Pr. BARGACH Samir  
 Pr. CHAARI Jilali\*  
 Pr. DIMOU M'barek\*  
 Pr. DRISSI KAMILI Med Nordine\*  
 Pr. EL MESNAOUI Abbes  
 Pr. ESSAKALI HOUSSYNI Leila  
 Pr. HDA Abdelhamid\*  
 Pr. IBEN ATTYA ANDALOUSSI Ahmed  
 Pr. OUAZZANI CHAHDI Bahia  
 Pr. SEFIANI Abdelaziz  
 Pr. ZEGGWAGH Amine Ali

### **Décembre 1996**

Pr. AMIL Touriya\*  
 Pr. BELKACEM Rachid  
 Pr. BOULANOUAR Abdelkrim  
 Pr. EL ALAMI EL FARICHA EL Hassan  
 Pr. GAOUZI Ahmed  
 Pr. MAHFOUDI M'barek\*  
 Pr. OUADGHIRI Mohamed  
 Pr. OUZEDDOUN Naima  
 Pr. ZBIR EL Mehdi\*

Traumato-Orthopédie  
 Radiologie  
 Chirurgie Générale- **Directeur CHIS**  
 Immunologie  
 Chirurgie Pédiatrique  
 Médecine Interne  
 Dermatologie  
 Chirurgie Générale  
 Anatomie Pathologique  
 Traumatologie – Orthopédie  
 Gynécologie –Obstétrique  
 Dermatologie

Urologie  
 Chirurgie – Pédiatrique  
 Neurologie  
 Pédiatrie  
 Gynécologie – Obstétrique  
 Traumatologie – Orthopédie  
 Radiologie  
 Ophtalmologie  
 Chirurgie Générale  
 Gynécologie Obstétrique  
 Pédiatrie

Réanimation Médicale  
 Chirurgie Générale  
 Gynécologie Obstétrique  
 Gynécologie Obstétrique  
 Médecine Interne  
 Anesthésie Réanimation  
 Anesthésie Réanimation  
 Chirurgie Générale  
 Oto-Rhino-Laryngologie  
 Cardiologie - **Directeur HMI Med V**  
 Urologie  
 Ophtalmologie  
 Génétique  
 Réanimation Médicale

Radiologie  
 Chirurgie Pédiatrie  
 Ophtalmologie  
 Chirurgie Générale  
 Pédiatrie  
 Radiologie  
 Traumatologie-Orthopédie  
 Néphrologie  
 Cardiologie



### **Novembre 1997**

Pr. ALAMI Mohamed Hassan  
Pr. BEN SLIMANE Lounis  
Pr. BIROUK Nazha  
Pr. ERREIMI Naima  
Pr. FELLAT Nadia  
Pr. HAIMEUR Charki\*  
Pr. KADDOURI Nouredine  
Pr. KOUTANI Abdellatif  
Pr. LAHLOU Mohamed Khalid  
Pr. MAHRAOUI CHAFIQ  
Pr. TAOUFIQ Jallal  
Pr. YOUSFI MALKI Mounia

Gynécologie-Obstétrique  
Urologie  
Neurologie  
Pédiatrie  
Cardiologie  
Anesthésie Réanimation  
Chirurgie Pédiatrique  
Urologie  
Chirurgie Générale  
Pédiatrie  
Psychiatrie  
Gynécologie Obstétrique

### **Novembre 1998**

Pr. AFIFI RAJAA  
Pr. BENOMAR ALI  
Pr. BOUGTAB Abdesslam  
Pr. ER RIHANI Hassan  
Pr. BENKIRANE Majid\*  
Pr. KHATOUI ALI\*

Gastro-Entérologie  
Neurologie – *Doyen de la FMP Abulcassis*  
Chirurgie Générale  
Oncologie Médicale  
Hématologie  
Cardiologie

### **Janvier 2000**

Pr. ABID Ahmed\*  
Pr. AIT OUMAR Hassan  
Pr. BENJELLOUN Dakhama Badr.Sououd  
Pr. BOURKADI Jamal-Eddine  
Pr. CHARIF CHEFCHAOUNI Al Montacer  
Pr. ECHARRAB El Mahjoub  
Pr. EL FTOUH Mustapha  
Pr. EL MOSTARCHID Brahim\*  
Pr. ISMAILI Hassane\*  
Pr. MAHMOUDI Abdelkrim\*  
Pr. TACHINANTE Rajae  
Pr. TAZI MEZALEK Zoubida

Pneumophtisiologie  
Pédiatrie  
Pédiatrie  
Pneumo-phtisiologie  
Chirurgie Générale  
Chirurgie Générale  
Pneumo-phtisiologie  
Neurochirurgie  
Traumatologie Orthopédie- *Dir. Hop. Av. Marr.*  
Anesthésie-Réanimation *Inspecteur du SSM*  
Anesthésie-Réanimation  
Médecine Interne



### **Novembre 2000**

Pr. AIDI Saadia  
Pr. AJANA Fatima Zohra  
Pr. BENAMR Said  
Pr. CHERTI Mohammed  
Pr. ECH-CHERIF EL KETTANI Selma  
Pr. EL HASSANI Amine  
Pr. EL KHADER Khalid  
Pr. EL MAGHRAOUI Abdellah\*  
Pr. GHARBI Mohamed El Hassan  
Pr. MAHASSINI Najat  
Pr. MDAGHRI ALAOUI Asmae  
Pr. ROUIMI Abdelhadi\*

Neurologie  
Gastro-Entérologie  
Chirurgie Générale  
Cardiologie  
Anesthésie-Réanimation  
Pédiatrie *Directeur Hop. Chekikh Zaied*  
Urologie  
Rhumatologie  
Endocrinologie et Maladies Métaboliques  
Anatomie Pathologique  
Pédiatrie  
Neurologie

## Décembre 2000

Pr. ZOHAIR ABDELAH\*

ORL

## Décembre 2001

Pr. BALKHI Hicham\*  
Pr. BENABDELJLIL Maria  
Pr. BENAMAR Loubna  
Pr. BENAMOR Jouda  
Pr. BENELBARHDADI Imane  
Pr. BENNANI Rajae  
Pr. BENOUECHANE Thami  
Pr. BEZZA Ahmed\*  
Pr. BOUCHIKHI IDRISSE Med Larbi  
Pr. BOUMDIN El Hassane\*  
Pr. CHAT Latifa  
Pr. DAALI Mustapha\*  
Pr. DRISSI Sidi Mourad\*  
Pr. EL HIJRI Ahmed  
Pr. EL MAAQILI Moulay Rachid  
Pr. EL MADHI Tarik  
Pr. EL OUNANI Mohamed  
Pr. ETTAIR Said  
Pr. GAZZAZ Miloudi\*  
Pr. HRORA Abdelmalek  
Pr. KABBAJ Saad  
Pr. KABIRI EL Hassane\*  
Pr. LAMRANI Moulay Omar  
Pr. LEKEHAL Brahim  
Pr. MAHASSIN Fattouma\*  
Pr. MEDARHRI Jalil  
Pr. MIKDAME Mohammed\*  
Pr. MOHSINE Raouf  
Pr. NOUINI Yassine  
Pr. SABBABH Farid  
Pr. SEFIANI Yasser  
Pr. TAOUFIQ BENCHEKROUN Soumia

Anesthésie-Réanimation  
Neurologie  
Néphrologie  
Pneumo-phtisiologie  
Gastro-Entérologie  
Cardiologie  
Pédiatrie  
Rhumatologie  
Anatomie  
Radiologie  
Radiologie  
Chirurgie Générale  
Radiologie  
Anesthésie-Réanimation  
Neuro-Chirurgie  
Chirurgie-Pédiatrique  
Chirurgie Générale  
Pédiatrie **Directeur. Hop.d'Enfants**  
Neuro-Chirurgie  
Chirurgie Générale  
Anesthésie-Réanimation  
Chirurgie Thoracique  
Traumatologie Orthopédie  
Chirurgie Vasculaire Périphérique  
Médecine Interne  
Chirurgie Générale  
Hématologie Clinique  
Chirurgie Générale  
Urologie **Directeur Hôpital Ibn Sina**  
Chirurgie Générale  
Chirurgie Vasculaire Périphérique  
Pédiatrie



## Décembre 2002

Pr. AL BOUZIDI Abderrahmane\*  
Pr. AMEUR Ahmed \*  
Pr. AMRI Rachida  
Pr. AOURARH Aziz\*  
Pr. BAMOU Youssef \*  
Pr. BELMEJDOUB Ghizlene\*  
Pr. BENZEKRI Laila  
Pr. BENZZOUBEIR Nadia  
Pr. BERNOUSSI Zakiya  
Pr. BICHRA Mohamed Zakariya\*  
Pr. CHOHO Abdelkrim \*

Anatomie Pathologique  
Urologie  
Cardiologie  
Gastro-Entérologie  
Biochimie-Chimie  
Endocrinologie et Maladies Métaboliques  
Dermatologie  
Gastro-Entérologie  
Anatomie Pathologique  
Psychiatrie  
Chirurgie Générale

Pr. CHKIRATE Bouchra  
 Pr. EL ALAMI EL FELLOUS Sidi Zouhair  
 Pr. EL HAOURI Mohamed \*  
 Pr. FILALI ADIB Abdelhai  
 Pr. HAJJI Zakia  
 Pr. IKEN Ali  
 Pr. JAAFAR Abdeloihab\*  
 Pr. KRIOUILE Yamina  
 Pr. LAGHMARI Mina  
 Pr. MABROUK Hfid\*  
 Pr. MOUSSAOUI RAHALI Driss\*  
 Pr. OUJILAL Abdelilah  
 Pr. RACHID Khalid \*  
 Pr. RAISS Mohamed  
 Pr. RGUIBI IDRISSE Sidi Mustapha\*  
 Pr. RHOU Hakima  
 Pr. SIAH Samir \*  
 Pr. THIMOU Amal  
 Pr. ZENTAR Aziz\*

#### **Janvier 2004**

Pr. ABDELLAH El Hassan  
 Pr. AMRANI Mariam  
 Pr. BENBOUZID Mohammed Anas  
 Pr. BENKIRANE Ahmed\*  
 Pr. BOUGHALEM Mohamed\*  
 Pr. BOULAADAS Malik  
 Pr. BOURAZZA Ahmed\*  
 Pr. CHAGAR Belkacem\*  
 Pr. CHERRADI Nadia  
 Pr. EL FENNI Jamal\*  
 Pr. EL HANCHI ZAKI  
 Pr. EL KHORASSANI Mohamed  
 Pr. EL YOUNASSI Badreddine\*  
 Pr. HACHI Hafid  
 Pr. JABOUIRIK Fatima  
 Pr. KHARMAZ Mohamed  
 Pr. MOUGHIL Said  
 Pr. OUBAAZ Abdelbarre\*  
 Pr. TARIB Abdelilah\*  
 Pr. TIJAMI Fouad  
 Pr. ZARZUR Jamila

#### **Janvier 2005**

Pr. ABBASSI Abdellah  
 Pr. AL KANDRY Sif Eddine\*  
 Pr. ALLALI Fadoua  
 Pr. AMAZOUZI Abdellah  
 Pr. AZIZ Nouredine\*  
 Pr. BAHIRI Rachid

Pédiatrie  
 Chirurgie Pédiatrique  
 Dermatologie  
 Gynécologie Obstétrique  
 Ophtalmologie  
 Urologie  
 Traumatologie Orthopédie  
 Pédiatrie  
 Ophtalmologie  
 Traumatologie Orthopédie  
 Gynécologie Obstétrique  
 Oto-Rhino-Laryngologie  
 Traumatologie Orthopédie  
 Chirurgie Générale  
 Pneumophtisiologie  
 Néphrologie  
 Anesthésie Réanimation  
 Pédiatrie  
 Chirurgie Générale

Ophtalmologie  
 Anatomie Pathologique  
 Oto-Rhino-Laryngologie  
 Gastro-Entérologie  
 Anesthésie Réanimation  
 Stomatologie et Chirurgie Maxillo-faciale  
 Neurologie  
 Traumatologie Orthopédie  
 Anatomie Pathologique  
 Radiologie  
 Gynécologie Obstétrique  
 Pédiatrie  
 Cardiologie  
 Chirurgie Générale  
 Pédiatrie  
 Traumatologie Orthopédie  
 Chirurgie Cardio-Vasculaire  
 Ophtalmologie  
 Pharmacie Clinique  
 Chirurgie Générale  
 Cardiologie

Chirurgie Réparatrice et Plastique  
 Chirurgie Générale  
 Rhumatologie  
 Ophtalmologie  
 Radiologie  
 Rhumatologie





Pr. BARKAT Amina  
 Pr. BENYASS Aatif  
 Pr. BERNOUSSI Abdelghani  
 Pr. DOUDOUH Abderrahim\*  
 Pr. EL HAMZAOUI Sakina\*  
 Pr. HAJJI Leila  
 Pr. HESSISSEN Leila  
 Pr. JIDAL Mohamed\*  
 Pr. LAAROUSSI Mohamed  
 Pr. LYAGOUBI Mohammed  
 Pr. NIAMANE Radouane\*  
 Pr. RAGALA Abdelhak  
 Pr. SBIHI Souad  
 Pr. ZERAIDI Najia

### **Décembre 2005**

Pr. CHANI Mohamed

### **Avril 2006**

Pr. ACHEMLAL Lahsen\*  
 Pr. AKJOUJ Said\*  
 Pr. BELMEKKI Abdelkader\*  
 Pr. BENCHEIKH Razika  
 Pr. BIYI Abdelhamid\*  
 Pr. BOUHAFS Mohamed El Amine  
 Pr. BOULAHYA Abdellatif\*  
 Pr. CHENGUETI ANSARI Anas  
 Pr. DOGHMI Nawal  
 Pr. FELLAT Ibtissam  
 Pr. FAROUDY Mamoun  
 Pr. HARMOUCHE Hicham  
 Pr. HANAFI Sidi Mohamed\*  
 Pr. IDRIS LAHLOU Amine\*  
 Pr. JROUNDI Laila  
 Pr. KARMOUNI Tariq  
 Pr. KILI Amina  
 Pr. KISRA Hassan  
 Pr. KISRA Mounir  
 Pr. LAATIRIS Abdelkader\*  
 Pr. LMIMOUNI Badreddine\*  
 Pr. MANSOURI Hamid\*  
 Pr. OUANASS Abderrazzak  
 Pr. SAFI Soumaya\*  
 Pr. SEKKAT Fatima Zahra  
 Pr. SOUALHI Mouna  
 Pr. TELLAL Saida\*  
 Pr. ZAHRAOUI Rachida

### **Octobre 2007**

Pr. ABIDI Khalid  
 Pr. ACHACHI Leila

Pédiatrie  
 Cardiologie  
 Ophtalmologie  
 Biophysique  
 Microbiologie  
 Cardiologie (mise en disponibilité)  
 Pédiatrie  
 Radiologie  
 Chirurgie Cardio-vasculaire  
 Parasitologie  
 Rhumatologie  
 Gynécologie Obstétrique  
 Histo-Embryologie Cytogénétique  
 Gynécologie Obstétrique

Anesthésie Réanimation

Rhumatologie  
 Radiologie  
 Hématologie  
 O.R.L  
 Biophysique  
 Chirurgie - Pédiatrique  
 Chirurgie Cardio – Vasculaire  
 Gynécologie Obstétrique  
 Cardiologie  
 Cardiologie  
 Anesthésie Réanimation  
 Médecine Interne  
 Anesthésie Réanimation  
 Microbiologie  
 Radiologie  
 Urologie  
 Pédiatrie  
 Psychiatrie  
 Chirurgie – Pédiatrique  
 Pharmacie Galénique  
 Parasitologie  
 Radiothérapie  
 Psychiatrie  
 Endocrinologie  
 Psychiatrie  
 Pneumo – Phtisiologie  
 Biochimie  
 Pneumo – Phtisiologie



Pr. ACHOUR Abdessamad\*  
 Pr. AIT HOUSSA Mahdi\*  
 Pr. AMHAJJI Larbi\*  
 Pr. AOUI Sarra  
 Pr. BAITE Abdelouahed\*  
 Pr. BALOUCH Lhoussaine\*  
 Pr. BENZIANE Hamid\*  
 Pr. BOUTIMZINE Nourdine  
 Pr. CHARKAOUI Naoual\*  
 Pr. EHIRCHIOU Abdelkader\*  
 Pr. ELABSI Mohamed  
 Pr. EL MOUSSAOUI Rachid  
 Pr. EL OMARI Fatima  
 Pr. GHARIB Noureddine  
 Pr. HADADI Khalid\*  
 Pr. ICHOU Mohamed\*  
 Pr. ISMAILI Nadia  
 Pr. KEBDANI Tayeb  
 Pr. LALAOUI SALIM Jaafar\*  
 Pr. LOUZI Lhoussain\*  
 Pr. MADANI Naoufel  
 Pr. MAHI Mohamed\*  
 Pr. MARC Karima  
 Pr. MASRAR Azlarab  
 Pr. MRABET Mustapha\*  
 Pr. MRANI Saad\*  
 Pr. OUZZIF Ez zohra\*  
 Pr. RABHI Monsef\*  
 Pr. RADOUANE Bouchaib\*  
 Pr. SEFFAR Myriame  
 Pr. SEKHSOKH Yessine\*  
 Pr. SIFAT Hassan\*  
 Pr. TABERKANET Mustafa\*  
 Pr. TACHFOUTI Samira  
 Pr. TAJDINE Mohammed Tariq\*  
 Pr. TANANE Mansour\*  
 Pr. TLIGUI Houssain  
 Pr. TOUATI Zakia

### **Décembre 2007**

Pr. DOUHAL ABDERRAHMAN

### **Décembre 2008**

Pr ZOUBIR Mohamed\*  
 Pr TAHIRI My El Hassan\*

Chirurgie générale  
 Chirurgie cardio vasculaire  
 Traumatologie orthopédie  
 Parasitologie  
 Anesthésie réanimation **Directeur ERSM**  
 Biochimie-chimie  
 Pharmacie clinique  
 Ophtalmologie  
 Pharmacie galénique  
 Chirurgie générale  
 Chirurgie générale  
 Anesthésie réanimation  
 Psychiatrie  
 Chirurgie plastique et réparatrice  
 Radiothérapie  
 Oncologie médicale  
 Dermatologie  
 Radiothérapie  
 Anesthésie réanimation  
 Microbiologie  
 Réanimation médicale  
 Radiologie  
 Pneumo phtisiologie  
 Hématologique  
 Médecine préventive santé publique et hygiène  
 Virologie  
 Biochimie-chimie  
 Médecine interne  
 Radiologie  
 Microbiologie  
 Microbiologie  
 Radiothérapie  
 Chirurgie vasculaire périphérique  
 Ophtalmologie  
 Chirurgie générale  
 Traumatologie orthopédie  
 Parasitologie  
 Cardiologie



Ophtalmologie

Anesthésie Réanimation  
 Chirurgie Générale

### **Mars 2009**

Pr. ABOUZAHIR Ali\*  
Pr. AGDR Aomar\*  
Pr. AIT ALI Abdelmounaim\*  
Pr. AIT BENHADDOU El hachmia  
Pr. AKHADDAR Ali\*  
Pr. ALLALI Nazik  
Pr. AMINE Bouchra  
Pr. ARKHA Yassir  
Pr. BELYAMANI Lahcen\*  
Pr. BJIJOU Younes  
Pr. BOUHSAIN Sanae\*  
Pr. BOUI Mohammed\*  
Pr. BOUNAIM Ahmed\*  
Pr. BOUSSOUGA Mostapha\*  
Pr. CHAKOUR Mohammed \*  
Pr. CHTATA Hassan Toufik\*  
Pr. DOGHMI Kamal\*  
Pr. EL MALKI Hadj Omar  
Pr. EL OUENNASS Mostapha\*  
Pr. ENNIBI Khalid\*  
Pr. FATHI Khalid  
Pr. HASSIKOU Hasna \*  
Pr. KABBAJ Nawal  
Pr. KABIRI Meryem  
Pr. KARBOUBI Lamya  
Pr. L'KASSIMI Hachemi\*  
Pr. LAMSAOURI Jamal\*  
Pr. MARMADE Lahcen  
Pr. MESKINI Toufik  
Pr. MESSAOUDI Nezha \*  
Pr. MSSROURI Rahal  
Pr. NASSAR Ittimade  
Pr. OUKERRAJ Latifa  
Pr. RHORFI Ismail Abderrahmani \*

### **PROFESSEURS AGREGES :**

#### **Octobre 2010**

Pr. ALILOU Mustapha  
Pr. AMEZIANE Taoufiq\*  
Pr. BELAGUID Abdelaziz  
Pr. BOUAITY Brahim\*  
Pr. CHADLI Mariama\*  
Pr. CHEMSI Mohamed\*  
Pr. DAMI Abdellah\*  
Pr. DARBI Abdellatif\*  
Pr. DENDANE Mohammed Anouar  
Pr. EL HAFIDI Naima  
Pr. EL KHARRAS Abdennasser\*

Médecine interne  
Pédiatre  
Chirurgie Générale  
Neurologie  
Neuro-chirurgie  
Radiologie  
Rhumatologie  
Neuro-chirurgie  
Anesthésie Réanimation  
Anatomie  
Biochimie-chimie  
Dermatologie  
Chirurgie Générale  
Traumatologie orthopédique  
Hématologie biologique  
Chirurgie vasculaire périphérique  
Hématologie clinique  
Chirurgie Générale  
Microbiologie  
Médecine interne  
Gynécologie obstétrique  
Rhumatologie  
Gastro-entérologie  
Pédiatrie  
Pédiatrie  
Microbiologie ***Directeur Hôpital My Ismail***  
Chimie Thérapeutique  
Chirurgie Cardio-vasculaire  
Pédiatrie  
Hématologie biologique  
Chirurgie Générale  
Radiologie  
Cardiologie  
Pneumo-phtisiologie



Anesthésie réanimation  
Médecine interne  
Physiologie  
ORL  
Microbiologie  
Médecine aéronautique  
Biochimie chimie  
Radiologie  
Chirurgie pédiatrique  
Pédiatrie  
Radiologie

Pr. EL MAZOUZ Samir  
 Pr. EL SAYEGH Hachem  
 Pr. ERRABIH Ikram  
 Pr. LAMALMI Najat  
 Pr. MOSADIK Ahlam  
 Pr. MOUJAHID Mountassir\*  
 Pr. NAZIH Mouna\*  
 Pr. ZOUAIDIA Fouad

Chirurgie plastique et réparatrice  
 Urologie  
 Gastro entérologie  
 Anatomie pathologique  
 Anesthésie Réanimation  
 Chirurgie générale  
 Hématologie  
 Anatomie pathologique

### **Mai 2012**

Pr. AMRANI Abdelouahed  
 Pr. ABOUELALAA Khalil\*  
 Pr. BELAIZI Mohamed\*  
 Pr. BENCHEBBA Driss\*  
 Pr. DRISSI Mohamed\*  
 Pr. EL ALAOUI MHAMDI Mouna  
 Pr. EL KHATTABI Abdessadek\*  
 Pr. EL OUAZZANI Hanane\*  
 Pr. ER-RAJI Mounir  
 Pr. JAHID Ahmed  
 Pr. MEHSSANI Jamal\*  
 Pr. RAISSOUNI Maha\*

Chirurgie Pédiatrique  
 Anesthésie Réanimation  
 Psychiatrie  
 Traumatologie Orthopédique  
 Anesthésie Réanimation  
 Chirurgie Générale  
 Médecine Interne  
 Pneumophtisiologie  
 Chirurgie Pédiatrique  
 Anatomie pathologique  
 Psychiatrie  
 Cardiologie

### **Février 2013**

Pr. AHID Samir  
 Pr. AIT EL CADI Mina  
 Pr. AMRANI HANCHI Laila  
 Pr. AMOUR Mourad  
 Pr. AWAB Almahdi  
 Pr. BELAYACHI Jihane  
 Pr. BELKHADIR Zakaria Houssain  
 Pr. BENCHEKROUN Laila  
 Pr. BENKIRANE Souad  
 Pr. BENNANA Ahmed\*  
 O.  
 Pr. BENSGHIR Mustapha\*  
 Pr. BENYAHIA Mohammed\*  
 Pr. BOUATIA Mustapha  
 Pr. BOUABID Ahmed Salim\*  
 Pr. BOUTARBOUCH Mahjouba  
 Pr. CHAIB Ali\*  
 Pr. DENDANE Tarek  
 Pr. DINI Nouzha\*  
 Pr. ECH-CHERIF EL KETTANI Mohamed Ali  
 Pr. ECH-CHERIF EL KETTANI Najwa  
 Pr. ELFATEMI Nizare  
 Pr. EL GUERROUJ Hasnae  
 Pr. EL HARTI Jaouad

Pharmacologie – Chimie  
 Toxicologie  
 Gastro-Entérologie  
 Anesthésie Réanimation  
 Anesthésie Réanimation  
 Réanimation Médicale  
 Anesthésie Réanimation  
 Biochimie-Chimie  
 Hématologie  
 Informatique Pharmaceutique

Anesthésie Réanimation  
 Néphrologie  
 Chimie Analytique  
 Traumatologie Orthopédie  
 Anatomie  
 Cardiologie  
 Réanimation Médicale  
 Pédiatrie  
 Anesthésie Réanimation  
 Radiologie  
 Neuro-Chirurgie  
 Médecine Nucléaire  
 Chimie Thérapeutique



Pr. EL JOUDI Rachid\*  
 Pr. EL KABABRI Maria  
 Pr. EL KHANNOUSSI Basma  
 Pr. EL KHLOUFI Samir  
 Pr. EL KORAICHI Alae  
 Pr. EN-NOUALI Hassane\*  
 Pr. ERREGUIG Laila  
 Pr. FIKRI Meryim  
 Pr. GHFIR Imade  
 Pr. IMANE Zineb  
 Pr. IRAQI Hind  
 Pr. KABBAJ Hakima  
 Pr. KADIRI Mohamed\*  
 Pr. LATIB Rachida  
 Pr. MAAMAR Mouna Fatima Zahra  
 Pr. MEDDAH Bouchra  
 Pr. MELHAOUI Adyl  
 Pr. MRABTI Hind  
 Pr. NEJJARI Rachid  
 Pr. OUBEJJA Houda  
 Pr. OUKABLI Mohamed\*  
 Pr. RAHALI Younes  
 Pr. RATBI Ilham  
 Pr. RAHMANI Mounia  
 Pr. REDA Karim\*  
 Pr. REGRAGUI Wafa  
 Pr. RKAIN Hanan  
 Pr. ROSTOM Samira  
 Pr. ROUAS Lamiaa  
 Pr. ROUIBAA Fedoua\*  
 Pr. SALIHOUN Mouna  
 Pr. SAYAH Rochde  
 Pr. SEDDIK Hassan\*  
 Pr. ZERHOUNI Hicham  
 Pr. ZINE Ali\*

Toxicologie  
 Pédiatrie  
 Anatomie Pathologie  
 Anatomie  
 Anesthésie Réanimation  
 Radiologie  
 Physiologie  
 Radiologie  
 Médecine Nucléaire  
 Pédiatrie  
 Endocrinologie et maladies métaboliques  
 Microbiologie  
 Psychiatrie  
 Radiologie  
 Médecine Interne  
 Pharmacologie  
 Neuro-chirurgie  
 Oncologie Médicale  
 Pharmacognosie  
 Chirurgie Pédiatrique  
 Anatomie Pathologique  
 Pharmacie Galénique  
 Génétique  
 Neurologie  
 Ophtalmologie  
 Neurologie  
 Physiologie  
 Rhumatologie  
 Anatomie Pathologique  
 Gastro-Entérologie  
 Gastro-Entérologie  
 Chirurgie Cardio-Vasculaire  
 Gastro-Entérologie  
 Chirurgie Pédiatrique  
 Traumatologie Orthopédie

### **Avril 2013**

Pr. EL KHATIB Mohamed Karim\*  
 Pr. GHOUNDALE Omar\*  
 Pr. ZYANI Mohammad\*

Stomatologie et Chirurgie Maxillo-faciale  
 Urologie  
 Médecine Interne

**\*Enseignants Militaires**



## **MARS 2014**

ACHIR ABDELLAH  
BENCHAKROUN MOHAMMED  
BOUCHIKH MOHAMMED  
EL KABBAJ DRISS  
EL MACHTANI IDRISSE SAMIRA  
HARDIZI HOUYAM  
HASSANI AMALE  
HERRAK LAILA  
JANANE ABDELLA TIF  
JEAIDI ANASS  
KOUACH JAOUAD  
LEMNOUER ABDELHAY  
MAKRAM SANAA  
OULAHYANE RACHID  
RHISSASSI MOHAMED JMFAR  
SABRY MOHAMED  
SEKKACH YOUSSEF  
TAZL MOUKBA. :LA.KLA.

**\*Enseignants Militaires**

Chirurgie Thoracique  
Traumatologie- Orthopédie  
Chirurgie Thoracique  
Néphrologie  
Biochimie-Chimie  
Histologie- Embryologie-Cytogénétique  
Pédiatrie  
Pneumologie  
Urologie  
Hématologie Biologique  
Génécologie-Obstétrique  
Microbiologie  
Pharmacologie  
Chirurgie Pédiatrique  
CCV  
Cardiologie  
Médecine Interne  
Génécologie-Obstétrique

## **DECEMBRE 2014**

ABILKACEM RACHID'  
AIT BOUGHIMA FADILA  
BEKKALI HICHAM  
BENAZZOU SALMA  
BOUABDELLAH MOUNYA  
BOUCHRIK MOURAD  
DERRAJI SOUFIANE  
DOBLALI TAOUFIK  
EL AYOUBI EL IDRISSE ALI  
EL GHADBANE ABDEDAIM HATIM  
EL MARJANY MOHAMMED  
FEJJAL NAWFAL  
JAHIDI MOHAMED  
LAKHAL ZOUHAIR  
OUDGHIRI NEZHA  
Rami Mohamed  
SABIR MARIA  
SBAI IDRISSE KARIM

**\*Enseignants Militaires**

Pédiatrie  
Médecine Légale  
Anesthésie-Réanimation  
Chirurgie Maxillo-Faciale  
Biochimie-Chimie  
Parasitologie  
Pharmacie Clinique  
Microbiologie  
Anatomie  
Anesthésie-Réanimation  
Radiothérapie  
Chirurgie Réparatrice et Plastique  
O.R.L  
Cardiologie  
Anesthésie-Réanimation  
Chirurgie Pédiatrique  
Psychiatrie  
Médecine préventive, santé publique et Hyg.



## **AOUT 2015**

Meziane meryem  
Tahri latifa

Dermatologie  
Rhumatologie

## **JANVIER 2016**

BENKABBOU AMINE  
EL ASRI FOUAD  
ERRAMI NOUREDDINE  
NITASSI SOPHIA

Chirurgie Générale  
Ophtalmologie  
O.R.L  
O.R.L

## **2- ENSEIGNANTS – CHERCHEURS SCIENTIFIQUES**

### **PROFESSEURS / PRs. HABILITES**

Pr. ABOUDRAR Saadia  
Pr. ALAMI OUHABI Naima  
Pr. ALAOUI KATIM  
Pr. ALAOUI SLIMANI Lalla Naïma  
Pr. ANSAR M'hammed  
Pr. BOUHOUCHE Ahmed  
Pr. BOUKLOUZE Abdelaziz  
Pr. BOURJOUANE Mohamed  
Pr. CHAHED OUAZZANI Lalla Chadia  
Pr. DAKKA Taoufiq  
Pr. DRAOUI Mustapha  
Pr. EL GUESSABI Lahcen  
Pr. ETTAIB Abdelkader  
Pr. FAOUZI Moulay El Abbes  
Pr. HAMZAOUI Laila  
Pr. HMAMOUCHE Mohamed  
Pr. IBRAHIMI Azeddine  
Pr. KHANFRI Jamal Eddine  
Pr. OULAD BOUYAHYA IDRISSE Med  
Pr. REDHA Ahlam  
Pr. TOUATI Driss  
Pr. ZAHIDI Ahmed  
Pr. ZELLOU Amina

Physiologie  
Biochimie – chimie  
Pharmacologie  
Histologie-Embryologie  
Chimie Organique et Pharmacie Chimique  
Génétique Humaine  
Applications Pharmaceutiques  
Microbiologie  
Biochimie – chimie  
Physiologie  
Chimie Analytique  
Pharmacognosie  
Zootechnie  
Pharmacologie  
Biophysique  
Chimie Organique  
Biologie moléculaire  
Biologie  
Chimie Organique  
Chimie  
Pharmacognosie  
Pharmacologie  
Chimie Organique

*Mise à jour le 14/12/2016 par le  
Service des Ressources Humaines*





# *Dédicaces*



*JE DEDIE CETTE THESE ...*

*A mon père et à ma mère*

*Aucune dédicace ne saurait exprimer tout le respect, la reconnaissance et  
l'amour que je vous porte.*

*A mon grand frère Amine*

*Je te souhaite une vie pleine de bonheur, de santé et de prospérité.*

*A tous les membres de ma famille petits et grands*

*Veillez trouver dans ce modeste travail l'expression de mon affection la  
plus sincère.*

*A tous mes amis et mes collègues*

*En guise de reconnaissance et d'amitié pure, je vous dédie ce modeste  
travail.*

*A tous ceux qui me sont chers*



# *Remerciements*

*Notre maître et président de thèse*

*Mr le professeur DAKKA T.*

*Professeur de physiologie*

*Nous sommes très touchés par l'honneur que vous nous faites en acceptant  
de présider le jury de cette thèse.*

*C'est avec grande joie que nous avons accueilli votre accord.*

*Que ce travail soit pour nous une occasion de vous exprimer notre  
admiration et notre profond respect.*

*A notre maître et rapporteur de thèse*

*Mr le professeur CHOKAIRI O.*

*Professeur d'Histologie et Embryologie.*

*Nous vous remercions de nous faire avoir fait l'honneur de nous confier ce  
travail.*

*Acceptez, cher maître, l'hommage de notre gratitude qui, si grande qu'elle  
puisse être, ne sera jamais à la hauteur de votre dévouement.*

*A notre maître et juge de thèse*

*Mr. le professeur ZOUIDIA F.*

*Professeur d'Anatomie et Pathologie*

*Vous avez accepté de siéger parmi le jury de notre thèse. Ce geste dénote non seulement de votre gentillesse mais surtout de votre souci du devoir envers vos étudiants.*

*Veillez accepter Monsieur le Professeur, ma profonde reconnaissance et mes remerciements les plus sincères.*

*Soyez assuré que c'est une fierté pour nous de vous compter parmi les membres de notre jury.*

*A notre maître et juge de thèse*

*Mr. le professeur RAHALI Y.*

*Professeur de Pharmacie Galénique*

*C'est pour nous un immense privilège de vous voir accepter de juger ce  
travail.*

*Veillez croire cher maître à notre très haute considération et notre profond  
respect.*



*Enfin qu'il nous soit permis d'exprimer globalement nos remerciements à  
tous ceux et à toutes celles qui ont manifesté leur appui et ont facilité notre  
travail.*

*Merci*

*Mohamed En-nouali*

## Liste des abréviations

|              |                                                                                                                                          |
|--------------|------------------------------------------------------------------------------------------------------------------------------------------|
| ADN          | : Acide désoxyribonucléique                                                                                                              |
| a-hémolysine | : Alpha-Hémolysine                                                                                                                       |
| ARN          | : Acide ribonucléique                                                                                                                    |
| CAD          | : Coronary artery disease                                                                                                                |
| CARGO        | : The Cardiac Allograft Rejection Gene Expression Observation (L'observation de l'expression du gène de rejet de l'allogreffe cardiaque) |
| CD           | : Cyclodextrin                                                                                                                           |
| ERS          | : Evaluation standard des risques pour la santé                                                                                          |
| FDA          | : The US Food and Drug Administration                                                                                                    |
| GWAS         | : Genome-wide association studies (études d'association à l'échelle du génome)                                                           |
| HGP          | : Human Genome Project (projet du génome humain)                                                                                         |
| HSF          | : Historique de santé familiale                                                                                                          |
| IMAGE        | : Invasive Monitoring Attenuation Through Gene Expression                                                                                |
| miRNAs       | : micrornas                                                                                                                              |
| MsPA         | : Mycobacterium smegmatis porin A                                                                                                        |
| NGS          | : Next generation sequencing (séquenceur de nouvelle génération)                                                                         |
| ONT          | : Oxford Nanopore Technologies                                                                                                           |

|        |                                                                                  |
|--------|----------------------------------------------------------------------------------|
| PBMC   | : Peripheral blood mononuclear cell (Cellule mononucléaire du sang périphérique) |
| PheWAS | : Phenome-wide association studies (Etudes d'association à l'échelle du phenome) |
| phi29  | : Bacillus phage phi29                                                           |
| READNA | : The REvolutionary Approaches and Devices for Nucleic Acid analysis             |
| REC    | : Recherche sur l'efficacité comparative                                         |
| SCA    | : Syndromes coronariens aigus                                                    |
| siRNAs | : Small interfering RNAs                                                         |
| SNP    | : Single-nucleotide polymorphisms                                                |
| SSD    | : Systèmes de soutien à la décision clinique                                     |
| UE     | : Union européenne                                                               |
| WGS    | : Whole genome sequencing (le séquençage complet du génome)                      |



# *Sommaire*

|                                                                                    |           |
|------------------------------------------------------------------------------------|-----------|
| <b>I. Introduction .....</b>                                                       | <b>2</b>  |
| <b>II. Rappel.....</b>                                                             | <b>6</b>  |
| 1. Définition.....                                                                 | 7         |
| 2. Propriétés .....                                                                | 10        |
| 3. Fonctions biologiques .....                                                     | 13        |
| 4. Gènes et génomes.....                                                           | 14        |
| 5. Transcription et traduction .....                                               | 15        |
| 6. Réplication .....                                                               | 17        |
| 7. Histoire de la recherche sur l'ADN .....                                        | 17        |
| <b>III. Le séquençage de l'ADN.....</b>                                            | <b>21</b> |
| 1. Caractéristiques des générations de séquençage de l'ADN .....                   | 24        |
| 1.1. Séquençage de 1ère génération.....                                            | 24        |
| 1.2. Séquençage de 2ème génération.....                                            | 25        |
| 1.3. Séquençage de 2.5 génération.....                                             | 26        |
| 1.4. Séquençage de 3ème génération.....                                            | 27        |
| 1.5. Séquençage de 4ème génération.....                                            | 27        |
| 2. Dispositifs de séquençage d'ADN de 2ème et 2.5 génération .....                 | 28        |
| 2.1. Dispositifs de séquençage d'ADN de 2ème génération qui ont réussi ..          | 28        |
| 2.2. Dispositifs de séquençage d'ADN de 2e génération qui n'ont jamais réussi..... | 29        |

|                                                                                                                                                             |           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 2.3. Méthodes de détection innovantes .....                                                                                                                 | 29        |
| 2.4. L'analyse des données .....                                                                                                                            | 31        |
| 3. Séquençage de 3ème génération – nano pores .....                                                                                                         | 32        |
| 3.1. Introduction aux technologies nano pores :.....                                                                                                        | 33        |
| 3.2. Acteurs commerciaux clés du séquençage nano poreux.....                                                                                                | 36        |
| 4. READNA et séquençage d'ADN .....                                                                                                                         | 37        |
| 4.1. Aperçu de READNA.....                                                                                                                                  | 37        |
| 4.2. Séquençage de nano pores d'Oxford (séquençage d'exonucléases nano poreuses et séquençage de brins de nano pores) - Séquençage de 3ème génération ..... | 38        |
| 4.2.1. Séquençage d'exonucléase de nanopores .....                                                                                                          | 38        |
| 4.2.2. Séquençage de brin de nano pores .....                                                                                                               | 40        |
| 4.3. Séquençage in situ (séquençage des cellules individuelles) - Séquençage de 4ème génération .....                                                       | 41        |
| 4.4. Manipulation de chromosomes complets, techniques de cartographie optique .....                                                                         | 41        |
| <b>IV. La médecine personnalisée .....</b>                                                                                                                  | <b>44</b> |
| 1. Qu'est-ce que la médecine personnalisée ?.....                                                                                                           | 45        |
| 1.1. La Médecine Personnalisée et l'Histoire de santé familiale .....                                                                                       | 45        |
| 1.2. Évaluation du risque de maladie chronique .....                                                                                                        | 47        |
| 1.3. Système de soutien à la décision clinique .....                                                                                                        | 48        |

|                                                                   |    |
|-------------------------------------------------------------------|----|
| 1.4. La médecine personnalisée et le génome humain .....          | 48 |
| 2. La boîte à outils du génome humain .....                       | 50 |
| 2.1. Variation à l'échelle du génome .....                        | 51 |
| 2.2. Transcriptomique .....                                       | 54 |
| 2.3. Protéomique .....                                            | 56 |
| 2.4. Métabolomique .....                                          | 57 |
| 2.5. Épigénomique .....                                           | 57 |
| 2.6. Biologie des systèmes et médecine des systèmes.....          | 58 |
| 3. De l'information de génome à la médecine personnalisée .....   | 59 |
| 3.1. Tests préconceptionnels et dépistage des hétérozygotes ..... | 59 |
| 3.2. Diagnostic prénatal et préimplantatoire .....                | 61 |
| 3.3. Prédiction du risque.....                                    | 63 |
| 3.4. Diagnostic de la maladie et caractérisation moléculaire..... | 64 |
| 3.5. Pronostic de la maladie.....                                 | 66 |
| 3.6. Pharmacogénomique .....                                      | 68 |
| 3.7. Surveillance de la réponse des maladies à la thérapie .....  | 70 |
| 4. Génomes interactives et microbiomes .....                      | 71 |
| 4.1. Génomes pathogènes.....                                      | 72 |
| 4.2. Le microbiome humain .....                                   | 73 |
| 5. Intégration de la génomique dans la pratique clinique .....    | 76 |
| 5.1. Education .....                                              | 77 |

|                                                            |           |
|------------------------------------------------------------|-----------|
| 5.2. Système de soutien à la décision clinique .....       | 77        |
| 5.3. Vie privée.....                                       | 78        |
| 5.4. Politique réglementaire.....                          | 79        |
| 5.5. Normes.....                                           | 79        |
| 5.6. Recherche sur l'efficacité comparative .....          | 81        |
| 5.7. Propriété intellectuelle.....                         | 81        |
| 5.8. Remboursement .....                                   | 82        |
| 6. Quand la "médecine personnalisée" est "médecine "?..... | 83        |
| <b>V. Conclusion.....</b>                                  | <b>86</b> |
| <b>VI. Résumé .....</b>                                    | <b>90</b> |
| <b>VII. bibliographie .....</b>                            | <b>94</b> |



" Le fait que j'ai un risque génétique pour la maladie d'Alzheimer et la cécité n'est pas une bonne nouvelle. Mais la réalité est que chacun d'entre nous aura des dizaines de ces risques, et ce que nous devons apprendre est de savoir comment les traiter. "

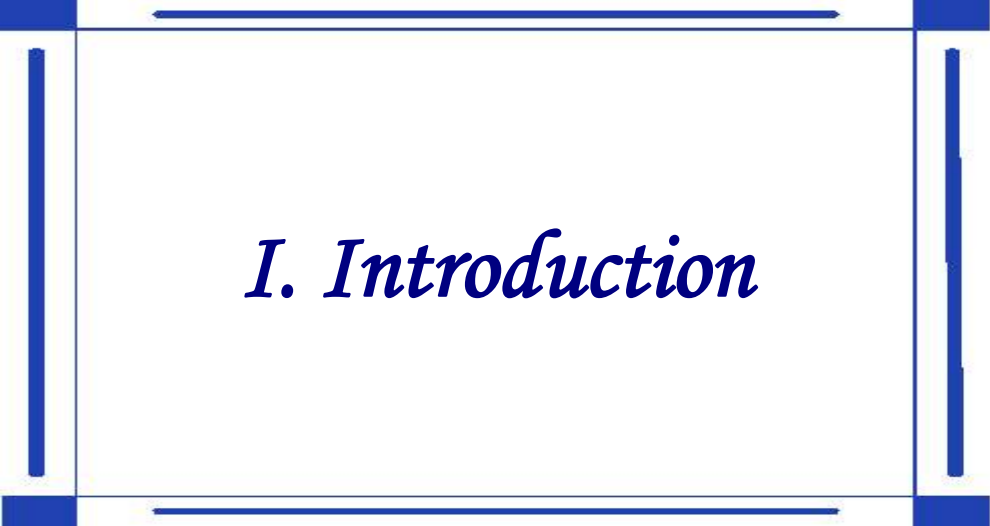
**Craig Venter**

"Les individus ne sont pas des choses stables, ils sont éphémères. Les chromosomes aussi sont mélangés dans l'oubli, comme des mains de cartes peu après qu'ils soient distribués. Mais les cartes elles-mêmes survivent au brassage. Les cartes sont les gènes. Les gènes ne sont pas détruits par la traversée, ils changent simplement de partenaires et continuent leur chemin. Bien sûr, ils continuent. C'est leur affaire. Ils sont les répliqueurs et nous sommes leurs machines de survie. Quand nous avons servi notre but, nous sommes mis de côté. Mais les gènes sont des habitants du temps géologique: les gènes sont éternels. "

**Richard Dawkins, Le gène égoïste**

" Au cours de la dernière décennie, nous avons débloqué de nombreux mystères concernant l'ADN et l'ARN. Cette connaissance ne se limite pas aux livres sur les étagères et ne se limite pas aux établis des laboratoires. Nous avons utilisé ces résultats de recherche pour identifier les causes de nombreuses maladies. De plus, les scientifiques ont traduit ces connaissances génétiques en plusieurs traitements et thérapies, ce qui a permis de créer un pont entre le laboratoire et le patient. ”

**Le président Barack Obama** Sur la Loi sur la génomique et la médecine personnalisée (S.976), le 23 mars 2007



# *I. Introduction*

La génomique est l'étude du matériel génétique complet, y compris les gènes et leurs fonctions, de l'organisme. Il se concentre sur l'obtention et l'analyse des données du génome, c'est-à-dire, les données au niveau de l'ADN en essayant de découvrir les modèles spécifiques d'expressions génétiques ou de molécules biologiques les dits biomarqueurs (molécule de signature et marqueur moléculaire) qui différencient l'état maladie de l'état normal et également déterminer la variabilité dans la réponse des patients aux médicaments .La génomique transforme la pratique médicale par l'individualisation des traitements médicaux en disposant des données génétiques des patients. La médecine est individuelle ; Nous ne sommes pas tous les mêmes et nos variations génétique détermine comment une personne réagira à la thérapie. Ces variations dans la séquence d'un gène peuvent entraîner des différences dans les enzymes qui métabolisent les médicaments. C'est ce qui explique pourquoi les enzymes apparaissent sous différentes formes chez les individus. C'est aussi pourquoi des personnes différentes métabolisent un ou le même médicament différemment. Pourquoi une drogue fonctionne efficacement pour certaines personnes, mais pas pour d'autres ? Pourquoi quelqu'un requiert-il deux fois la quantité régulièrement prescrite des mêmes médicaments pour fonctionner efficacement ? Pourquoi certaines personnes ressentent des effets secondaires des médicaments, mais d'autres ne le font pas ? Pourquoi certaines personnes vont être affectés par des cancers /n'importe quelle maladie, mais d'autres pas. Voici les raisons pour lesquelles la médecine personnalisée et individualisée est importante pour tous.

Le séquençage des acides nucléiques est le pilier de la recherche biologique. Il existe plusieurs générations de technologies de séquençage de l'ADN qui peuvent être bien caractérisées par leur nature et le type de production qu'elles fournissent. Le séquençage des terminateurs didésoxy développé par Sanger a dominé pendant 30 ans et a été le cheval de bataille du projet sur le génome humain. En 2005, le premier séquenceur de 2e génération a été présenté avec des ordres de sortie beaucoup plus élevés que le séquençage de Sanger et une réduction spectaculaire des coûts. Nous sommes maintenant à l'aube de la 3ème génération avec des systèmes nano pores qui sont en cours de développement pour le séquençage de l'ADN.

Au cours de la dernière décennie, de hauts fonctionnaires, des dirigeants de l'industrie, des fournisseurs de soins de santé et le public ont adopté la médecine génomique et personnalisée. La médecine génomique, qui est l'utilisation de l'information provenant des génomes et de leurs dérivés (ARN, protéines et métabolites) pour guider la prise de décision médicale, est un élément clé de la médecine personnalisée, qui est un domaine de soins de santé en évolution rapide, qui s'appuie sur les informations cliniques, génétiques, génomiques et environnementales uniques à chaque personne. Comme la médecine commence à utiliser des outils génomiques qui permettent une prédiction et un traitement plus précis, incluant l'interrogation du génome entier sur la variation de séquence, la transcription, les protéines et les métabolites, les principes de la médecine génomique et personnalisée nécessiteront le développement, la standardisation, et l'intégration de plusieurs outils importants dans les systèmes de santé et les flux de travail cliniques. Ces outils comprennent l'évaluation des risques pour la santé, les antécédents médicaux de la famille et l'aide à la

décision clinique pour les risques complexes et les informations prédictives. Conjugués à l'information génomique, ces outils permettront un changement de paradigme vers une approche globale qui identifiera les risques individuels et guidera la prise en charge clinique et la prise de décision, qui formeront la base d'une approche plus efficace et éclairée des soins aux patients. L'évaluation des risques liés à l'ADN pour la maladie complexe commune, les signatures moléculaires pour le diagnostic et le pronostic du cancer, la thérapie guidée par le génome et la sélection des doses sont parmi les exemples les plus importants. En outre, les informations provenant de génomes individuels, qui constituent un domaine en évolution rapide dans le domaine du développement technologique, engendrent une révolution sociale et de l'information parmi les consommateurs qui affectera sans aucun doute la prise de décision en matière de soins de santé. Bien que ces découvertes scientifiques et d'autres fassent leur chemin du génome à la clinique, L'application intégrale de la médecine génomique et personnalisée dans les soins de santé exigera des changements radicaux dans les politiques de réglementation et de remboursement, ainsi que des protections législatives pour la protection de la vie privée à l'échelle du système. Ainsi, il existe des défis à la fois scientifiques et politiques pour les soins de santé personnalisés; Cependant, ils seront confrontés et résolus avec la certitude que la science derrière la médecine génomique est solide et la pratique de la médecine qu'elle informe est basée sur des preuves.

L'objectif principal de cette étude est d'examiner les applications et le rôle de la génomique dans la médecine personnalisée.



## *II. Rappel*

## 1. Définition

L'acide désoxyribonucléique ( l'ADN ) (1) est une molécule qui porte les instructions génétiques utilisées dans la croissance, le développement, le fonctionnement et la reproduction de tous les organismes vivants connus et de nombreux virus. L'ADN et l'acide ribonucléique (ARN) sont des acides nucléiques ; aux côtés des protéines , des lipides et des glucides complexes ( polysaccharides ), ils constituent l'un des quatre principaux types de macromolécules essentiels à toutes les formes de vie connues . La plupart des molécules d'ADN sont constituées de deux brins de bio polymère enroulés l'un autour de l'autre pour former une double hélice .

Les deux brins d'ADN sont appelés poly nucléotides car ils sont composés d'unités monomères plus simples appelées nucléotides (2,3). Chaque nucléotide est composé de l'un des quatre nucléobases contenant de l'azote - cytosine (C), guanine (G), adénine (A) ou thymine (T) - un sucre appelé désoxyribose et un groupe phosphate . Les nucléotides sont joints les uns aux autres dans une chaîne par des liaisons covalentes entre le sucre d'un nucléotide et le phosphate du suivant, résultant en une alternance épine dorsale de sucre-phosphate . Les bases azotées des deux brins poly nucléotidiques séparés sont liées ensemble, selon les règles d'appariement des bases (A avec T et C avec G), avec des liaisons hydrogène pour former de l'ADN double brin. La quantité totale de paires de bases d'ADN apparentées sur Terre est estimée à  $5,0 \times 10^{37}$  et pèse 50 milliards de tonnes (4). En comparaison, la masse totale de la biosphère a été estimée à 4 billions de tonnes de carbone (TtC) (5).

L'ADN stocke des informations biologiques. Le squelette de l'ADN est résistant au clivage, et les deux brins de la structure à double brin stockent la même information biologique. Cette information est répliquée au fur et à mesure que les deux brins se séparent. Une grande partie de l'ADN (plus de 98% pour les humains) est non codante, ce qui signifie que ces sections ne servent pas de modèles pour les séquences protéiques.

Les deux brins d'ADN vont dans des directions opposées l'un à l'autre et sont donc antiparallèles. Attaché à chaque sucre est l'un des quatre types de nucléobases (informellement, *bases*). C'est la séquence de ces quatre nucléobases le long du squelette qui code l'information biologique. Les brins d'ARN sont créés en utilisant des brins d'ADN comme matrice dans un processus appelé transcription. Sous le code génétique, ces brins d'ARN sont traduits pour spécifier la séquence des acides aminés dans les protéines dans un processus appelé traduction.

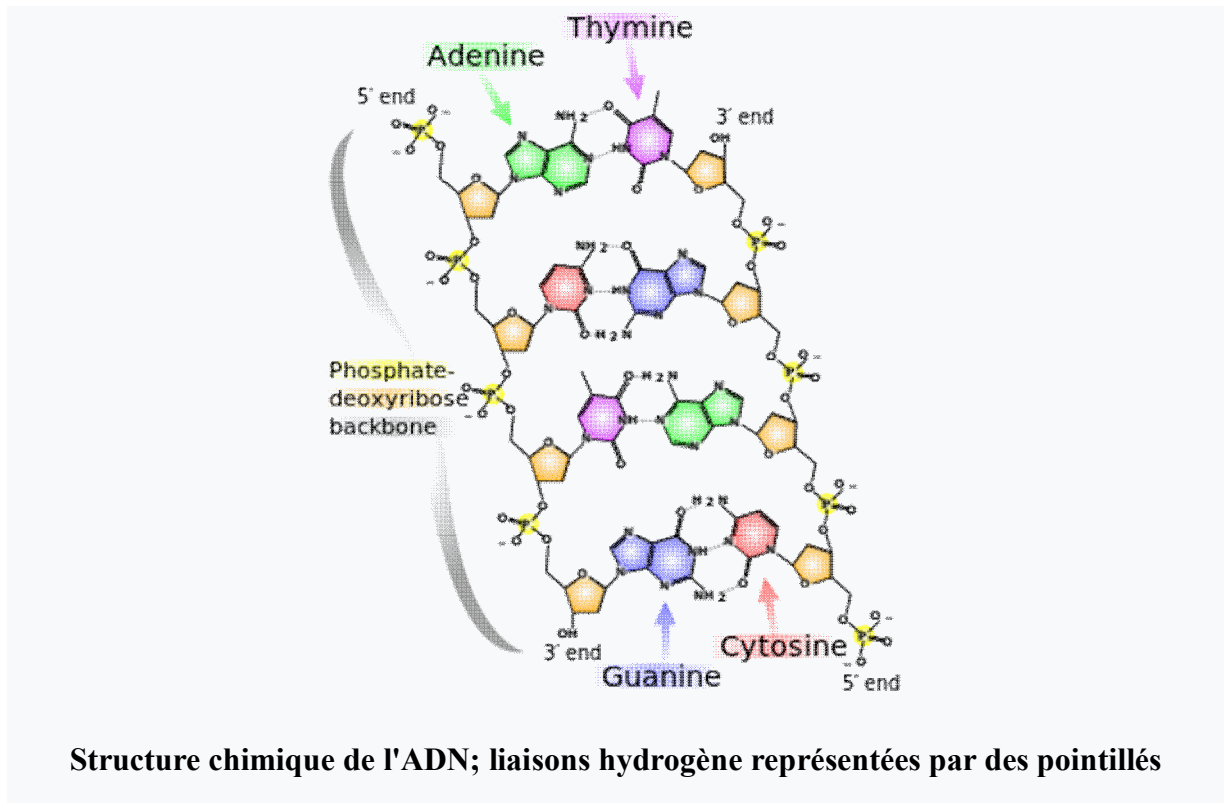
Au sein des cellules eucaryotes, l'ADN est organisé en longues structures appelées chromosomes. Au cours de la division cellulaire, ces chromosomes sont dupliqués dans le processus de réplication de l'ADN, fournissant à chaque cellule son propre ensemble complet de chromosomes. Les organismes eucaryotes (animaux, plantes, champignons et protistes) stockent la plus grande partie de leur ADN à l'intérieur du noyau cellulaire et une partie de leur ADN dans des organites, tels que les mitochondries ou les chloroplastes (6). En revanche, les procaryotes (bactéries et archaea) ne stockent leur ADN que dans le cytoplasme. Au sein des chromosomes eucaryotes, les protéines de la chromatine telles que les histones se concentrent et organisent l'ADN. Ces



structures compactes guident les interactions entre l'ADN et les autres protéines, aidant à contrôler quelles parties de l'ADN sont transcrites.

L'ADN a été isolé par Friedrich Miescher en 1869. Sa structure moléculaire a été identifiée par James Watson et Francis Crick au Laboratoire Cavendish de l' Université de Cambridge en 1953, dont les efforts de modélisation ont été guidés par les données de diffraction des rayons X acquises par Raymond Gosling , qui était un étudiant de troisième cycle de Rosalind Franklin . L'ADN est utilisé par les chercheurs comme un outil moléculaire pour explorer les lois physiques et les théories, telles que le théorème ergodique et la théorie de l'élasticité. Les propriétés matérielles uniques de l'ADN en ont fait une molécule attrayante pour les scientifiques et les ingénieurs intéressés par la micro et la nano fabrication. Parmi les avancées notables dans ce domaine, citons l'ADN origami et les matériaux hybrides à base d'ADN (7).

## 2. Propriétés



L'ADN est un long polymère fabriqué à partir d'unités répétées appelées nucléotides (8,9). La structure de l'ADN est dynamique le long de sa longueur, étant capable de s'enrouler dans des boucles serrées, et d'autres formes (10). Dans toutes les espèces il est composé de deux chaînes hélicoïdales, liées les unes aux autres par des liaisons d'hydrogène. Les deux chaînes sont enroulées autour du même axe et ont la même hauteur de 34 angströms (3,4 nanomètres). La paire de chaînes a un rayon de 10 angströms (1,0 nanomètre) (11). Selon une autre étude, lorsqu'elle est mesurée dans une solution différente, la chaîne d'ADN mesure de 22 à 26 nm de large (2,2 à 2,6 nanomètres), et une unité nucléotidique mesure 3,3 Å (0,33 nm) de longueur (12). Bien que chaque unité

de répétition de nucléotide individuelle soit très petite, les polymères d'ADN peuvent être de très grandes molécules contenant des millions à des centaines de millions de nucléotides. Par exemple, l'ADN du plus grand chromosome humain, le chromosome numéro 1, comprend environ 220 millions de paires de bases et aurait une longueur de 85 mm s'il était redressé (13).

Dans les organismes vivants, l'ADN n'existe généralement pas comme une seule molécule, mais plutôt comme une paire de molécules qui sont maintenues étroitement ensemble (14,15). Ces deux longs brins s'entrelacent comme des vignes, dans la forme d'une double hélice. Le nucléotide contient à la fois un segment du squelette de la molécule (qui maintient la chaîne ensemble) et une nucléobase (qui interagit avec l'autre brin d'ADN dans l'hélice). Une nucléobase liée à un sucre est appelée un nucléoside et une base liée à un sucre et un ou plusieurs groupes de phosphate est appelé un nucléotide. Un polymère comprenant plusieurs nucléotides liés (comme dans l'ADN) est appelé polynucléotide (16).

Le squelette du brin d'ADN est fabriqué à partir de résidus de phosphate et de sucre alternés (17). Le sucre dans l'ADN est le 2- désoxyribose, qui est un pentose (cinq- carbone). Les sucres sont réunis par des groupes phosphate qui forment des liaisons phosphodiester entre les troisième et cinquième atomes de carbone des anneaux de sucre adjacents. Ces liaisons asymétriques signifient qu'un brin d'ADN a une direction. Dans une double hélice, la direction des nucléotides dans un brin est opposée à leur direction dans l'autre brin: les brins sont *antiparallèles*. Les extrémités asymétriques des brins d'ADN sont dites avoir une directionnalité de amorce *cinq* (5') et amorce *trois* (3'), avec l'extrémité 5' ayant un groupe phosphate terminal et l'extrémité 3' un groupe hydroxyle

terminal. Une différence majeure entre l'ADN et l'ARN est le sucre, avec le 2-désoxyribose dans l'ADN étant remplacé par l'alternative pentose sucre ribose dans l'ARN (15).



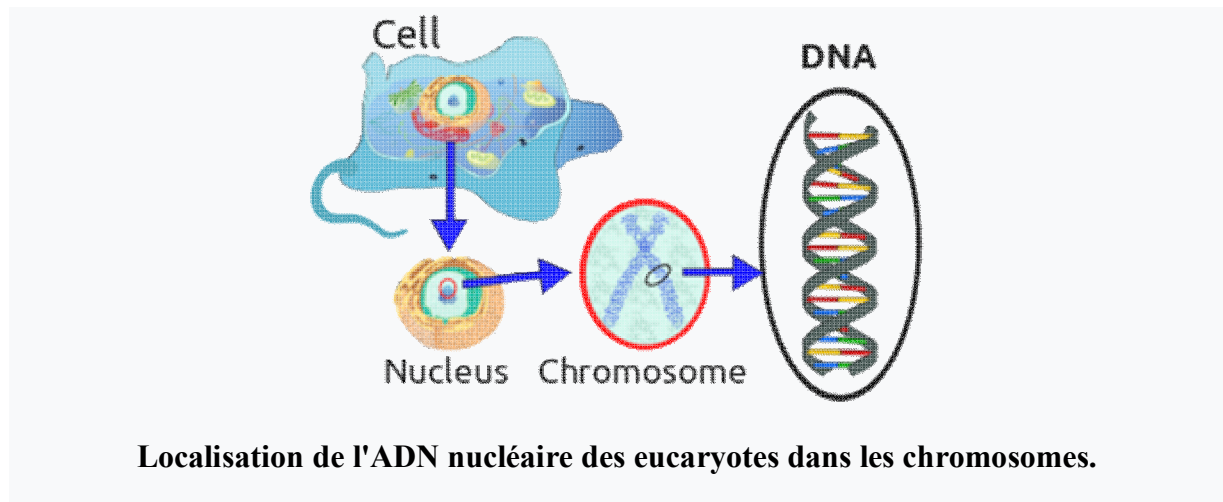
**Une section d'ADN. Les bases se trouvent horizontalement entre les deux brins en spirale (18).**

La double hélice d'ADN est stabilisée principalement par deux forces: les liaisons hydrogène entre les nucléotides et les interactions d'empilement de bases entre les nucléobases aromatiques (19). Dans l'environnement aqueux de la cellule, les liaisons  $\pi$  conjuguées des bases nucléotidiques s'alignent perpendiculairement à l'axe de la molécule d'ADN, minimisant leur interaction avec l'enveloppe de solvation. Les quatre bases trouvées dans l'ADN sont l'adénine (A), la cytosine (C), la guanine (G) et la thymine (T). Ces quatre bases sont attachées au sucre-phosphate pour former le nucléotide complet, comme indiqué pour l'adénosine mono phosphate. Des paires adénine avec de la

thymine et des paires guanine avec de la cytosine . Il était représenté par des paires de bases A-T et des paires de bases G-C (20,21).

### 3. Fonctions biologiques

---



L'ADN se présente généralement sous forme de chromosomes linéaires chez les eucaryotes et de chromosomes circulaires chez les procaryotes . L'ensemble des chromosomes dans une cellule constitue son génome ; le génome humain a environ 3 milliards de paires de bases d'ADN disposées en 46 chromosomes (22). L'information portée par l'ADN est contenue dans la séquence de morceaux d'ADN appelés gènes .La transmission de l'information génétique dans les gènes est réalisée via l'appariement de bases complémentaires. Par exemple, en transcription, lorsqu'une cellule utilise l'information dans un gène, la séquence d'ADN est copiée dans une séquence d'ARN complémentaire par l'intermédiaire de l'attraction entre l'ADN et les nucléotides d'ARN corrects. Habituellement, cette copie d'ARN est ensuite utilisée pour construire une séquence de protéine correspondante dans un processus appelé traduction , qui dépend de la même interaction entre les

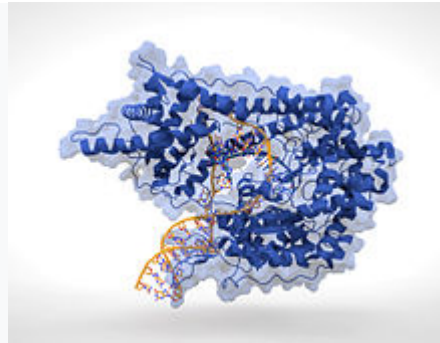
nucléotides d'ARN. En variante, une cellule peut simplement copier son information génétique dans un processus appelé réplication de l'ADN; Ici, l'accent est mis sur les interactions entre l'ADN et les autres molécules qui interviennent dans la fonction du génome.

#### **4. Gènes et génomes**

L'ADN génomique est étroitement et soigneusement emballé dans le processus appelé condensation de l'ADN , pour s'adapter aux petits volumes disponibles de la cellule. Chez les eucaryotes, l'ADN est situé dans le noyau de la cellule , avec de petites quantités dans les mitochondries et les chloroplastes . Chez les procaryotes, l'ADN est maintenu dans un corps de forme irrégulière dans le cytoplasme appelé nucléoïde (23) .L'information génétique dans un génome est contenue dans des gènes, et l'ensemble complet de cette information dans un organisme est appelé son génotype . Un gène est une unité de l'hérédité et est une région de l'ADN qui influence une caractéristique particulière dans un organisme. Les gènes contiennent un cadre de lecture ouvert qui peut être transcrit, et des séquences régulatrices telles que des promoteurs et des amplificateurs , qui contrôlent la transcription du cadre de lecture ouvert.

Dans de nombreuses espèces , seule une petite fraction de la séquence totale du génome code pour la protéine. Par exemple, seulement environ 1,5% du génome humain est constitué d'exons codant pour des protéines, avec plus de 50% d'ADN humain constitué de séquences répétitives non codantes (24). Les raisons de la présence de tant d' ADN non codant dans les génomes eucaryotes et les différences extraordinaires dans la taille du génome , ou *valeur-C* , entre les espèces, représentent un puzzle de longue date connu sous le nom « énigme valeur C » (25). Cependant, certaines séquences d'ADN qui ne codent pas les

protéines peuvent encore coder pour des molécules ARN fonctionnels non codants, qui sont impliquées dans la régulation de l'expression des gènes (26).



**T7 ARN polymérase (bleu) produisant un ARNm (vert) à partir d'une matrice d'ADN (orange) (27).**

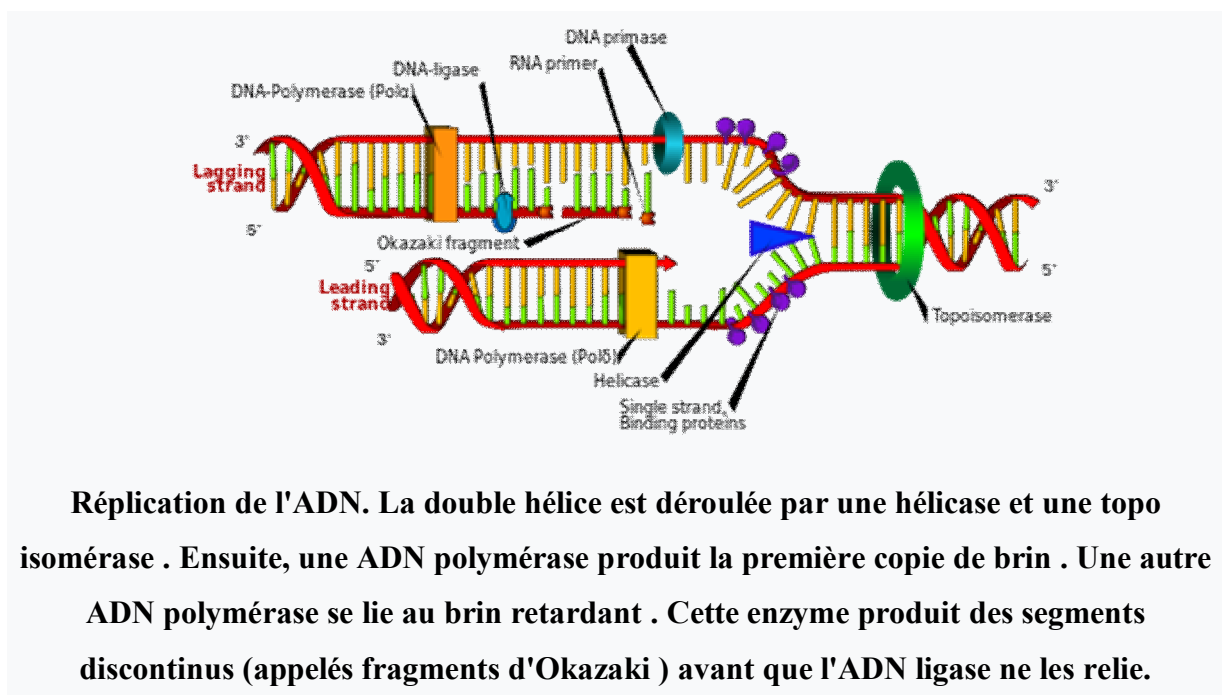
Certaines séquences d'ADN non codantes jouent un rôle structurel dans les chromosomes. Les télomères et les centromères contiennent généralement peu de gènes mais sont importants pour la fonction et la stabilité des chromosomes (28). Une forme abondante d'ADN non codant chez les humains est des pseudos gènes , qui sont des copies de gènes qui ont été désactivés par mutation (29). Ces séquences ne sont généralement que des fossiles moléculaires , bien qu'elles puissent parfois servir de matériel génétique brut pour la création de nouveaux gènes à travers le processus de duplication et de divergence des gènes (30) .

## **5. Transcription et traduction**

Un gène est une séquence d'ADN qui contient des informations génétiques et peut influencer le phénotype d'un organisme. Au sein d'un gène, la séquence de bases le long d'un brin d'ADN définit une séquence d'ARN messenger , qui définit alors une ou plusieurs séquences protéiques. La relation entre les séquences nucléotidiques des gènes et les séquences d'acides aminés des

protéines est déterminée par les règles de traduction , connues collectivement sous le nom de code génétique . Le code génétique consiste en des «mots» à trois lettres appelés *codons* formés à partir d'une séquence de trois nucléotides (par exemple, ACT, CAG, TTT).

En transcription, les codons d'un gène sont copiés dans l'ARN messenger par l'ARN polymérase . Cette copie d'ARN est ensuite décodée par un ribosome qui lit la séquence d'ARN en appariant l'ARN messenger par transfert de base , ce qui porte des acides aminés. Comme il y a 4 bases en combinaisons de 3 lettres, il y a 64 codons possibles ( $4^3$  combinaisons). Ceux-ci codent pour les vingt acides aminés standards , donnant à la plupart des acides aminés plus d'un codon possible. Il existe également trois codons «stop» ou «non-sens» signifiant la fin de la région codante; ce sont les codons TAA, TGA et TAG.





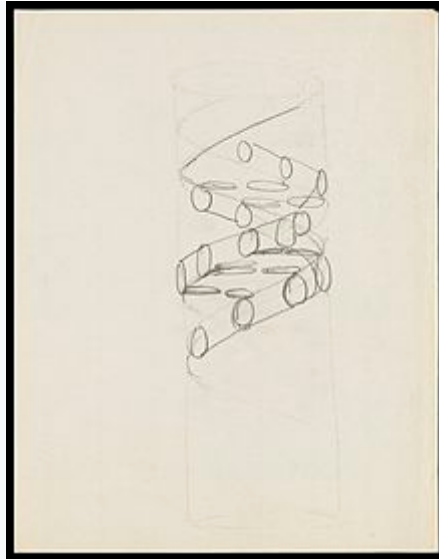
## 6. Réplication

La division cellulaire est essentielle à la croissance d'un organisme, mais, quand une cellule se divise, elle doit répliquer l'ADN dans son génome de sorte que les deux cellules filles aient la même information génétique que leur parent. La structure double-brin de l'ADN fournit un mécanisme simple pour la réplication de l'ADN. Ici, les deux brins sont séparés puis la séquence d'ADN complémentaire de chaque brin est recrée par une enzyme appelée ADN polymérase. Cette enzyme construit le brin complémentaire en trouvant la base correcte par appariement de bases complémentaires et en le liant sur le brin d'origine. Comme les ADN polymérases peuvent seulement étendre un brin d'ADN dans une direction 5 'vers 3', différents mécanismes sont utilisés pour copier les brins antiparallèles de la double hélice (31). De cette façon, la base sur le vieux brin dicte quelle base apparaît sur le nouveau brin, et la cellule finit avec une copie parfaite de son ADN.

## 7. Histoire de la recherche sur l'ADN



**James Watson et Francis Crick (à droite), co-auteurs du modèle en double hélice, avec Maclyn McCarty (à gauche).**



**Croquis au crayon de la double hélice d'ADN par Francis Crick en 1953**

L'ADN a été isolé pour la première fois par le physicien suisse Friedrich Miescher qui, en 1869, a découvert une substance microscopique dans le pus des bandages chirurgicaux jetés. Comme il résidait dans les noyaux des cellules, il l'appelait "nucléine" (32,33). En 1878, Albrecht Kossel a isolé le composant non-protéine de "nucléine", acide nucléique, et a isolé plus tard ses cinq nucléobases primaires (34,35). En 1919, Phoebus Levene a identifié la base, le sucre, et le phosphate de l'unité de nucléotide (36). Levene a suggéré que l'ADN consistait en une chaîne d'unités nucléotidiques reliées entre elles par les groupes phosphate. Levene pensait que la chaîne était courte et les bases répétées dans un ordre fixe. En 1937, William Astbury a produit les premiers modèles de diffraction des rayons X qui ont montré que l'ADN avait une structure régulière (37).

En 1927, Nikolaï Koltsov a proposé que les traits hérités soient hérités par une «molécule héréditaire géante» composée de «deux brins miroirs qui se répliqueraient de manière semi-conservatrice en utilisant chaque brin comme modèle» (38,39). En 1928, l'expérience de Frederick Griffith dans a découvert que les traits de la forme "lisse" de *Pneumocoques* pourraient être transférés à la forme "rugueuse" des mêmes bactéries en mélangeant des bactéries "lisses" tuées avec la forme "rugueux" vivant ". Ce système a fourni la première indication claire que l'ADN porte l'information génétique - l'expérience Avery-MacLeod-McCarty - quand Oswald Avery, avec ses collègues Colin MacLeod et Maclyn McCarty, a identifié l'ADN comme le principe de transformation en 1943 (40). Le rôle de l'ADN dans l' hérédité a été confirmé en 1952 lorsque Alfred Hershey et Martha Chase dans l' expérience Hershey-Chase ont montré que l'ADN est le matériau génétique du phage T2 (41) .

À la fin de 1951, Francis Crick a commencé à travailler avec James Watson au Laboratoire Cavendish de l'Université de Cambridge . En 1953, Watson et Crick ont suggéré ce qui est maintenant accepté comme le premier modèle correct de structure d'ADN en double hélice dans la revue *Nature*(11). Leur modèle moléculaire d'ADN en double hélice était alors basé sur une image de diffraction des rayons X (appelée " Photo 51 ") (42) prise par Rosalind Franklin et Raymond Gosling en mai 1952, et l'information que les bases d'ADN sont appariées. Le 28 février 1953, Crick interrompit l'heure du déjeuner de son patron au pub The Eagle à Cambridge pour annoncer que lui et Watson avaient "découvert le secret de la vie" (43).

Des preuves expérimentales soutenant le modèle de Watson et Crick ont été publiées dans une série de cinq articles dans le même numéro de *Nature* (44). Parmi ceux-ci, l'article de Franklin et Gosling fut la première publication de leurs propres données de diffraction des rayons X et de leur méthode d'analyse originale qui soutenait en partie le modèle de Watson et Crick (45,46) ; ce numéro contenait également un article sur la structure d'ADN par Maurice Wilkins et deux de ses collègues, dont l'analyse et les modèles de rayons X de l'ADN B *in vivo* ont également soutenu la présence *in vivo* des configurations d'ADN double hélice proposé par Crick et Watson pour leur modèle moléculaire double-hélice de l'ADN dans les deux pages précédentes de *Nature* (47) . En 1962, après la mort de Franklin, Watson, Crick et Wilkins ont reçu conjointement le prix Nobel de physiologie ou de médecine (48).les prix Nobel sont attribués uniquement aux bénéficiaires vivants. Un débat se poursuit sur qui devrait être crédité de la découverte (49).

Dans une présentation influente en 1957, Crick exposait le dogme central de la biologie moléculaire , qui prédisait la relation entre l'ADN, l'ARN et les protéines, et articulait «l'hypothèse de l'adaptateur» (50). La confirmation finale du mécanisme de réplication impliqué par la structure en double hélice suivie en 1958 par l'expérience Meselson-Stahl (51). D'autres travaux de Crick et ses collègues ont montré que le code génétique était basé sur des triplets de bases non-chevauchants, appelés codons, permettant à Har Gobind Khorana , Robert W. Holley et Marshall Warren Nirenberg de déchiffrer le code génétique (52). Ces résultats représentent la naissance de la biologie moléculaire (53).



### *III. Le séquençage de l'ADN*

Un document historique publié par Frederick Sanger en 1977 a présenté une technique de séquençage de l'ADN qui allait devenir l'étalon-or pour les 30 prochaines années (54). La méthode originale de séquençage de l'ADN du terminateur didéoxy a été mise en œuvre avec une électrophorèse sur gel automatisée et une chimie de terminaison fluorescente, initialement sur gel en plaque et plus tard sur des systèmes à base de gel capillaire (55). La plupart des améliorations technologiques de cette méthode ont eu lieu lors du projet du génome humain (HGP) qui a été lancé en 1990 et au cours des 13 prochaines années, dans un effort international a séquencé le premier génome humain complet à un coût proche de trois milliards de dollars (56). Simultanément, une entreprise privée concurrente a également séquencé un génome humain. En 2001, les premiers résultats de cet effort gargantuesque ont été publiés (57), et en 2004 la séquence finie a été publiée (58). Le projet du génome humain a prouvé que le séquençage complet du génome (WGS) pouvait être réalisé, mais à un coût prohibitif à effectuer de façon régulière. Un effet du HGP a été une révolution technologique et l'arrivée de ce qu'on appelle le séquençage de prochaine génération (NGS). Le but de ces technologies était de réduire le temps, l'effort et le coût de WGS à un niveau où le génome humain pourrait être séquencé de façon routinière pour la recherche et les applications cliniques. Aujourd'hui, nous pouvons distinguer au moins quatre générations avec des caractéristiques distinctes. La transition du séquençage d'ADN de première génération à la deuxième génération a été perturbante, avec une production augmentant de plus de 5 ordres de grandeur et une baisse des coûts de plus de 5 ordres de grandeur. Avec les systèmes de séquençage de 2ème génération actuels, le WGS d'un génome humain à 30\* de couverture peut être atteint en 10 jours pour un coût inférieur à 10 000 \$.

Le développement technologique continue de faire baisser le coût des WGS humains en dessous de 1000 \$. L'incitation supplémentaire de la Fondation X Prize qui s'est engagée à attribuer 10 millions de dollars à la première équipe qui séquence avec précision le génome entier de 100 sujets en 30 jours pour 1000 \$ ou moins par génome a accentué l'intensité des développements technologiques dans le domaine (<http://genomics.xprize.org/>). L'objectif central du séquençage de 2ème génération est de développer des systèmes très bon marché et très rapides avec pour objectif principal de générer de gros volumes de séquences. En raison de la nature de la chimie, une séquence du génome se compose de milliards de fragments courts entre 70 et 400 bases de long, ce qui est un inconvénient et est traité par calcul. La qualité et la quantité des données sont des problèmes majeurs car les systèmes produisent maintenant 60 G Bases de séquences par jour sur un seul séquenceur d'ADN. En conséquence, le calcul pour l'analyse des données est ce qui fait le plus peser sur le séquençage de 2ème génération. L'augmentation de la production de données ne fera qu'accentuer ce problème et de nouvelles méthodes de traitement de cette quantité de données sont en cours d'élaboration. On peut s'attendre à ce que les techniques qui seront disponibles à l'avenir fourniront une séquence avec des caractéristiques nettement différentes de celles de la 2e génération actuelle. Il est également probable que des outils supplémentaires et raffinés pour certaines applications de séquençage seront nécessaires pour compléter les données produites à l'aide de séquenceurs de deuxième génération.

Même si le séquençage de deuxième génération n'est pas une ancienne technologie, on annonce des systèmes de troisième génération qui sont à nouveau perturbateurs. L'Union européenne a réalisé que de nombreux

problèmes dans le domaine du séquençage de l'ADN doivent être résolus et un projet européen intitulé READNA (Approches et outils de conversion pour l'analyse des acides nucléiques) a été financé avec 12 millions d'euros pour développer de nouvelles technologies d'analyse des acides nucléiques. La revue présentera les technologies de 3ème et 4ème génération qui sont développées dans le cadre de ce projet innovant.

## **1. Caractéristiques des générations de séquençage de l'ADN**

Il existe au moins quatre générations de méthodes de séquençage de l'ADN qui peuvent être classées selon les principales caractéristiques distinctives (tableau 1). D'autres classifications sont utilisées, cependant, nous nous concentrons sur des caractéristiques particulières pour cette classification qui rendent la catégorisation claire.

### **1.1. Séquençage de 1ère génération**

Pour le séquençage de 1ère génération, des échelles de fragments de la séquence sont générées soit par extension enzymatique d'une amorce hybridée à un pool de molécules modèles et en introduisant des terminaisons T, C, G ou A spécifiques le long du modèle (54). Alternativement, des clivages spécifiques à la base ont été introduits dans une méthode conçue par Maxam et Gilbert (59). Les échelles à fragments sont séparées par électrophorèse sur gel. Les produits sont soit révélés par un marquage radioactif et une exposition à un film ou par l'introduction de colorants fluorescents dans la réaction de terminaison et l'imagerie par fluorescence. Ceci est maintenant réalisé au cours d'une séquence de séquençage en temps réel.



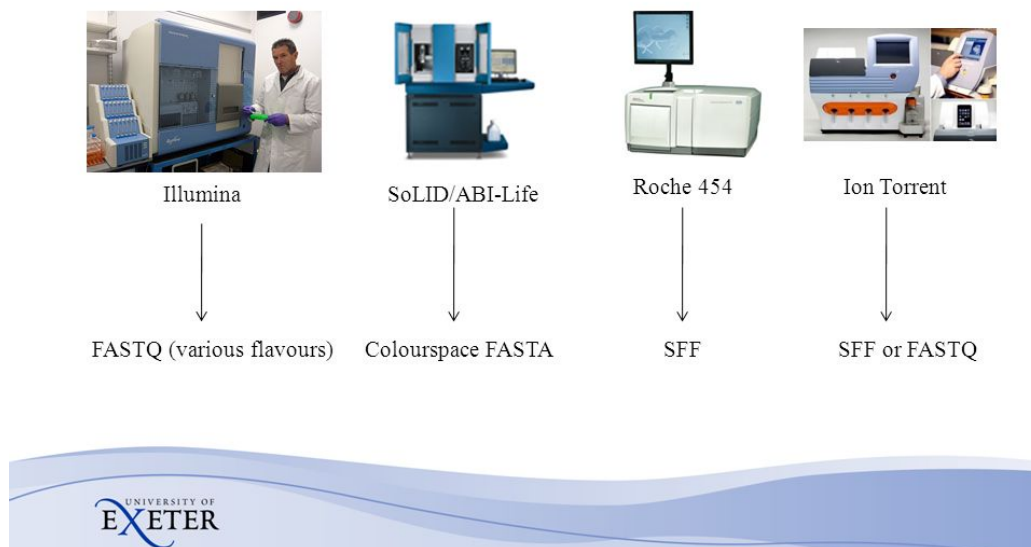


**Figure1 : Séquenceur de 1 ère génération**

## **1.2. Séquençage de 2ème génération**

Un changement de paradigme sous-tend le séquençage de 2ème génération (60). Des modèles d'ADN clonal individuels sont générés et tous séquencés en parallèle en utilisant des micros fluidiques. Les réactions de séquençage sont réalisées avec des cycles d'additions de bases et d'imagerie. Pour déterminer la séquence, des images consécutives sont superposées et des bases sont appelées aux positions de toutes les molécules modèles. Les additions de base peuvent être réalisées par polymérisation ou ligature de l'ADN et la visualisation se fait par chimiluminescence (Chimie de pyroséquençage utilisée dans l'instrument Roche FLX), fluorescence (Illumina et Life Technologies 5500) ou avec des changements de pH (dans l'instrument Ion Torrent). Les caractéristiques clés du séquençage de 2ème génération sont l'utilisation de nombreux modèles clonaux en parallèle et un processus de détermination de séquence utilisant la réplication enzymatique.

## 2<sup>nd</sup> generation sequencing output formats



**Figure2 : Séquenceurs de 2ème génération**

### 1.3. Séquençage de 2.5 génération

Le séquenceur SMRT de Pacific Biosciences est parfois appelé séquenceur de 3ème génération, mais il satisfait aux mêmes caractéristiques que les instruments de séquençage de 2ème génération en ce que les molécules clonales individuelles sont séquencées en utilisant un système de réplication de modèle enzymatique (61). Cependant, il rompt avec le concept de réseau aléatoire des systèmes de «deuxième génération» «traditionnels». Le système enzymatique polymérase est positionné au fond des puits des guides d'ondes en mode zéro (62). Les molécules d'ADN modèle sont capturées et l'incorporation de base est

surveillée par clivage du pyrophosphate lié au colorant fluorescent dans la fenêtre d'observation limitée en volume.



**Figure3 : Le séquenceur SMRT de Pacific Biosciences**

### **1.4. Séquençage de 3ème génération**

La troisième génération est définie par la lecture directe des molécules individuelles. Aucun système enzymatique de réplication n'est utilisé pour identifier la séquence. Les séquenceurs nano pores (décrits ci-dessous) en sont des exemples.

### **1.5. Séquençage de 4ème génération**

La quatrième génération est une méthode qui est actuellement encore très expérimentale. Il définit la réalisation d'une expérience de séquençage en contexte, telle que l'ADN de cellules individuelles dans une coupe histologique, donc dans leur contexte. Les applications de cette méthode seraient l'interrogation de séquences d'ADN définies telles que celles qui sont susceptibles d'avoir subi une mutation somatique, plutôt que la génération de séquences complètes de chaque cellule.

## **2. Dispositifs de séquençage d'ADN de 2ème et 2.5 génération (63)**

### **2.1. Dispositifs de séquençage d'ADN de 2ème génération qui ont réussi**

Le premier instrument à atteindre le marché était de 454 Corporation (64). Ce système utilise une PCR en émulsion avec des oligonucléotides attachés à des billes pour l'amplification clonale. La réaction de séquençage est réalisée en utilisant la chimie de pyroséquençage développée à l'origine par Ronaghi et al. (65-67). Pour cette méthode, les bases sont ajoutées une à la fois. La libération de pyrophosphate est convertie en luminescence. La quantité de lumière émise correspond au nombre de bases incorporées de ce type (T, C, G ou A). 454 Corporation a ensuite été achetée par Roche et le système a évolué vers la Roche GS FLX et une version miniaturisée appelée Roche GS FLX junior. Le deuxième système à atteindre le marché était le système 1G de Solexa (68). Ce système utilise la PCR en pont pour l'amplification clonale avec des oligonucléotides greffés en surface et l'extension d'une seule base d'amorce (également appelée SBE) pour le séquençage. Solexa a fusionné avec Illumina et le système a évolué vers l'Illumina HiSeq 2500, avec lequel un génome humain peut être séquençé à 30 \* en un peu plus d'une journée. Récemment, Illumina a sorti l'instrument de table MiSeq qui offre des temps d'exécution plus courts pour des volumes de données plus faibles. Le troisième système, le SOLiD a été publié par Applied Biosystems en 2007 (69). Dans ce système, la PCR en émulsion est combinée avec la ligature oligonucléotidique pour la détermination de la séquence. Après une fusion avec Invitrogen, le système est devenu le Life Technologies 5500. Life Technologies a récemment acquis Ion Torrent qui a

développé un système de séquençage de 2<sup>e</sup> génération avec un système de détection innovant (voir ci-dessous).

## **2.2. Dispositifs de séquençage d'ADN de 2<sup>e</sup> génération qui n'ont jamais réussi**

Il y a plusieurs systèmes de séquençage d'ADN de 2<sup>ème</sup> génération qui n'ont pas réussi. Le séquençage des polynômes a été implémenté dans un dispositif " open source " appelé le Polonator (70). Les concepts sous-jacents sont similaires à la ligation oligonucléotidique du système SOLiD utilisée pour la réaction de séquençage (71). L'Heliscope d'Helicos était un dispositif qui n'appliquait pas l'amplification clonale, mais capturait seulement les molécules d'ADN individuelles à la surface d'une lame de verre (72). Le séquençage a été réalisé par une extension d'amorce à base unique. Pour contourner les incorporations de base erronées qui entraînent des erreurs de séquençage si des molécules isolées sont analysées, le cycle de séquençage complet est répété, générant ainsi deux fois la séquence pour chaque molécule d'ADN matrice. Les différences entre les deux exécutions sont des erreurs de séquençage qui ne sont pas utilisées pour le traitement en aval. Tous ces systèmes utilisent la détection optique du processus de séquençage avec un système d'imagerie à haute résolution, suivi par un traitement d'image pour déterminer toutes les séquences individuelles d'une série.

## **2.3. Méthodes de détection innovantes**

Récemment, des systèmes de séquençage d'ADN de deuxième génération sont apparus utilisant des systèmes de détection alternatifs. IonTorrent utilise une incorporation de base enzymatique, une à la fois, similaire au système

Roche FLX. Les incorporations de base réussies sont déterminées par les changements de pH dus à la libération de  $H^+$  pendant l'incorporation de la base. Un microsystème avec plusieurs capteurs de pH en parallèle est utilisé pour détecter les bases. Complete Genomics utilise un schéma d'amplification aléatoire pour la préparation des échantillons. Les matrices de clones sont capturées avec un système sonde-ancrage et les séquences sont déterminées en utilisant la ligature oligonucléotidique (73). Complete Genomics ne vend pas ses systèmes mais les utilise exclusivement pour effectuer le séquençage de service. Pacific Biosciences a développé une technologie qui va au-delà du séquençage de 2ème génération et de ce que nous avons appelé la génération 2.5. Dans cette technologie, une ADN polymérase est greffée au fond d'un micro puits. La détection se fait par le bas en tirant parti d'un principe physique quantique. Le diamètre du micro puits ne permet pas à la lumière de pénétrer profondément dans le puits (guides d'onde de mode zéro) et limite ainsi le volume d'irradiation (62). L'entrée de la lumière dans le puits est suffisamment profonde pour exciter le système de séquençage, mais pas les autres composants (fluorescents) qui sont en solution. Une molécule d'ADN modèle est capturée par l'ADN polymérase attachée au fond du puits. L'incorporation des nucléotides est suivie par la fluorescence suivante qui est liée au pyrophosphate qui est clivé du nucléotide dans la réaction de polymérisation. L'incorporation est suivie en temps réel. Cette technologie permet de séquencer de très longues bandes d'ADN. Des longueurs de lecture de plusieurs kb ont été rapportées. Il y a quelques années, Life Technologies a annoncé qu'elle développait une technologie similaire, sauf qu'elle n'utilisait pas les guides d'onde à mode zéro, mais qu'elle utilisait plutôt des nano cristaux et un système FRET. Cependant, ce système très intéressant n'a pas encore atteint le marché.

## **2.4. L'analyse des données**

### Défis d'analyse de données associés à la 2e génération de séquençage

Les défis d'analyse de données associés aux résultats de séquençage de deuxième génération ont trois composantes principales: (1) Appels de base, (2) Alignement et (3) Appel de variantes. Pour la plupart des systèmes, l'appel de base est étroitement lié au système de séquençage et est effectué à l'aide d'un logiciel fourni par le fournisseur du dispositif de séquençage. L'alignement est une stratégie qui est choisie lorsque l'échantillon en cours d'examen a une séquence de référence. Les séquences produites sont comparées à la séquence de référence et la position la plus probable dans le génome pour une lecture donnée est déterminée. Les séquences qui se lient uniquement au génome et ne présentent pas de discordances par rapport à la référence sont les plus faciles à positionner. Évidemment, les séquences qui diffèrent d'un degré acceptable de la référence sont les plus intéressantes car elles sont utilisées pour identifier la variabilité. Il y a une batterie de différents outils qui ont été développés pour faire l'alignement, pour entrer dans les détails des avantages et des inconvénients peut exploser cette revue. Ce qu'il faut dire, c'est qu'il y a quatre critères d'équilibrage dans un aligneur, (a) est-ce qu'un aligneur est déterministe ou non, (b) est-il exhaustif, (c) quel pourcentage de lectures est-il capable de mapper à une référence? (d) quel effort de calcul est requis pour le faire. L'appel de variantes est intimement lié à l'alignement. Plus l'alignement est fiable, plus la tâche de l'appel de variante est facile. Encore une fois, il existe plusieurs variantes appelants disponibles et les détailler irait trop loin. De nos jours, la plupart des aligneurs parviennent assez bien à identifier des variants de nucléotides simples et de petits indels. Cependant, l'identification d'indels plus

importants et de variantes structurelles est plus exigeante et les logiciels pour ce faire sont encore soumis à de nombreux changements. Il est crucial, en termes d'effort de calcul, que l'étape d'alignement soit celle qui nécessite le plus de ressources de calcul.

Les stratégies d'alignement reposent sur l'hypothèse que la position la plus proche de la référence est l'endroit correct pour une séquence lue. Dans de nombreux cas, cela sera correct. Cependant, il existe une probabilité résiduelle que cela puisse être faux et que la position correcte soit à un endroit différent dans le génome où la correspondance est plus éloignée. Evidemment, cela influence les résultats. Une stratégie alternative à l'alignement est l'assemblage de séquence de novo. Plusieurs logiciels pour l'assemblage de novo à partir de lectures courtes de séquenceurs de 2ème génération ont été présentés. Cependant, ces approches nécessitent des architectures matérielles qui ne sont pas des produits et du temps de calcul, une sophistication requise et des efforts considérables. Beaucoup de développement est en cours dans ce domaine. Des stratégies utilisant une combinaison d'alignement et d'assemblage de novo pourraient également être bénéfiques. Dans tous les cas, lorsque la troisième génération deviendra disponible avec différentes caractéristiques de lecture, les stratégies d'analyse des données devront être adaptées. Ils ne seront pas nécessairement plus exigeants que pour le séquençage de 2ème génération.

### **3. Séquençage de 3ème génération – nano pores**

Les seuls systèmes notables répondant aux critères que nous avons définis ci-dessus pour le séquençage des acides nucléiques de 3ème génération sont les nano pores. Les méthodes montrées à ce jour mesurent directement la molécule



d'ADN sans avoir besoin de la soumettre à un système de réplication enzymatique.



**Figure 4 : Séquenceur ONT**

### **3.1. Introduction aux technologies nano pores :**

Le principe général du séquençage d'ADN à base de nano pores est extrêmement simple et a été démontré pour la première fois en 1996 (74). Essentiellement, le système consiste en un nano pore incorporé dans une membrane (biologique ou synthétique) immergée dans une solution saline. Un courant est appliqué au système qui entraîne les ions à travers le nano pore. L'arrivée d'une biomolécule chargée telle qu'une base d'ADN créera une résistance dans le flux d'ions qui conduira à des changements dans le courant électrique qui peut être mesuré. Des molécules de taille différente bloquent le nano pore par un degré différent et pour un temps différent (75). Si les quatre bases d'ADN (T, C, G et A) peuvent être distinguées lorsqu'elles traversent le nano pore, cela permettrait de mesurer une séquence d'ADN. En substance, le principe mis en évidence ci-dessus pointe vers une méthode capable de lectures d'ADN à lecture longue ( $> 2$  kb) (76). La clé du prix du séquençage des nano

pores est la taille nanométrique de ces pores, la possibilité de parallélisation et le faible coût des consommables (77).

Actuellement, deux voies principales sont poursuivies pour le séquençage des nano pores, la première est l'utilisation de nano pores biologiques et la seconde est l'utilisation de nano pores à l'état solide (78). Ci-dessous l'accent est mis sur les nano pores biologiques que les entreprises tirant parti des nano pores biologiques envisagent de commercialiser des plates-formes de séquençage nano pores en 2012 et 2013 (Oxford Nanopore Technologies (ONT) et Genia).

Actuellement, les nano pores biologiques ont la haute main car deux articles ont été publiés qui prouvent que le séquençage de l'ADN est possible en utilisant soit un pore  $\alpha$ -hémolysine (79) soit un pore MsPA (77). L'avantage des nano pores biologiques est qu'ils ont une taille bien définie, peuvent être modifiés très facilement et produits en masse restant homogène en taille et en structure. Dans les expériences de nano pores d' $\alpha$ -hémolysine, il faut noter que pour obtenir un système entièrement fonctionnel capable de discriminer chacune des quatre bases d'ADN, le nano pore doit être génétiquement modifié et en plus une molécule adaptatrice importante pour ralentir la translocation des bases d'ADN à travers le nano pore devait être fixé de manière non covalente dans le pore (80). Essentiellement, des expériences de mutagenèse ont conclu qu'une mutation spécifique au sein du bêta-baril était nécessaire pour obtenir une discrimination de base. Il décrit que l'attachement d'un dérivé de cyclo dextrine (CD) était essentiel pour la détection continue des bases circulant à travers le nano pore artificiel (81), en particulier la séparation optimale des bases a été obtenue en attachant le CD au résidu cystéine L135C dans le canon bêta du pore (82). Les conditions pour la construction d'hémolysine / CD modifiée ont été

optimisées pour les conditions requises par l'exonucléase pour être actives tout en maintenant la capacité de discrimination de base. En fin de compte, un équilibre a été recherché entre la tension appliquée et la concentration en sel utilisée pour s'assurer que l'exonucléase fonctionne correctement et que la majorité des bases d'ADN circulent à travers le pore (pour éviter les doubles lectures de bases) à une vitesse non trop rapide pour permettre la détection de molécules individuelles. Le papier MspA utilise les mêmes concepts généraux mais tire parti du plus petit site de constriction dans le pore MspA (diamètre 1,2 nm et longueur 0,5 nm) comparé à l' $\alpha$ -hémolysine (diamètre 1,5 nm et longueur 5 nm) qui montre un meilleur potentiel de discrimination entre les bases. Un inconvénient de la méthode est que les molécules d'ADN double brin sont utilisées pour bloquer la translocation des molécules d'ADN (par rapport à l'utilisation de la molécule CD dans le travail d' $\alpha$ -hémolysine), l'ADN doit donc être converti de façon à séparer les nucléotides monocaténares doivent être mesurés, ce qui entraîne une forte préparation de pré-séquençage de l'ADN. Cependant, dans un développement récent, l'intégration de l'ADN polymérase de bactériophage (DNAP) phi29 dans le pore agit comme un moteur pour tirer le modèle monocaténaire à travers l'ouverture d'un nucléotide à la fois, permettant à un signal distinct d'être généré et chacun des nucléotides peut être mesuré (83). Un résultat similaire a été publié en utilisant le phi29 avec le pore  $\alpha$ -hémolysine et ce travail a été autorisé par l'ONT (84).

En ce qui concerne les nano pores à l'état solide, le graphène est un matériau qui pourrait être utilisé dans le futur (85), bien que dans un article crucial de Hagan Bayley, il reste encore du chemin à parcourir avant que les nano pores solides et en particulier le graphène une étape où ils peuvent être

utilisés pour séquencer régulièrement l'ADN. Il convient de mentionner que la technologie des nano pores à l'état solide produira des pores plus stables, dont la longueur et le diamètre peuvent être contrôlés, possèdent des propriétés de surface ajustables et présentent un plus grand potentiel d'intégration dans les dispositifs que des matrices biologiques (86).

À la veille de la commercialisation de deux systèmes nano pores (Oxford Nanopore Technologies et Genia Technologies, Inc.), il semble que de nombreux défis technologiques ont été résolus. Bien que les résultats attendus du séquençage de l'ADN soient la seule façon de décider si le défi des nano pores a été fissuré, il faudra encore quelques années pour avoir un système de séquençage de l'ADN qui offre le type de données requis par les chercheurs.

### **3.2. Acteurs commerciaux clés du séquençage nano poreux**

Avant de se concentrer sur READNA et son partenaire ONT, voici une courte compilation d'entreprises développant des plateformes de séquençage nano pores.

Genia Technologies Inc. \* - Séquençage de nanopores à l'aide de nano pores biologiques et contrôle électronique pour la détection du pyrophosphate marqué issu d'un nucléotide triphosphate lors de la réplication avec une polymérase fixée à l'entrée du nano pore (date de diffusion: 2013).

IBM / Roche \* - Le séquençage nano pore à l'aide de nano pores à l'état solide et le contrôle électronique déplace l'ADN à travers des nano pores intégrés dans une puce de silicium et la détection du courant de tunnel électronique (pas de date de diffusion).

NABsys Inc. \* - Le séquençage de nano pores utilisant des nano pores à l'état solide et le contrôle électronique déplace l'ADN avec des sondes oligonucléotidiques hybridées à travers des nano pores incorporés dans une puce de silicium et la détection du courant ionique (aucune date de diffusion).

NobleGen Biosciences Inc. \* - Plateforme " Optipore " - Nanopores Le séquençage à l'aide de nano pores à l'état solide et le contrôle électronique déplacent le gabarit d'ADN synthétique avec des balises oligonucléotidiques hybridées à travers des nano pores dans une puce de silicium, déplaçant les balises et détectant par fluorescence.

Oxford Nanopore Technologies \* - Le séquençage des nano pores à l'aide de nano pores biologiques et d'une enzyme de processus déplace l'ADN à travers des nano pores incorporés dans une bicouche lipidique et la détection du courant ionique (2012).

## **4. READNA et séquençage d'ADN**

### **4.1. Aperçu de READNA**

Le projet «Approches et dispositifs révolutionnaires pour l'analyse des acides nucléiques» READNA est un projet financé par le cadre de l'UE 7 qui a débuté le 1er juin 2008. Le consortium READNA comprend des projets visant à accélérer de nouvelles technologies de séquençage d'ADN révolutionnaires et des méthodes pour améliorer les méthodes d'analyse des acides nucléiques existantes. Des informations sur tous les développements de READNA peuvent être trouvées sur (<http://www.cng.fr/READNA>). Le but ultime sera de faire progresser les technologies permettant le séquençage d'un génome humain entier pour 1000 \$ en moins d'une journée. Ici, nous allons nous concentrer sur une technologie de séquençage d'ADN de 3ème génération qui a été développée par Oxford Nanopore Technologies au sein de READNA.

## **4.2. Séquençage de nano pores d'Oxford (séquençage d'exonucléases nano poreuses et séquençage de brins de nano pores) - Séquençage de 3ème génération**

Les technologies Nanopore sont développées par plusieurs partenaires READNA. Nous nous concentrerons ici sur les efforts de deux partenaires READNA, Oxford Nanopore Technologies (ONT) et l'Université d'Oxford (Professeur Hagan Bayley, également cofondateur de l'ONT).

Oxford Nanopore Technologies (ONT) poursuit deux types différents de technologie de l'ADN de troisième génération basée sur l'utilisation de nano pores pour séquencer l'ADN, à l'aide d'un nano pore d'exonucléase et d'un séquençage direct des brins. En février 2012, à l'occasion de la réunion annuelle Advances in Genome Biology & Technology à Marco Island, l'ONT a annoncé sa nouvelle plate-forme ADN " GridION " et présenté en même temps une version jetable de la même plate-forme. 'Min ION'. La plate-forme MinION a suscité beaucoup d'intérêt car lorsqu'elle sera commercialisée, elle deviendra le premier séquenceur jetable miniaturisé aussi petit qu'une clé USB et le coût sera d'environ 900 \$. L'ONT prévoit commercialiser la «première génération» de sa GridION et de sa MinION auprès d'un nombre restreint de clients.

### **4.2.1. Séquençage d'exonucléase de nanopores**

Dans la première méthode, une molécule d'ADN est capturée par une exonucléase processive qui est attachée de manière covalente à l'entrée du nanopore. L'exonucléase clive l'ADN base par base et le laisse tomber dans le nano pore. Les bases sont identifiées lorsqu'elles atteignent un confinement dans un nano pore d'a-hémolysine à une vitesse assurant une résolution de base

unique. La vitesse de translocation des bases d'ADN voyageant à travers un nano pore de 5 nm a été un problème majeur dans le passé car la résolution de base unique n'était pas réalisable. Essentiellement, un nano pore d'une longueur de 5 nm aura entre 10 et 15 nucléotides traversant le nano pore à un moment donné (76). En outre, la vitesse de translocation est de l'ordre de 1 nucléotide par ms, ce qui est trop rapide pour résoudre une base individuelle. Pour atteindre la détection de base unique, la vitesse doit être de l'ordre de 1 nucléotide / ms. Une façon de ralentir le débit est d'utiliser des adaptateurs non covalents et en particulier la cyclo dextrine (CD) a été l'adaptateur le plus efficace en combinaison avec le pore de protéine  $\alpha$ -hémolysine conçu et utilisé par ONT, un autre est l'adaptation de la concentration en sel (87). La découverte historique publiée par Clarke et al. montre que CD a été attaché à des cystéines spécifiques dans le nano pore pour réaliser l'identification de base (79). Les bases sont clivées par l'exonucléase et entrent immédiatement en contact avec l'adaptateur CD, se liant brièvement avant de se déplacer sur et à travers le nano pore en raison d'une différence de concentration en sel. La liaison temporaire crée un événement de perturbation qui est lu par une puce électronique associée au nano pore permettant de déterminer l'ordre des bases en raison d'un changement de courant. Les différences de courant sont telles que le système est capable de différencier les quatre bases d'ADN A, G, C et T et en plus le système est capable de distinguer les bases méthyl cytosine qui sont importantes dans l'analyse de la méthylation de l'ADN (79). Bien que ce soit une preuve de principe que les nano pores dans les bonnes conditions (modifications du pore, ajout d'un adaptateur, modification de la concentration en sel et contrôle de la température) puissent distinguer les quatre bases, d'autres développements seront nécessaires l'arène commerciale. De plus, le débit exigera des réseaux de

ces nano pores pour s'assurer que le séquençage des nano pores répond aux attentes actuelles d'une séquence complète du génome humain en une seule journée.

#### **4.2.2. Séquençage de brin de nano pores**

La deuxième méthode en cours de développement par ONT est le séquençage des brins de nano pores. La différence majeure avec le séquençage des exonucléases nano pores est qu'au lieu d'individualiser les bases en utilisant une enzyme exonucléase pour la détection par le nano pore, la molécule d'ADN est rendue monocaténaire et transloquée à travers le nano pore. Le problème qui doit être résolu pour que cette approche fonctionne pour le séquençage est la résolution d'une base individuelle dans le nano pore. La taille du nucléotide par rapport à la taille du nano pore résulte en plusieurs bases se trouvant dans la zone de constriction du nano pore et le courant est bloqué par un groupe de bases. Lorsque le brin d'ADN transloque à travers les nano pores, il doit être traduit en séquence. ONT ont publiquement présenté qu'ils ont réussi à utiliser cette approche pour séquencer 1 phage, en déconvoluant les changements de courant à travers le nanopore en utilisant un algorithme de Viterbi. La méthode a été privilégiée par ONT car elle est potentiellement plus rapide et plus précise car chaque base est physiquement attachée à la précédente ce qui permet de lire correctement l'ordre des bases. Le principal facteur limitant deviendra la processivité de l'enzyme utilisée, mais si suffisamment de créneaux sont disponibles, le débit dépassera certainement celui de toute autre plateforme de séquençage actuellement disponible à un coût espéré de 1000 \$ pour un génome complet. Les ONT doivent encore publier les résultats du séquençage de l'ADN, mais ont confirmé le séquençage d'un génome viral de 5 kb ayant un taux



d'erreur d'appel de base de 4%. La société vise à avoir un taux d'erreur maximum de 2% lors de la commercialisation en 2012. Une caractéristique impressionnante de cette technologie est que les lectures individuelles sont de l'ordre de 10s de Kbases. Les données de ce type révolutionnent le séquençage de l'ADN.

#### **4.3. Séquençage in situ (séquençage des cellules individuelles) - Séquençage de 4ème génération**

Le séquençage de 1ère génération a prévalu pendant près de 30 ans, la 2ème génération se poursuit depuis 7 ans, la 3ème génération est proche et la 4ème génération est déjà à l'horizon. Les cellules voisines dans le contexte de la tumeur ne sont pas identiques et il deviendra de plus en plus important de pouvoir mesurer cela. Le groupe de Mats Nilsson a fourni des exemples d'analyses de transcrits de nombreuses cellules différentes en parallèle dans leur contexte cellulaire (88). Ceci est une preuve de concept du séquençage de 4ème génération.

#### **4.4. Manipulation de chromosomes complets, techniques de cartographie optique**

En raison des procédures de préparation des échantillons utilisés dans le séquençage de deuxième génération, les informations contextuelles sont perdues. L'information haplotypique est particulièrement utile dans l'assemblage de nouveaux génomes, mais elle est également importante pour les problèmes biologiques. La présence de deux variants mutants sur des chromosomes identiques ou opposés est biologiquement pertinente. Les méthodes de prédiction des haplotypes utilisent le contexte familial ou l'inférence statistique.

Dans les molécules d'ADN réorganisées aléatoires, telles qu'elles sont rencontrées dans de nombreuses cellules tumorales, ces stratégies se décomposent. Cependant, la détermination moléculaire est possible en utilisant des méthodes telles que la cartographie optique (89). Au sein de READNA, nous avons consacré beaucoup d'efforts aux techniques de manipulation des molécules d'ADN individuelles et avons développé des systèmes micro- et nano-fluidiques pour les étirer et les rendre disponibles pour une étude plus approfondie (90).

En conclusion, le séquençage de l'ADN a atteint un point où il peut être appliqué de manière holistique. Les transitions entre les générations de méthodes de séquençage de l'ADN ont perturbé les changements dans les ordres de grandeur de la production et les changements complets dans la nature du type de séquence fourni par une méthode et la diminution spectaculaire des coûts. Alors que le séquençage d'ADN de première génération a prévalu pendant 30 ans, les dernières années ont vu une vague de changements avec la 2ème génération et déjà l'apparition de méthodes d'analyse des acides nucléiques de 3ème et 4ème génération. Chaque nouvelle génération n'a pas remplacé la précédente. Il existe un chevauchement important de différentes méthodes qui aident à compléter l'ensemble d'outils requis pour l'analyse génomique complète. La recherche ou la question clinique à portée de main détermine le dispositif optimal ou son utilisation. La deuxième génération qui est l'état de l'art actuel et la méthodologie la plus répandue, par la nature de sa séquence (beaucoup de lectures courtes) met une pression considérable sur les procédures de calcul et d'analyse des données. Au fur et à mesure que nous avançons dans le séquençage de 3ème génération, avec des lectures plus longues, les demandes de

calcul vont probablement diminuer. Cependant, en raison de l'augmentation de la production globale, le traitement bio-informatique en aval des résultats deviendra encore plus important. Les méthodes d'analyse des acides nucléiques de quatrième génération remettront en question les procédures de traitement de l'image et la résolution qu'elles peuvent atteindre.

TABLE 1

Overview of sequencing methods and characteristics

|                                           | Generation | Input          | Template enrichment           | Nature of sequencing template | Sequencing chemistry                  | Detection                              | Particularity                 | Reference                                                             |
|-------------------------------------------|------------|----------------|-------------------------------|-------------------------------|---------------------------------------|----------------------------------------|-------------------------------|-----------------------------------------------------------------------|
| Sanger sequencing                         | 1st        | DNA            | PCR                           | Pool                          | Primer extension, dideoxy termination | Radioactivity, fluorescence            | Used for Human Genome Project | Sanger [1]                                                            |
| Maxam-Gilbert                             | 1st        | DNA            | Sufficient template required  | Pool                          | Chemical base-specific cleavage       | Radioactivity                          |                               | Maxam-Gilbert [6]                                                     |
| 454-Roche                                 | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Primer extension one base at a time   | Real-time chemiluminescence            | Pyrosequencing                | Margulies [11]                                                        |
| Polonator                                 | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Oligonucleotide ligation              | Fluorescence imaging                   |                               | Shendure [17]                                                         |
| Solexa-Illumina                           | 2nd        | DNA            | Bridge PCR                    | Amplified Clone               | Single-base primer extension          | Fluorescence imaging                   |                               | Bentley [15]                                                          |
| Helicos                                   | 2nd        | DNA            | None                          | Individual DNA molecule       | Single-base primer extension          | Fluorescence imaging                   | Single molecule               | Harris [19]                                                           |
| Applied Biotechnologies-Life Technologies | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Oligonucleotide ligation              | Fluorescence imaging                   | Two-base encoding             | McKernan [39]                                                         |
| Complete Genomics                         | 2nd        | DNA            | Rolling circle replication    | Amplified Clone               | Oligonucleotide ligation              | Fluorescence imaging                   | Proprietary system            | Drmanac [20]                                                          |
| IonTorrent                                | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Primer extension one base at a time   | pH measurement in semiconductor device | Non-optical                   | Rothberg [18]                                                         |
| Pacific Biosciences                       | 2.5th      | DNA            | None                          | Individual DNA molecule       | Primer extension                      | Real-time fluorescence imaging         | Zero-mode wave guides         | Eid [9]                                                               |
| Oxford Nanopore (Exonuclease)             | 3rd        | DNA            | None                          | Individual DNA molecule       | Exonuclease digestion                 | Nanopore                               | Exonuclease nanopore          | <a href="http://www.nanoporetech.com">http://www.nanoporetech.com</a> |
| Oxford Nanopore (Strand)                  | 3rd        | DNA            | None                          | Individual DNA molecule       | Direct measurement                    | Nanopore                               |                               | <a href="http://www.nanoporetech.com">http://www.nanoporetech.com</a> |
| <i>In situ</i>                            | 4th        | Tissue section | <i>In situ</i> rolling circle | Rolling circle amplicon       | Primer extension                      | Fluorescence imaging                   |                               | Larsson [36]                                                          |

Tableau1 : Vue d'ensemble des méthodes de séquençage et des caractéristiques



## *IV. La médecine personnalisée*

## **1. Qu'est-ce que la médecine personnalisée ?**

La médecine personnalisée est un domaine vaste, à développement rapide des soins de santé fondé sur des informations uniques de chaque personne cliniques, génétiques, génomiques et environnementaux. Les soins de santé qui embrasse la médecine personnalisée est une approche intégrée, coordonnée et fondées sur des preuves pour l'individualisation des soins aux patients dans tout le processus de la état normal à la maladie .La médecine personnalisée utilise notre compréhension moléculaire de la maladie pour améliorer les stratégies de soins de santé préventives alors que les gens sont toujours bien et commencent les traitements médicamenteux dans les premiers stades de la maladie. L'objectif global de la médecine personnalisé est d'optimiser les soins médicaux et les résultats pour chaque individu, ce qui entraîne une personnalisation sans précédent des soins aux patients.

### **1.1. La Médecine Personnalisée et l'Histoire de santé familiale**

L'Histoire de santé familiale (HSF) est un outil simple mais précieux pour la fourniture de renseignements personnels sur la santé et sur les risques à certains maladies. Représentant la combinaison complexe de facteurs partagés génétiques, environnementaux, et de style de vie, un HSF robuste peut approximativement donner des renseignements sur les risques génétiques/génomiques et l'intègrent dans les soins aux patients. L'évaluation du HSF aiderait à identifier les personnes à risque plus élevé de maladie, permettant des étapes préemptives et préventives, y compris les changements de style de vie, des projections de la santé, l'essai et le traitement précoce selon le cas (voir tableau 2). Par conséquent, HSF détient un énorme potentiel pour améliorer les soins de santé préventifs parmi les populations larges et

diversifiées de façon personnellement Pertinente. Toutefois, l'évaluation et intégration de l'HSF n'ont pas été embrassé par la communauté sanitaire et demeure une ressource largement inexploitée (91). Le défi de l'intégration de l'HSF dans la santé du public implique trois composantes essentielles: (a) des méthodes de collecte standard et accessibles; (b) l'accès du fournisseur de soins de santé; et (c) des conseils cliniques pour l'interprétation et l'utilisation.

Actuellement, la collecte d'informations de l'HSF peuvent être incomplète, peuvent être difficile à interpréter, et / ou peuvent varier considérablement dans le contenu entre les fournisseurs de soins de santé du patient. En outre, les fournisseurs peuvent avoir une connaissance et une formation insuffisante afin d'interpréter fidèlement l'HSF pour déterminer le risque des syndromes génétiques héréditaires. (92)

**Tableau 2 : Risques associés au fait d'avoir un parent au premier degré atteint d'une maladie (93)**

| <b>Condition</b>     | <b>Rapport de cotes de prévalence</b> | <b>Références</b> |
|----------------------|---------------------------------------|-------------------|
| Diabète de type 2    | 2.3-2.8                               | 84, 85            |
| Maladie coronarienne | 4.5                                   | 186               |
| Cancer du côlon      | 2 – 2,6                               | 90                |
| Le cancer du sein    | 1.6 – 2                               | 90                |
| Cancer du poumon     | 1.7 – 2,5                             | 90                |

## **1.2. Évaluation du risque de maladie chronique**

Un élément fondamental de la médecine personnalisée est une évaluation standard des risques pour la santé (ERS) visant à évaluer la probabilité qu'un individu développe les maladies chroniques les plus courantes (ou les événements pathologiques). Les ERS fondées sur des données probantes couplées à des modèles prédictifs faciliteront l'évaluation du risque de maladie d'un patient et la priorisation des stratégies cliniques pour y remédier. L'un des ERS les plus largement reconnus est le modèle de maladie cardiaque coronarienne de Framingham, développé à partir de l'étude de Framingham sur le cœur débuté en 1948 (94). L'évaluation du risque de cancer du sein selon le modèle Gail et ses versions modifiées sont également des outils largement acceptés (95).

Malgré la valeur prouvée de nombreux outils ERS, ils n'ont généralement pas été inclus dans l'évaluation formelle des patients en raison du manque de normes des données cliniques requises ou les algorithmes utilisés, ainsi que le manque d'intégration dans les systèmes de technologie de l'information sur la santé (96). Cependant, les ERS devraient constituer la base de la prévision et de la stratification des risques qui sont fondamentales pour les soins de santé personnalisés. Et à mesure que de nouvelles informations biologiques et des marqueurs moléculaires prédictifs sont découverts et validés, il est essentiel que ces nouveaux marqueurs soient évalués itérativement pour leur capacité à augmenter la capacité prédictive des modèles cliniques existants. S'ils le font, ils devraient être incorporés dans la ERS et devenir le point de départ pour des soins de santé personnalisés.

### **1.3. Système de soutien à la décision clinique**

Pour optimiser l'utilisation de la HSF et des ERS, les systèmes de soutien à la décision clinique (SSD) fournissent aux médecins et aux patients des connaissances filtrées conçues pour orienter et améliorer les soins de santé sur le lieu des soins (97). Les systèmes SSD informatisés sont de plus en plus utilisés et des informations spécifiques au patient — poids, médicaments actuels, antécédents médicaux et même l'HSF — peuvent être intégrées dans des formulaires conviviaux pour aider à gérer le diagnostic et le traitement.

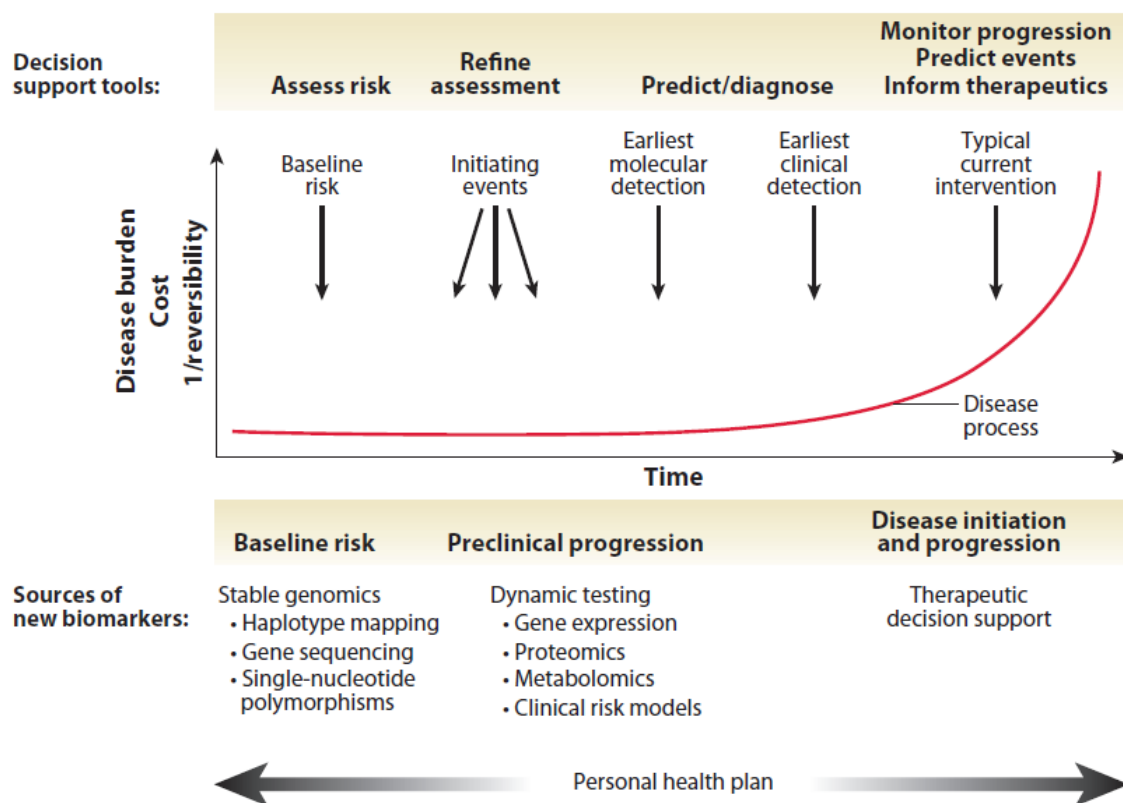
Il a été démontré que les systèmes ERS améliorent les pratiques de prescription, améliorent les soins préventifs et améliorent le respect des normes de soins fondées sur des données probantes (98, 99, 100). À mesure que la ERS s'intégrera aux dossiers de santé électroniques, l'utilisation des données personnelles sur la santé deviendra transparente, donnant aux cliniciens accès à l'information et aux connaissances des patients pour aider à créer des plans de santé adaptés aux besoins individuels.

### **1.4. La médecine personnalisée et le génome humain**

La médecine personnalisée est maintenant étayée par l'incorporation d'informations provenant des génomes, de leurs séquences et de leurs produits dérivés (tels que l'ARN, les protéines et les métabolites) dans la prise de décision médicale: les tests de santé basés sur le génome peuvent être utilisés pour prédire le risque, dépister les porteurs, établir des diagnostics cliniques et des pronostics pour les individus, et diriger la gestion clinique (101). Ainsi, l'information sur le génome humain permet maintenant aux fournisseurs de créer des plans de soins optimisés à chaque stade d'une maladie, en mettant l'accent sur les soins de santé réactifs plutôt que préventifs (102, 103, 104).



Tout au long du continuum de la santé à la maladie (figure 1), il existe maintenant plusieurs moments importants au cours desquels les applications génomiques personnalisent les soins de santé (105, 106). La susceptibilité à la maladie et le risque peuvent maintenant être quantifiés et anticipés pendant la santé et même à la naissance en utilisant des évaluations basées sur l'ADN qui ne changent pas au cours de la vie d'une personne. Les profils de transcription, les profils d'expression des protéines et les niveaux de métabolites combinés avec les modalités d'imagerie dynamique peuvent fournir des moyens plus précis de dépister les personnes à haut risque de développer une maladie dès les premières manifestations moléculaires alors que la maladie est subclinique. Cette même information peut fournir un diagnostic définitif et une classification moléculaire qui prédit le pronostic. De même, la pharmacogénomique - l'étude de la variation génétique et de l'efficacité et de la toxicité des médicaments - devient un exemple important d'information génomique appliquée à la médecine clinique (108). La sélection des médicaments peut être guidée à la fois par la constitution génétique sous-jacente du patient ainsi que l'architecture moléculaire de la maladie chez l'individu. Un bon exemple de cette approche peut être vu avec le dosage de warfarine, où les fournisseurs peuvent maintenant utiliser les systèmes de SSD en ligne (<http://www.warfarindosing.org>) pour déterminer la dose appropriée basée sur l'information génotypique et clinique d'un patient (109).



**Figure 5 : Le rôle de l'information génomique dans le continuum de la santé à la maladie.(Référence 93)**

## 2. La boîte à outils du génome humain

Le tableau 3 présente les principales plateformes technologiques du génome formant une boîte à outils pour l'exploration de l'impact du génome et de ses produits exprimés sur la santé et les états pathologiques. L'application clinique de ces technologies constitue la base de soins médicaux personnalisés. (Référence 103)

## **2.1. Variation à l'échelle du génome**

La variation dans le génome humain provient de plusieurs types d'altérations de séquence, y compris les mutations ponctuelles impliquant des changements de paires de bases mono nucléotidiques [polymorphismes mono nucléotidiques (SNP)] et les réarrangements structuraux de la séquence d'ADN (109, 110).

La plupart de nos connaissances sur la variation du génome humain proviennent de données de séquences d'ADN. La technologie de séquençage de nouvelle génération prend en charge le séquençage à haut débit et réduit les coûts au point où un génome peut être séquencé pour 10 000 \$, et on estime que ces coûts seront de 1 000 \$ en un an ou deux (111). À ce niveau de prix, le séquençage du génome d'un patient se situera dans la plage des tests médicaux standards (112, 113). Le séquençage de nouvelle génération a permis le reséquençage de plusieurs génomes humains, mettant en évidence l'évolution des plateformes de séquençage et leur capacité à étendre la couverture de lecture avec une plus grande résolution (114, 115, 116, 117). Le séquençage de nouvelle génération est maintenant utilisé pour comprendre plusieurs maladies - par exemple, le cancer, les maladies génétiques rares et les infections microbiennes - dans le but d'élucider les variantes géniques fonctionnelles sous-jacentes qui sont peut-être responsables de l'état pathologique (118, 119, 120, 121, 122, 123). L'utilisation du séquençage du génome entier avance à la clinique, où les données de séquences d'ADN ont été utilisées pour établir des diagnostics définitifs et même guider le traitement (124, 125, 126).

Le projet international HapMap (<http://hapmap.ncbi.nlm.nih.gov>), achevé en 2005, a développé une carte haplotype du génome humain pour détecter les

régions de déséquilibre de liaison élevé et de faible diversité haplotypique, identifiant les quelques SNP qui sont les variantes causales de la maladie (127, 128, 129). Ces données sont à la base des études d'association pangénomique (GWAS), une stratégie basée sur le génome qui utilise des comparaisons de SNP entre les populations pathogènes et témoins et fournit une méthode non biaisée de relier la variation du génome humain à la maladie clinique. En mai 2011, plus de 800 GWAS avaient été publiés sur 150 maladies et caractères humains, signalant plus de 2 400 SNP avec des associations et des rapports de cotes statistiquement significatifs (130, 131). Par exemple, dans la maladie de Crohn, des études d'association ont découvert 32 variants contribuant à la maladie, avec des variants dans les gènes NOD2 et IL23R associés à une augmentation du risque de maladie (132). Dans le diabète de type 2, au moins 20 variants ont été décrits, avec des variants plus récents émergeant de plus grands ensembles de données GWAS dans le contexte de l'homéostasie du glucose à jeun (133, 134). De même, les GWAS dans la maladie coronarienne (CAD) ont identifié plus de 100 variantes pouvant contribuer au risque de coronaropathie (135, 136, 137, 138, 139). Nombre de ces études ont indépendamment vérifié l'importance du locus 9p21.3. Les gènes contenus dans cette région codent pour l'ANRIL, un ARN régulateur non codant supposé réguler le silençage épigénétique par action sur le locus p15 / CDKN2B-p16 / CDKN2A-p14 / ARF, et les protéines kinases jouant un rôle inflammatoire induit par le TGF- $\beta$  dans l'athérosclérose (140, 141, 142, 143). D'autres ont suggéré que les SNP dans des amplificateurs distincts trouvés dans la région affectent la liaison STAT1 d'une manière dépendant de l'interféron  $\gamma$  dans les cellules endothéliales vasculaires humaines, reliant la susceptibilité CAD à la signalisation inflammatoire dans des cellules vasculaires spécifiques (144). Enfin, une utilisation précoce réussie du développement de

thérapeutiques basées sur les données GWAS dans la dégénérescence maculaire liée à l'âge où des thérapeutiques ciblées ont été développées pour les voies inflammatoires médiées par le complément identifiées par les GWAS (145,146).

Étonnamment, pour la grande majorité des GWAS qui ont été menées, les rapports de cotes des allèles associés au phénotype de la maladie ont été  $<2$  et peuvent ne pas contribuer significativement au risque de maladie. Il y a cependant quelques exceptions: Dans un GWAS de Ge et al. (147), un polymorphisme sur le chromosome 19, 3 kb en amont de l'IL-28B, codant pour l'interféron lambda-3, était associé à un changement de deux fois de la réponse au traitement. Cette découverte a été cliniquement confirmée dans une étude de Thompson et al. (148) ont génotypé des patients recevant un traitement contre le virus de l'hépatite C et ont trouvé que le même polymorphisme était un bon prédicteur d'une réponse virologique soutenue (rapport de cotes 5,2, intervalle de confiance à 95%, 4,1-6,7). Dans l'étude SEARCH Collaborative Group, un polymorphisme localisé sur SLCO1B1, un gène régulant l'absorption hépatique des statines, était associé à un risque accru de myopathie suite à un traitement par statine (odds ratio 4,5, intervalle de confiance 95%, 2,6-7,7) (149). Bien que la pertinence clinique des GWAS ne soit pas aussi répandue que nous l'avions imaginé, ces études ont révélé de nombreux nouveaux gènes contribuant à la pathogenèse qui ont fourni de nouvelles perspectives sur la biologie sous-jacente, les mécanismes de la maladie et les approches thérapeutiques potentiellement nouvelles.

**Tableau 3 : La boîte à outils du génome humain (Référence 93)**

Table 2 The human genome toolbox

| “-omics” approach                          | Generated information                                  | Applications                                                                     | Notable examples                                                                                  |
|--------------------------------------------|--------------------------------------------------------|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|
| Human genome sequence (genomics)           | Whole-genome sequence, SNPs, and CNVs (~10–15 million) | Disease mechanisms<br>Disease diagnosis<br>Pharmacogenomics                      | Age-related macular degeneration (120), HCV virologic response (1), AML (32), warfarin dosing (6) |
| Gene expression profiles (transcriptomics) | Microarrays and RNA sequencing (~25,000 transcripts)   | Disease mechanisms<br>Disease diagnosis<br>Disease prognosis<br>Pharmacogenomics | AML (71), ALL (94), ACS (20), breast cancer (161)                                                 |
| Proteome (proteomics)                      | Protein profiles of specific protein products          | Disease diagnosis                                                                | ACS (143)                                                                                         |
| Metabolome (metabolomics)                  | Metabolic profiles (~1,000–10,000 metabolites)         | Disease mechanisms<br>Pharmacogenomics                                           | ACS (182), drug toxicity (44), cancer profiling (76), CAD (193)                                   |

Abbreviations: ACS, acute coronary syndromes; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CAD, coronary artery disease; CNV, copy number variation; HCV, hepatitis C virus; SNP, single-nucleotide polymorphism. Table adapted from Reference 66.

## 2.2. Transcriptomique

La transcriptomique est l'étude génomique des niveaux d'expression de l'ARN dans une cellule, un tissu ou un fluide biologique, et comprend un spectre de molécules allant de l'ARN messenger (ARNm) aux ARN non codants. Les analyses statistiques et la classification des gènes coexprimés à l'aide de méthodes paramétriques ou non paramétriques peuvent être appliquées pour classer la maladie ou prédire les états pathologiques futurs, ou pour comprendre les mécanismes biologiques sous-jacents de la maladie (150).

Dans le cancer, les données de biopuces ont été utilisées pour le diagnostic, le pronostic et la réponse au traitement (151,152). Ces données ont révélé des sous-ensembles discernables de classes moléculaires dans de nombreux cancers, y compris les cancers du sein, du poumon, du sang et du mélanome (153, 154, 155, 156). Chez les patients atteints de leucémies aiguës, ce diagnostic moléculaire est particulièrement utile pour la prise de décision clinique, où le traitement peut différer en fonction du sous-type de maladie en réseau (157).

L'application de la technologie des micros réseaux à la classification des maladies a été appliquée dans d'autres maladies complexes, y compris les maladies cardiovasculaires, les maladies rhumatismales et leurs voies inflammatoires, et les maladies neurologiques telles que la sclérose en plaques et les troubles psychiatriques tels que la schizophrénie, dépression (158, 159, 160, 161, 162, 163).

Une caractéristique du transcriptome qui a émergé est l'expression des profils d'ARN non codants, en particulier les petits ARN interférents (siRNA) et les microARN (miARN). D'abord identifiés en 1993 chez *Caenorhabditis elegans*, les miARN ont généralement une longueur de 22 nucléotides et inhibent l'expression génique à un niveau post-transcriptionnel en se liant à des régions non traduites complémentaires des ARNm cibles (164,165). Ils ont joué un rôle dans plusieurs maladies, de la régulation des cancers proliférants au développement des cardiomyopathies, de la réponse immunitaire et de la schizophrénie, et même à la promotion de la réplication du virus de l'hépatite C (166, 167, 168, 169, 170, 171). Les modèles de miARN différentiellement exprimés peuvent être utilisés cliniquement dans le diagnostic et le pronostic de la maladie (172, 173,174). Compte tenu de leur rôle dans la complexité de la régulation des gènes, les miARN et même les miARN et les siRNA sont testés (175). Un essai de patients humains avec des tumeurs solides réfractaires a montré des résultats prometteurs après la livraison de siRNA en utilisant un système de nanoparticules ciblées (176). Cependant, l'analyse des miARN peut être difficile en raison de leur manque de spécificité, conduisant à des effets hors cible: En raison de leur petite taille, ils peuvent potentiellement influencer l'expression de plusieurs ARNm différents (177, 178).

Enfin, les technologies de nouvelle génération utilisées pour faire progresser le séquençage du génome entier peuvent être exploitées pour analyser les transcriptomes: le séquençage de l'ARN (RNA-Seq), initialement utilisé pour examiner les courts ARN régulateurs, a été appliqué à l'analyse du transcriptome humain entier pour continuer notre compréhension de la régulation de la transcription (179, 180, 181, 182, 183). Par exemple, les transcriptomes liés au cancer ont été étudiés par RNA Seq, fournissant une nouvelle approche pour comprendre les événements de fusion et de translocation de gènes (184, 185, 186, 187).

### **2.3. Protéomique**

L'étude de la protéomique fait référence à l'étude à grande échelle des protéines, et le protéome se réfère souvent à l'ensemble complet des protéines et de leurs divers dérivés. Dans le contexte de la santé et de la maladie, la protéomique cherche à définir l'ensemble complet des protéines associées à un état physiologique ou pathologique particulier et continue d'être un domaine prioritaire pour la découverte de biomarqueurs et la médecine moléculaire.

Cette zone a été dominée par des approches isotopiques stables, mais les méthodes quantitatives sans marqueur récemment développées, qui reposent sur l'intensité mesurée d'un ion peptidique par rapport à son intensité dans d'autres échantillons, ont l'avantage d'un débit plus élevé et moins d'étapes de manipulation (188,189). Bien que cette technologie soit relativement immature dans ses applications à la santé humaine et à la maladie par rapport à l'ADN, l'ARN et le profil métabolique, ces méthodes, combinées au développement de la spectroscopie de masse, feront progresser la protéomique dans la classification des maladies, le diagnostic, le pronostic et la pharmacogénomique au cours des prochaines années.



## **2.4. Métabolomique**

La métabolomique utilise la spectroscopie de masse et la spectroscopie de résonance magnétique nucléaire pour mesurer les changements dans les petites molécules non protéiques liées à un état biologique ou physiologique. On estime que le métabolome humain contient environ 5 000 métabolites discrets de petites molécules, et l'identification des changements métaboliques associés à la maladie suggère immédiatement des cibles médicamenteuses enzymatiques. Cette approche a été appliquée à un certain nombre d'états pathologiques chroniques, tels que le diabète, l'obésité, les maladies cardiovasculaires, le cancer et les troubles mentaux (190, 191, 192, 193). Dans le domaine des maladies cardiovasculaires, deux études distinctes ont défini des profils de métabolites pour l'ischémie et la coronaropathie, respectivement (194, 195). Le profilage métabolomique a également été utilisé directement pour évaluer la toxicité des médicaments (196, 197, 198)

## **2.5. Épigénomique**

L'épigénomique fait référence à la programmation génétique qui résulte principalement de la méthylation de l'ADN (199). La base moléculaire de cette régulation implique la méthylation de la cytosine et est maintenue par une famille d'ADN méthyltransférases; des quatre bases comprenant l'ADN, seule la cytosine peut être méthylée et non méthylée, et seulement dans le contexte des dinucléotides CpG, qui chevauchent souvent la région promotrice des gènes. Les gènes transcrits activement sont généralement dans un état non méthylé, tandis que les gènes qui sont réduits au silence sont fortement méthylés (200, 201). L'analyse des tumeurs montre une hypométhylation chez les oncogènes et une hyperméthylation régionale chez les gènes suppresseurs de tumeurs (202, 203,

204). Des changements de méthylation ont également été mis en évidence dans d'autres maladies telles que le diabète de type 2, les pathologies cardiovasculaires et les maladies auto-immunes telles que le lupus érythémateux disséminé (205, 206, 207, 208). Les Instituts nationaux de la santé ont lancé le projet Human Epigenome pour identifier, cataloguer et interpréter les profils de méthylation de l'ADN de tous les gènes humains dans tous les principaux tissus, et ce projet a depuis été étendu au niveau international (209, 210, 211).

## **2.6. Biologie des systèmes et médecine des systèmes**

La biologie des systèmes et son domaine appliqué de la médecine systémique reposent sur le principe que les mesures de l'information biologique doivent être globales. Les différents niveaux d'information génomique doivent être étudiés de manière intégrée et les systèmes biologiques fonctionnent dynamiquement, non comme des événements discrets (212). La santé humaine et la maladie sont définies par des voies moléculaires interdépendantes qui peuvent être définies par le génome, l'épigénome, le transcriptome, le protéome et / ou le métabolome. Bien que chacun puisse être analysé individuellement, le cœur de la biologie des systèmes est la création d'un modèle qui intègre toutes les données d'expression et ses changements dynamiques. L'analyse des données de systèmes à partir d'un seul échantillon de sang ou de tissu fournit des informations sur la classification moléculaire et la prédiction pour un état pathologique particulier ainsi que sur le contexte biologique du développement de nouvelles cibles thérapeutiques (213, 214, 215, 216).

### **3. De l'information de génome à la médecine personnalisée**

La traduction d'informations génomiques en médecine clinique nécessite de répondre aux questions suivantes: Quels sont les facteurs moléculaires responsables de la pathogenèse d'une maladie particulière? Pouvons-nous déterminer quels patients développeront une maladie? Comment pouvons-nous développer un régime thérapeutique qui peut être adapté à la biologie unique d'un individu? Le tableau 4 donne des exemples de marqueurs génomiques utilisés cliniquement aujourd'hui dans le continuum de l'état normal à l'état maladie. Beaucoup de ces outils ont été traités dans les sections précédentes, et cette section mettra en évidence l'utilisation de ces outils dans la pratique médicale aujourd'hui.

#### **3.1. Tests préconceptionnels et dépistage des hétérozygotes**

En France, le dépistage préconceptionnel des couples à risque de maladies héréditaires monogéniques s'effectue dans le cadre d'un conseil génétique s'il existe un cas index dans la famille. Mais dans quelques populations présentant une forte prévalence de certaines affections héréditaires le dépistage des hétérozygotes est devenu depuis longtemps systématique :  $\beta$  thalassémies dans les pays méditerranéens et maladie de Tays-Sachs dans la communauté juive Ashkénase des États-Unis. À Chypre, le dépistage de la  $\beta$  thalassémie a fait baisser le nombre de cas de 51 en 1974 à simplement 5 entre 1991 et 2001 ; aucun cas n'a été observé entre 2002 et 2007. Aux États-Unis 1379 couples à risque de Tays-Sachs ont été identifiés (aboutissant soit à la non réalisation de l'union projetée, soit à un diagnostic prénatal, soit à un diagnostic préimplantatoire) et l'incidence de la maladie a diminué de 90 % ; dans cette communauté juive le programme de dépistage a été étendu à 16 autres maladies

récessives. En Israël, la détection systématique est recommandée pour 8 maladies dont la fréquence des hétérozygotes est supérieure à 1/60 et gratuite pour 4 d'entre elles (mucoviscidose, Tays-Sachs, dysautonomie familiale, thalassémie) ; elle est également encouragée pour des maladies dont la fréquence des hétérozygotes est supérieure à 1/110 (Nieman Pick, glycogénoses, etc.). Aux États-Unis le dépistage des hétérozygotes est recommandé pour des maladies fréquentes comme la mucoviscidose ou l'amyotrophie spinale ; la firme « Genzyme genetics » propose ainsi un test pour 98 des mutations les plus fréquentes du gène CFTR (un couple à risque sur 1 400).

En France, le dépistage néonatal systématique concerne cinq maladies : mucoviscidose, phénylcétonurie, hypothyroïdies, hyperplasies congénitales des surrénales et drépanocytose ; il s'agit d'un dépistage biologique des homozygotes. L'apparition des techniques de séquençage à haut débit est susceptible d'entraîner le passage à un dépistage neonatal génotypique, identifiant également de ce fait les hétérozygotes et éventuellement étendu à plusieurs centaines de maladies dont les gènes et les mutations pathogènes sont connus et ceci pour moins de 400 dollars (217). Les NGS rendent ainsi possible le diagnostic préconceptionnel généralisé, ce qui soulève de multiples questions tant éthiques que socio-économiques. Sur un tel marché potentiel se positionnent dès à présent des firmes commercialisant des DTC (Direct Test to Consumer). Ainsi « Amby Screen » propose le dépistage de 75 maladies pédiatriques à transmission récessive ou liée à l'X. « 23 and me » offre un test plus étendu incluant, outre le génotypage de mutations responsables de maladies monogéniques, la recherche de polymorphismes associés à des maladies dites plurifactorielles et un profil pharmacogénétique, et cela pour 99 dollars.

### 3.2. Diagnostic prénatal et préimplantatoire

Les diagnostics prénataux ne peuvent être pratiqués en France que dans un cadre réglementaire strict par des centres pluridisciplinaires et des laboratoires agréés ; dans le cas du diagnostic préimplantatoire seuls quatre centres sont autorisés par l'Agence de Biomédecine : Paris, Strasbourg, Montpellier et Nantes. Les indications diagnostiques sont restreintes puisque la loi n'autorise l'interruption de grossesse que pour une affection d'une particulière gravité et reconnue comme incurable. Le rapport 2010 de l'Agence fait état de 55 000 diagnostics prénataux en cytogénétique (essentiellement recherche de trisomie) dont près de 4 000 positifs et de 2 750 en génétique moléculaire dont 544 positifs. Le diagnostic prénatal est effectué sur biopsie de trophoblaste à 12 semaines d'aménorrhée (entraînant 1 à 2 % d'avortements), sur cellules amniotiques à 16 semaines (0,2 à 0,5 % d'avortements), sur sang foetal après 20 semaines (3 % d'avortements) ou sur tissu foetal au troisième trimestre de la grossesse. Un diagnostic préimplantatoire nécessite une stimulation ovarienne hormonale, la collecte d'ovules suivie d'insémination artificielle, une biopsie embryonnaire au troisième jour (J3), un diagnostic en 24-48 heures (caryotype avec hybridation *in situ* ou recherche d'une mutation) et un transfert *in utero* de 1 à 3 embryons sains à J4/J5 ; chacune de ces étapes comporte un pourcentage d'échecs important et en définitive le nombre de grossesses ainsi obtenu est faible (218).

Les pertes foetales liées aux prélèvements invasifs pratiqués pour le diagnostic prénatal sont estimées entre 300 et 600 par an, d'où l'intérêt majeur d'un diagnostic non invasif par analyse des acides nucléiques foetaux circulant dans le sang maternel. La présence d'ADN foetal ou cffDNA (cell free fetal

DNA) dans le sang maternel a été démontrée en 1997 et sa concentration estimée à plus de 10 % de l'ADN total circulant ; il disparaît rapidement du sang maternel après l'accouchement, contrairement aux cellules mononuclées foetales qui pourraient persister plusieurs années. Détectable dès 5 à 6 semaines d'aménorrhée, il a une origine trophoblastique ou placentaire, et provient sans doute de cellules en apoptose ; il est *Bull. Acad. Natle Méd.*, 2014, 198, no 1, 101-117, séance du 7 janvier 2014 106 constitué majoritairement de petits fragments dont la taille moyenne est de 143 bases, ce qui correspond à la longueur d'ADN enroulé sur un nucléosome. La première application de l'analyse de l'ADN foetal du sang maternel a été la détermination du sexe du fœtus et potentiellement des maladies liées à l'X : il s'agit d'un examen inscrit à la nomenclature des actes de biologie (B500). En 2011 la HAS a également donné un avis favorable à la prise en charge de la détermination du rhésus foetal chez les femmes rhésus négatif : actuellement environ 6000 examens par an, mais il s'agit d'un « marché » potentiel de 160 à 180 000 tests annuels. En utilisant des « primers » spécifiques il était également possible de faire différents diagnostics ciblés ( $\beta$  thalassémie, hémophilie A...). L'apparition des techniques de séquençage à haut débit, en séquençant à la fois l'ADN circulant maternel et l'ADN foetal et en déterminant leurs proportions relatives (par détermination des SNPs pour reconstituer les haplotypes), a ouvert de nouvelles perspectives : diagnostic des maladies génétiques et surtout des aneuploïdies (219, 220). La recherche d'une trisomie repose sur la mise en évidence d'un petit excès de matériel provenant de ce chromosome par rapport à l'ADN maternel ; il implique un séquençage aléatoire des petits fragments d'ADN avec une grande profondeur suivi d'une analyse bioinformatique complexe. Les premiers résultats récemment publiés sont très bons pour la trisomie 21 (sensibilité 100 %

et 0,1 % de faux positifs) et la trisomie 18 (sensibilité 100 % et 0,3 % de faux positifs), sensiblement moins performants en cas de trisomie 13 (sensibilité de 91,4 % et 0,9 % de faux positifs). Bien entendu les possibilités offertes par les NGS dans ce domaine vont se multiplier ; elles devraient pouvoir également s'appliquer dans l'avenir au diagnostic préimplantatoire, mais aussi à la détection des petites délétions et des variations du nombre de copies ou CNV (« Copy Number Variation »), restreignant ainsi le champ de la cytogénétique classique (221).

### **3.3. Prédiction du risque**

L'information génomique change le paysage médical en fournissant des outils qui transforment la pratique en prévision et en prévention du risque. Par exemple, les femmes porteuses de mutations dans les gènes BRCA1 et BRCA2 à forte pénétrance présentent un risque accru de susceptibilité au cancer du sein (222). Le dépistage clinique de ces mutations génétiques a conduit à des mesures préventives standard telles que des interventions chirurgicales ou radio thérapeutiques pour aider les patients à minimiser leur risque (223). Dans le cancer colorectal, les méthodes cliniques pour détecter l'instabilité des microsatellites dans les gènes de réparation des mésappariements MLH1 et MSH2 identifient les patients à risque et permettent aux médecins de les surveiller de près grâce au dépistage fréquent du cancer du côlon (224). Des facteurs de risque génétiques cliniquement utilisables ont également émergé pour le syndrome du QT long (LQTS) et le syndrome du QT court (225). LQTS est hérité selon un schéma autosomique dominant, et la variation du phénotype de la maladie est associée à des mutations dans 12 gènes différents de susceptibilité LQTS, ce qui justifie de tester génétiquement les patients pour ces

mutations (226). Par exemple, les  $\beta$ -bloquants sont un traitement efficace pour les patients avec LQTS1 (KCNQ1) mais pas pour ceux avec LQTS2 (KCNH2) ou LQTS3 (SCN5A), démontrant comment le génotypage de ces patients peut aider à guider le traitement (227). L'utilisation de tests de variation génétique utilisant le génotypage et le séquençage combinés avec des stratégies interventionnelles peut retarder la progression de la maladie ou prévenir complètement la maladie.

### **3.4. Diagnostic de la maladie et caractérisation moléculaire**

Une application précoce de la technologie du génome à la médecine clinique est dans la classification des états pathologiques. Les analyses génomiques et moléculaires ont révélé des sous-types distincts de la maladie, qui ont été traditionnellement définis par de larges phénotypes cliniques ou descriptifs. Parce que la réponse thérapeutique peut différer entre chaque sous-type, cela a des implications claires pour guider le traitement clinique.

L'oncologie offre de multiples exemples de la façon dont la médecine génomique a changé la compréhension de la maladie. Plusieurs génomes de cancer ont été séquencés, y compris la leucémie myéloïde aiguë, le cancer du poumon non à petites cellules (NSCLC) et le cancer du poumon à petites cellules (228, 229, 230). L'Atlas du génome du cancer, un effort multi-institutionnel de 1,5 milliard de dollars sur dix ans, vise à caractériser et à réaliser une analyse génomique multidimensionnelle du cancer (231). L'un des premiers cancers étudiés était le glioblastome multiforme (GBM); 206 échantillons ont été analysés pour le nombre de copies d'ADN, l'expression des gènes et les états de méthylation de l'ADN. De nouvelles mutations fréquentes du gène de la sous-unité régulatrice de la kinase PI3 ont été trouvées ainsi qu'une



association entre l'état de méthylation de MGMT (biomarqueur pour la réponse au témozolomide) et les mutations de réparation de mésappariement dans le GBM post-traitement (232). Une autre conclusion importante à tirer de ce projet est la caractérisation de sous-types distincts de GBM (classique, mésoenchymateuse, pro neurale) basés sur des mutations génétiques causales dans EGFR, NF1 et PDGFR / IDH1, respectivement (233).

La capacité à trouver plusieurs profils d'expression génique discrets dans un type de tumeur reflète l'hétérogénéité de la maladie. L'un des premiers exemples d'utilisation de profils de micro réseaux d'expression génique pour classer les tumeurs a été utilisé pour distinguer la leucémie myéloïde aiguë de la leucémie lymphoblastique aiguë (234). Des méthodes similaires ont également été appliquées avec succès pour caractériser, au niveau moléculaire, une large gamme de cancers du sein, de cancers du poumon et de lymphomes (235, 236, 237).

L'utilisation de la génomique pour stratifier la maladie n'est pas propre à l'oncologie. Les syndromes coronariens aigus (SCA) décrit une classe de maladies impliquant l'occlusion des artères coronaires, entraînant une lésion ischémique du cœur. Plusieurs études ont démontré l'utilité de la métabolomique et de la protéomique pour distinguer les états SCA de l'infarctus aigu du myocarde et de l'angor instable (238, 239). Une étude récente du protéome plaquettaire chez des patients avec et sans sus-décalage du segment ST a mis en évidence des différences d'activation plaquettaire et, en particulier, des taux élevés de la protéine SPARC (240). L'expression du gène du sang périphérique a également été associée à la sévérité de la coronaropathie, et un test clinique diagnostique a été développé sur la base d'observations similaires (CorusTM

CAD) et validé dans un essai prospectif multicentrique avec des résultats impressionnants, améliorant la précision de l'évaluation des maladies coronariennes de 16% -20% (241, 242). Ainsi, un certain nombre de plateformes de tests génomiques complexes sont utilisées dans la clinique ou sont préparées pour une utilisation clinique afin d'aider au diagnostic et d'éclairer la prise de décision thérapeutique.

**Tableau 4 : Diagnostics moléculaires disponibles cliniquement sur le continuum de la santé à la maladie dans le cancer et les maladies cardiovasculaires (référence 105)**

**Table 3 Clinically available molecular diagnostics along the continuum from health to disease in cancer and cardiovascular disease**

| Time point in clinical decision making | Cancer                                                                                                |                                                                                                 | Cardiovascular disease                                                            |                                                        |
|----------------------------------------|-------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------|
|                                        | Test                                                                                                  | Indication                                                                                      | Test                                                                              | Indication                                             |
| Risk/susceptibility                    | <i>BRCA1, BRCA2</i><br><i>HNPCC, MLH1, MSH2</i><br><i>TP53, PTEN</i>                                  | Breast<br>Colon<br>Sarcomas                                                                     | <i>KIF6, 9p21</i><br>Familion® 5-gene profile                                     | CAD<br>LQTS                                            |
| Screening                              | HPV genotypes                                                                                         | Cervical                                                                                        | Corus™ CAD                                                                        | CAD                                                    |
| Diagnosis                              | Lymphochip                                                                                            | Lymphoma                                                                                        | Corus CAD                                                                         | CAD                                                    |
| Prognosis                              | Oncotype DX® (21-gene assay)<br>MammaPrint® (70-gene assay)<br>Her2/neu, ER, PR                       | Breast                                                                                          | TnI, BNP, CRP                                                                     | ACS                                                    |
| Pharmacogenomics                       | Her2/neu<br><i>UGT1A1</i><br><i>KRAS</i><br><i>EGFR</i><br>Amplichip®; DMET™<br><i>CYP2D6/CYP2C19</i> | Herceptin<br>Irinotecan<br>Cetuximab<br>Erlotinib, gefitinib<br>Various others<br>(see Table 2) | <i>KIF6, SLCO1B1</i><br>Amplichip; DMET<br><i>CYP2D6/CYP2C19</i><br><i>VKORC1</i> | Statins<br>Warfarin<br>Various others<br>(see Table 2) |
| Monitoring                             | CTCs                                                                                                  | Tumor recurrence or progression                                                                 | AlloMap® gene profile                                                             | Transplant rejection                                   |

Abbreviations: ACS, acute coronary syndromes; BNP, brain natriuretic peptide; CAD, coronary artery disease; CRP, C-reactive protein; CTC, circulating tumor cell; ER, estrogen receptor; HPV, human papillomavirus; LQTS, long QT syndrome; PR, progesterone receptor. Table adapted from Reference 64.

### 3.5. Pronostic de la maladie

Les modèles de signatures tumeur-expression génique, combinés à des données cliniquement pertinentes telles que les issues de survie, fournissent l'un des exemples les plus clairs de l'utilisation clinique de l'information génomique

en tant qu'outils pronostiques (243, 244, 245, 246). MammaPrint R\_ et Oncotype DXR\_ sont deux de ces tests pronostiques qui sont maintenant cliniquement disponibles pour une utilisation dans le cancer. L'oncotype DX est une signature de 21 gènes dérivée d'une réaction en chaîne de la polymérase en temps réel d'échantillons inclus dans la paraffine et utilisée pour prédire la récurrence à distance sur une période de 10 ans (247, 248, 249). Il est actuellement en cours de validation dans un essai clinique de phase III, mais a une plus grande signification statistique par rapport aux facteurs de prédiction histologiques traditionnels (250, 251). MammaPrint, approuvé par la Food and Drug Administration américaine en 2007, a été développé au Netherlands Cancer Institute, qui a créé une signature à 70 gènes pour discriminer les patients à haut et faible risque de récurrence métastatique dans les cinq ans (252). Des essais cliniques basés sur cette signature ont depuis été lancés pour valider la signature, et les auteurs de ces études ont rapporté une prédiction précise (253, 254). Une méta-analyse a récemment montré que les informations pronostiques d'Oncotype DX jouent un rôle important dans la prise de décision clinique et ont conduit à une réduction de la chimiothérapie tout en augmentant les thérapies hormonales adjuvantes (255).

Les approches génomiques ont également contribué au pronostic après transplantation rénale. Une étude a été menée chez des receveurs de greffe de rein qui n'ont pas rejeté leurs greffes rénales malgré l'arrêt du traitement immunosuppresseur pour analyser les profils d'expression génique des globules blancs circulants. Cela a conduit à la découverte d'une expression génique spécifique dans les cellules B qui distinguait les patients tolérants des patients non tolérants (256). Les résultats de ceci impliquent que des biomarqueurs

peuvent être développés qui peuvent prédire quels patients devraient continuer à recevoir l'immunosuppression des années après la chirurgie de transplantation et quels patients peuvent avoir ceci discontinué.

### **3.6. Pharmacogénomique**

La pharmacogénomique est l'utilisation de l'information génomique pour déterminer ou prédire la réponse à la thérapie. En oncologie, le génome somatique de la tumeur a été utilisé pour identifier les biomarqueurs qui prédisent et guident la thérapie pour une réponse optimale. Par exemple, dans la leucémie lymphoblastique aiguë, environ 20% des patients deviendront réfractaires au traitement. Deux études utilisant l'expression génique différentielle ont montré que les gènes impliqués dans la chimiorésistance étaient également associés à un pronostic plus sombre, jetant les bases de futures études sur les mécanismes de résistance (257, 258).

La thérapie ciblée est souvent synonyme de pharmacogénomique. Le gène récepteur du facteur de croissance, le récepteur du facteur de croissance épithélial humain (HER2), s'est révélé amplifié dans 25% à 30% des cancers du sein, et sa surexpression corrélée avec un mauvais pronostic (259,260). Le Trastuzumab, un anticorps monoclonal qui cible HER2, est efficace pour réduire la charge tumorale, et le dépistage des patients pour l'expression de HER2 / neu est devenu une routine et une exigence pour l'utilisation du médicament (261, 262, 263). Dans le CBNPC, les effets antitumoraux ont été démontrés en utilisant des inhibiteurs du récepteur du facteur de croissance épidermique (EGFR), un récepteur de surface cellulaire fréquemment muté pour provoquer la prolifération du cancer (264, 265, 266). Bien que l'essai Iressa de survie au cancer du poumon - un essai de phase III randomisé et contrôlé par placebo - ait

montré que le gefitinib, un inhibiteur de l'EGFR, ne procurait aucun avantage de survie à la population générale, un sous-groupe de non-fumeurs asiatiques mettre en évidence le potentiel de la génomique pour diriger et focaliser la thérapie (267). Parce que seulement 5% à 15% des patients blancs portent une mutation somatique EGFR, comparativement à 25% -35% des patients asiatiques atteints de CPNPC, ces études soulignent le besoin d'un dépistage génétique supplémentaire avant d'administrer un traitement (268).

Un exemple important de pharmacogénétique concerne la prise en charge de la warfarine (269). En raison de l'index thérapeutique étroit du médicament, diverses complications sont associées à son traitement, même après ajustement de la dose en fonction de l'âge, du sexe, du poids, de l'état pathologique, du régime alimentaire et des médicaments concomitants. L'étude des propriétés pharmacocinétiques de la warfarine a indiqué l'implication de deux gènes dans la détermination de la dose de warfarine. L'un de ces gènes code le CYP2C9, responsable de la majeure partie de la clairance métabolique (~ 80%) de l'énantiomère S, plus puissant sur le plan pharmacologique, de la warfarine. Les deux CYP2C9 \* 2 et \* 3 provoquent une réduction de la clairance de la S-warfarine, avec une variation de 10 fois du génotype liée à l'activité la plus élevée (CYP2C9 \* 1 / \* 1) à la plus faible (CYP2C9 \* 3 / \* 3). De nombreuses études ont associé ces génotypes à une sensibilité initiale à la dose, à une stabilisation retardée de la dose d'entretien, à des retards de sortie de l'hôpital et à des complications hémorragiques accrues (270). Le second gène identifié comme prédicteur du dosage est la protéine complexe de la vitamine K époxyde réductase 1 (VKORC1), ciblée par la warfarine. La prise en compte du génotype VKORC1 ou de l'haplotype associé au génotype CYP2C9 et des facteurs tels

que l'âge et la taille corporelle sont estimés représenter 35% à 60% de la variabilité des besoins posologiques en warfarine (271).

Malgré le fait que plusieurs groupes indépendants ont reproduit ces données, des études cliniques prospectives sont nécessaires pour établir si la dose initiale peut être adaptée aux patients par le génotypage CYP2C9 et VKORC1, couplée avec des variables cliniques connues (272). Bien que ces études soient actuellement en cours, le comité consultatif de pharmacologie clinique de la FDA a reconnu l'importance et le potentiel du génotypage du CYP2C9 et du VKORC1 durant la phase précoce du traitement par warfarine, et l'étiquette du médicament a été modifiée en conséquence en août 2007.

### **3.7. Surveillance de la réponse des maladies à la thérapie**

Le profil d'expression génique des cellules mononuclées du sang périphérique (PBMC) est couramment utilisé dans certains centres pour surveiller l'état des greffes après transplantation d'organes solides, l'essentiel étant de développer de meilleurs tests pour évaluer le rejet du greffon (273, 274,275). La recherche en transplantation cardiaque, peut-être plus que tout autre organe, a ouvert la voie à la mise au point de nouveaux tests sanguins permettant de faire la distinction entre la quiescence et le rejet aigu. L'étude d'observation génique de rejet d'allo greffon cardiaque (CARGO), une étude multicentrique phare en trois étapes dans le domaine de la transplantation cardiaque, a montré que le profil d'expression génique des PBMC pouvait être utilisé pour distinguer avec précision les patients post-transplantation versus rejet aigu (276). Ceci a été réalisé en identifiant d'abord un ensemble de 11 gènes qui distinguaient le mieux entre le rejet aigu et la quiescence, puis en utilisant une cohorte indépendante pour valider le schéma de classification dans

un mode prospectif et aveugle. En utilisant un système de notation de 0 à 40, avec de faibles scores représentant un faible risque de rejet de greffe, les chercheurs ont utilisé un seuil de 30 sur les patients plus d'un an après la transplantation et la validation du test a donné une valeur prédictive négative de 99,6% et une valeur prédictive positive de 6,8% pour le rejet aigu. Ce test est disponible dans le commerce sous le nom AlloMap R<sub>+</sub> (XDx; Brisbane, CA) et coûte environ 3 000 \$ par test. Bien que coûteuse, l'analyse des coûts basée sur les données de la cohorte CARGO a montré que ce test de profil d'expression génique est moins cher que la biopsie et permettrait d'économiser jusqu'à 12 millions de dollars en coûts de santé par an (277). Pour que cela soit vrai, cependant, le besoin de biopsie devrait être évité, et le profil d'expression génique devrait suffire chez les patients considérés comme présentant un faible risque de rejet. Un essai prospectif, multicentrique, randomisé mais non aveugle appelé Atténuation Invasive de Surveillance par Expression Génique (IMAGE) a montré qu'une stratégie de surveillance non invasive utilisant le profil d'expression génique n'est pas inférieure à la biopsie endomyocardique pour diagnostiquer un dysfonctionnement d'allogreffe cardiaque dans le rejet. Les auteurs ont conclu que l'utilisation de cette stratégie de surveillance du rejet réduit significativement le nombre de biopsies sans risque accru d'événements indésirables, faisant de l'étude un exemple clé où le profil d'expression génique des PBMC peut être utilisé pour guider la prise de décision thérapeutique (278).

#### **4. Génomes interactives et microbiomes**

Bien que la génomique joue un rôle important dans le diagnostic, le pronostic, le traitement et la prévention de la maladie, il reste beaucoup à savoir sur l'interaction humaine avec l'environnement. Des études récentes de la

réponse de l'hôte aux pathogènes microbiologiques ont démontré des changements d'expression génique très spécifiques en réponse à différents virus; par exemple, une signature d'expression génique basée sur l'hôte a été récemment identifiée qui pourrait être utilisée pour la détection précoce d'une infection virale (279,280). Les mêmes données ont également été utilisées pour montrer que les profils d'expression génique distinguent les infections bactériennes et virales (Figure 6). Les profils d'expression des gènes dans les cellules hôtes ont également été modifiés en fonction d'autres facteurs de stress environnementaux tels que le tabagisme et les états pathologiques comme l'asthme et la bronchopneumopathie chronique obstructive, et pourraient permettre de comprendre la pathogenèse et la susceptibilité au risque (281, 282, 283). Ainsi, les interactions gène-environnement mesurées en utilisant la réponse du génome de l'hôte commencent à fournir de nouveaux biomarqueurs cliniquement utiles et potentiellement diagnostiques.

#### **4.1. Génomes pathogènes**

Les génomes des pathogènes humains sont également séquencés dans l'espoir de fournir des informations pertinentes sur la pathogenèse de l'infection ou des cibles thérapeutiques potentielles. En 1995, le premier génome bactérien achevé a été séquencé par *Haemophilus influenzae* (284). Depuis lors, plus de 1 000 génomes bactériens et plus de 3 000 génomes viraux ont été complètement séquencés (285). Cette information génomique a fourni des données précieuses pour les outils génomiques comparatifs, en comparant les souches pathogènes avec les souches non pathogènes pour identifier les protéines qui sont utilisées pour la pathogenèse, conduisant potentiellement à de nouveaux médicaments avec une toxicité limitée (286). L'analyse protéomique comparative entre les

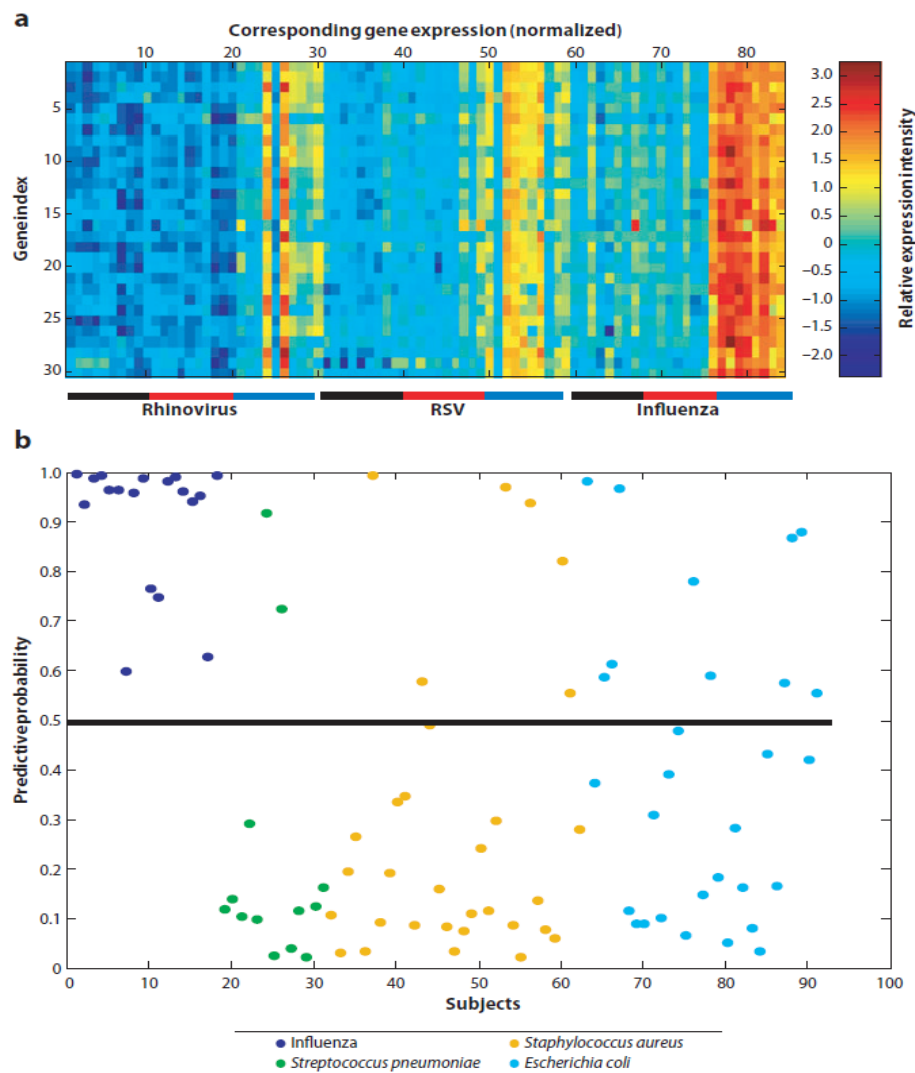


phénotypes virulents et avirulents de *Rickettsia Prowazekii* a révélé des mutations adaptatives dans huit gènes responsables de sa virulence (287). Le séquençage de génomes bactériens pour des cibles vaccinales telles que les protéines de surface bactériennes a également connu un grand succès clinique: Des séquençages vaccnologiques inversés des protéines cibles ont conduit au développement d'un vaccin prometteur pour le séro groupe B de *Neisseria meningitidis* récemment introduit dans les essais cliniques (288). Une autre approche pour comprendre comment les pathogènes développent une résistance a été de suivre les changements génomiques des isolats de *Staphylococcus aureus* au cours du temps chez un patient après l'exposition au traitement antibiotique jusqu'à la résistance, une stratégie qui a identifié 35 mutations dans 31 loci (289).

#### **4.2. Le microbiome humain**

Les interactions synergiques entre les humains et notre flore corporelle normale doivent être comprises, car il y a 10 fois plus de bactéries que le nombre de cellules humaines dans le corps: Le côlon humain contient à lui seul plus de 400 espèces bactériennes comprenant environ 10<sup>13</sup>-10<sup>14</sup> microorganismes (290,291). À travers l'étude de la variation génétique, le microbiome décrit l'écologie fragile qui existe entre la communauté des microbes symbiotiques, commensaux et pathologiques et notre espace corporel partagé. Potentiellement, cette connaissance pourrait être utilisée pour fournir des tests de diagnostic rapides, prévoir l'efficacité des antibiotiques et surveiller les environnements pour les contaminants (292, 293). La microbiomique a été utilisée pour identifier des infections virales inconnues ou pour diagnostiquer la résistance aux antibiotiques dans des infections telles que *Staphylococcus aureus*

résistant à la méthicilline (294, 295). De plus, plusieurs maladies ont été associées à des déséquilibres à grande échelle dans le microbiome intestinal, y compris la maladie inflammatoire de l'intestin, la diarrhée résistante aux antibiotiques et l'obésité (296, 297). Le Projet sur le microbiome humain est conçu pour comprendre ces interactions sur le plan physiologique et pathologique, en créant non seulement une base de données contenant des pathogènes, mais aussi un ensemble de contrôle de la flore normale (298). Au fur et à mesure que notre conscience de l'équilibre délicat entre l'hôte et les micro-organismes continue de croître, les applications du microbiome vont continuer de croître.



**Figure 2**

Gene expression signatures diagnose influenza and other respiratory viral infections in humans. (a) A microarray-based gene expression signature can accurately differentiate between symptomatic and asymptomatic individuals following exposure to rhinovirus, respiratory syncytial virus (RSV), and influenza. In the key, black represents the preinoculation baseline (healthy), red represents symptomatic postinoculation, and blue represents asymptomatic postinoculation. (b) A microarray-based gene expression signature can also distinguish individuals with bacterial infection from those with viral infection. The region above the black line is where we predict a high probability of viral infection; the region below is where the probability is low. Figure adapted from Reference 241.

**Figure 6 : Les signatures d'expression génique permettent de diagnostiquer la grippe et d'autres infections virales respiratoires chez les humains. (Référence 280)**

## **5. Intégration de la génomique dans la pratique clinique**

Plus l'information génomique est utilisée pour amplifier nos connaissances sur le mécanisme et le traitement de la maladie, mieux elle expose où les lacunes résident dans la méthodologie, l'infrastructure et la politique pour sa mise en œuvre en médecine clinique. L'abondance de la recherche utilisant des outils génomiques et la prospérité des données dont nous disposons nous obligent à faire des analyses techniques de ces résultats pour suivre leur production. Plusieurs suggestions ont été faites pour créer un cadre pour cette analyse, qui devrait impliquer ce qui suit pour chaque test: capacités techniques et normes opérationnelles, capacités de diagnostic et impact sur la prise de décision clinique, méthodes d'intégration comportementale et cognitive des médecins, rapport coût-efficacité, et les résultats pour la santé (299).

Une méta-analyse récente a révélé des zones de déficience dans le parcours de la clinique au chevet du patient, en commençant par la main-d'œuvre en soins primaires qui est à l'avant-garde de l'interaction avec le patient (300). Les médecins ne se sentent souvent pas préparés à administrer la médecine génomique, et les patients eux-mêmes ne comprennent pas clairement la valeur et l'intégration des tests génomiques dans leurs soins médicaux. Ce manque d'éducation s'accompagne d'un besoin accru d'évaluations des résultats pour les interventions génomiques, conduisant à une intégration limitée de la médecine génomique dans la pratique réelle. Avec ces obstacles identifiés, comment devrions-nous procéder?

## **5.1. Education**

L'un des principaux obstacles à la traduction de l'information génomique en pratique clinique est le manque de compréhension de la génomique par le fournisseur et l'accès limité aux outils appropriés. Plusieurs enquêtes suggèrent une déficience croissante des connaissances en génétique chez les médecins en moyenne (301, 302, 303). En plus de se familiariser avec les progrès rapides de la recherche génomique, les fournisseurs doivent également comprendre comment interpréter les résultats. Un baromètre direct de ceci est trouvé avec l'histoire de famille: Malgré le fait que 87% des médecins reconnaissent son importance, les médecins de premier recours se sont efforcés d'interpréter les schémas familiaux de la maladie et de mettre en œuvre l'information sur les antécédents familiaux dans leurs processus de travail (304, 305). Il est impératif d'améliorer la formation médicale entourant la médecine génétique et génomique pour le domaine et son intégration dans la pratique clinique pour avancer (306).

## **5.2. Système de soutien à la décision clinique**

SSD doit être utilisé pour promouvoir et intégrer les outils génomiques dans le flux de travail clinique. Dans les essais contrôlés randomisés, il a été démontré que les interventions de la SSD amélioraient significativement les soins prodigués aux patients et pouvaient avoir des implications significatives pour l'administration de recommandations basées sur la génomique aux médecins occupés (307, 308, 309). Si elle est élargie pour inclure les prévisions de risque pour la santé basées sur l'information génomique, des recommandations individualisées et en temps réel sur la pharmacogénomique ou les profils de risque pourraient être fournies sur la base d'une recherche à jour.

Le développement d'une infrastructure centrale et nationale pour la SSD et d'autres ressources génomiques, avec accès à l'information génomique et non génomique des patients, est essentiel pour toute mise en œuvre future (309). L'accès des médecins aux technologies de la SSD sera important pour son utilisation cohérente. De plus, ces systèmes ont eux-mêmes besoin de tests fondés sur des données probantes pour s'assurer qu'ils améliorent la prestation des soins de santé.

### **5.3. Vie privée**

À mesure que l'information génétique croît dans le contexte de la prise de décision clinique, les patients s'inquiètent progressivement de la discrimination génétique (310, 311). Au niveau fédéral, en 2005, le Congrès américain a adopté la Loi sur la transférabilité et la responsabilité en matière d'assurance maladie, qui prévoyait une protection contre la discrimination génétique pour les régimes de santé collectifs. La loi de 2008 sur la non-discrimination génétique a été adoptée pour fournir une protection supplémentaire (312). Plus de 30 états ont fourni leur propre législation traitant de cette question, créant une mosaïque de protection contre l'utilisation de l'information génétique par les assureurs pour discriminer les patients (313). Cependant, il faudra également élargir l'accès à l'éducation pour les patients, y compris des informations sur les composants génétiques spécifiques de leurs maladies et sur la façon dont ils affectent les droits au traitement et à la vie privée. Élargir le rôle des conseillers en génétique dans l'éducation des patients comblerait ce besoin et allégerait grandement le fardeau du système de soins de santé.

#### **5.4. Politique réglementaire**

La réglementation des produits pharmaceutiques à base de génomique et le niveau de rigueur auquel ils sont testés constituent un autre domaine de politique nécessaire (314). Ici, la FDA a été proactive, comme l'illustrent les plus de 200 étiquettes de produits qui pointent vers l'influence de la variation génétique sur la réponse aux médicaments (315). Ses efforts récents pour influencer les tests génétiques directement auprès des consommateurs, qui permettent aux consommateurs d'accéder à la variation génétique au sein de leur génome se rapportant à la maladie, sont peut-être encore plus remarquables (316). Bien que certains aient fait valoir que la validation adéquate fait défaut, il est clair qu'il existe un marché pour de telles entreprises (317). Cependant, la valeur clinique de leurs résultats n'a pas été définie, et la question à laquelle il faut répondre est, Quelle est la valeur de ces tests pour fournir aux cliniciens des données utilisables? Pour répondre à ces préoccupations, en juin 2010, la FDA a annoncé qu'elle commencerait à réglementer les tests directs aux consommateurs en tant qu'instrumentation médicale (318). Reste à savoir si ces mesures vont freiner le progrès scientifique ou protéger les consommateurs de la confusion par des informations cliniquement non pertinentes.

#### **5.5. Normes**

La réglementation croissante des outils génomiques dans la recherche et la clinique souligne le besoin de normes nationales de pratique pour toutes les infrastructures génomiques, des approches statistiques à la conception d'études cliniques et même à une infrastructure de bio banque qui fournit aux chercheurs des données provenant de multiples études. Les référentiels centralisés et normalisés pour le stockage, l'enregistrement et l'annotation d'échantillons

biologiques sont parmi les ressources les plus favorables à la médecine génomique (319). Des lignes directrices sur les pratiques exemplaires ont été publiées depuis 2005 pour orienter les efforts de bio banques, et des efforts ont déjà été déployés au niveau national pour normaliser la collecte d'échantillons biologiques: Le plan directeur du Réseau national de prélèvements biologiques a été proposé en 2005 et piloté par le Programme pilote de réseau national de bio spécificités du Programme spécialisé de recherche sur la prostate de l'Institut National du Cancer pour stocker des échantillons biologiques dans de grandes bio banques en réseau (320, 321). Les questions d'accessibilité devraient également être abordées, en accordant aux chercheurs extérieurs un accès gratuit à l'information tout en protégeant la propriété originale de l'échantillon. L'éthique de l'utilisation d'un large consentement du patient pour stocker les échantillons et les utiliser est toujours discutée, mais tous les efforts doivent être faits pour protéger la vie privée des patients et séparer les données sur l'identité des patients des données de chaque échantillon (322).

De plus, des normes communautaires sont nécessaires dans les approches statistiques et mathématiques des données génomiques. En raison du manque de grandes populations, des études statistiquement puissantes pour valider les résultats de la recherche sur les maladies chroniques sont difficiles à concevoir. Tous les efforts devraient être faits pour tester des modèles basés sur la génomique dans des essais cliniques, conçus avec des contrôles appropriés, avant que ces modèles soient introduits dans la clinique. La reproductibilité des résultats doit être démontrée avant leur mise en œuvre clinique (323). La randomisation et l'aveuglement sont également essentiels pour éliminer le biais des observateurs. Étant donné que le risque de faux positifs est élevé, des normes de contrôle et de normalisation doivent être en place pour chaque étude respective afin de pouvoir comparer les résultats.



## **5.6. Recherche sur l'efficacité comparative**

Étant donné le paysage actuel de la pratique fondée sur des données probantes, à quelle norme les approches génomiques de la médecine devraient-elles être appliquées? Pour résoudre ces problèmes, la recherche sur l'efficacité comparative (REC) est devenue une stratégie pour évaluer systématiquement les données à mesure que les résultats de la recherche entrent dans la pratique clinique. En 2009, 1,1 milliard de dollars de l'American Recovery and Reinvestment Act ont été affectés à la REC.

La REC est défini par l'Institute of Medicine comme «la génération et la synthèse de preuves qui comparent les avantages et les inconvénients des méthodes alternatives pour prévenir, diagnostiquer, traiter et surveiller un état clinique ou pour améliorer la prestation des soins, à la fois au niveau individuel et au niveau de la population "(324). En substance, la REC comprend une comparaison directe entre deux approches de soins et leur efficacité dans des contextes réels (325). Bien que certains ont fait valoir que la REC pourrait être utilisé à mauvais escient pour entraver l'avancement médical, la mise en œuvre de la médecine génomique exigera des preuves que la REC peut fournir, ce qui, nous l'espérons, encouragera l'innovation et réduira les coûts des soins de santé (326, 327). En comparant l'efficacité des outils génomiques dans le diagnostic, le pronostic, le traitement et la prédiction du risque par rapport aux normes d'or dans les soins actuels, des preuves peuvent être générées pour démontrer ou nier l'efficacité de plusieurs approches.

## **5.7. Propriété intellectuelle**

Le développement rapide des stratégies basées sur la génomique nécessite un investissement important et une coopération entre les intérêts publics et

privés. Malheureusement, au fur et à mesure que cette recherche se poursuit, l'insécurité entoure la propriété intellectuelle. Une question qui a retenu l'attention des juristes est la capacité des organisations, des entreprises et même des gouvernements à détenir des brevets sur des gènes nouvellement découverts. En 2009, un procès a été intenté au nom de l'American Civil Liberties Union et de la Public Patent Foundation pour invalider les brevets détenus sur les gènes BRCA1 et BRCA2, leur principal argument étant que ces gènes sont des produits de la nature (328). Les plaignants ont soutenu que les brevets (détenus par Myriad Genetics, l'Université de l'Utah Research Corporation et le gouvernement des États-Unis) entravaient la recherche. Fin mars 2010, un juge fédéral a révoqué des brevets sur les gènes BRCA1 et BRCA2 et potentiellement invalidé des centaines d'autres, confirmant qu'ils étaient des produits de la nature et non, comme le soutenaient les compagnies, de l'ADN isolé humain (329). Un appel a été déposé et ces contestations juridiques posent des questions intéressantes sur la façon dont la recherche en génomique devrait se dérouler.

## **5.8. Remboursement**

Si la médecine personnalisée doit être largement pratiquée, il est essentiel que les assureurs de santé, privés et gouvernementaux, disposés à fournir une couverture pour les tests génétiques, se dirigent vers l'intégration (330). À ce stade, les assureurs-maladie ont refusé de le faire parce qu'ils n'avaient pas reçu d'incitatifs qui les aideraient à obtenir des avantages à long terme en matière de coûts. Cependant, des progrès récents ont été réalisés pour encourager la médecine personnalisée: De nouveaux modèles ont montré une méthode de coordination de la prévention avec des résultats de maladie favorables et des coûts réduits (331). Comme les preuves renforcent l'importance de prévenir la

maladie en plus de la traiter, un nouveau rôle pourrait émerger pour la médecine génomique qui pourrait être soutenu par les compagnies d'assurance. Le gouvernement fédéral fournit une motivation supplémentaire, qui a joué un rôle accru dans l'introduction d'initiatives législatives visant à encourager l'adoption de la médecine personnalisée (332).

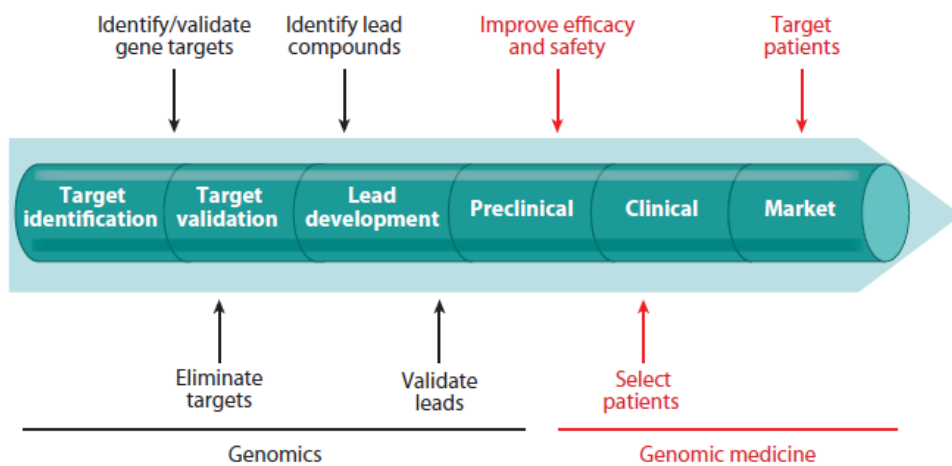
## **6. Quand la "médecine personnalisée" est "médecine" ?**

Il semble que les principaux obstacles se situent dans trois catégories: la collecte d'informations adéquates et appropriées sur les patients, le dosage de ces informations pour qu'elles deviennent cliniquement pertinentes et l'interprétation de ces tests par les professionnels de la santé en charge des soins aux patients.

Malgré les défis qui nous attendent pour l'intégration complète de la médecine personnalisée dans la prise de décision clinique, une revue du paysage de la pratique médicale suggère que la médecine personnalisée est arrivée. Bien qu'il soit facile de se demander pourquoi les progrès scientifiques de la recherche génomique ne se sont pas exactement traduits par les résultats révolutionnaires attendus en médecine, il serait stupide de négliger l'impact profond qu'elle a déjà fait. De plus, l'effet subjectif perçu peut être attribué à des technologies en retard et à des approches analytiques combinées à une lutte pour mettre en œuvre des résultats éprouvés en clinique. Bien que des améliorations puissent être apportées à la base de connaissances et à l'infrastructure, il est également important de garder une perspective historique: Les progrès de la médecine fondés sur une pensée scientifique novatrice prennent beaucoup de temps et d'investissements. Un rappel important de ce fait est l'histoire de 30 ans du paclitaxel, qui a débuté en 1962 avec la collecte d'écorce d'arbre et qui n'est entrée dans les essais cliniques de phase I qu'en 1984 (333). En moyenne, il faut

17 ans pour que 14% de la recherche scientifique soit incorporée dans la pratique clinique quotidienne (334, 335). L'utilisation de la génomique pour comprendre la biologie de la maladie et sélectionner des cibles pour le développement de médicaments et l'utilisation de la médecine génomique pour concevoir des essais cliniques et sélectionner des patients pour des thérapies ciblées devraient augmenter le taux de réussite et réduire le temps de développement de médicaments (Figure7).

Comme Kuhn l'affirme clairement, les changements dans la pensée scientifique ne peuvent se produire que par ce qui doit sembler être des exemples disjoints, chacun ayant son propre impact. Mais s'ils sont pris ensemble, il devient clair que la médecine personnalisée n'était pas une promesse faite à la hâte, ni un objectif inaccessible; l'ensemble des preuves concernant l'efficacité de l'utilisation de la médecine génomique pour prescrire un traitement individualisé continue de croître. Bien que l'investissement dans la médecine génomique ne soit pas trivial, il fournit la base et les méthodes systématiques pour les progrès futurs. Plus de travail est nécessaire, plus de réflexion est nécessaire, et plus de lignes directrices doivent être établies, mais nous sommes constamment en train de traverser dans un nouveau paradigme où l'optimisation de la pratique médicale peut s'adresser spécifiquement à chaque patient nécessitant un traitement.



**Figure 3**

The pathway for identifying and validating drug targets for pharmacological therapies, showing a schematic representation of the process from biomarker discovery to a marketable drug. Genomic studies can potentially reduce the time it takes for drug discovery through enhanced understanding of the underlying biology and the discovery of relevant biomarkers; genomic medicine can enable drug development and improve efficacy and safety by identifying and targeting clinical studies to patient subsets most likely to respond to the drug.

**Figure 7 : La voie pour l'identification et la validation des cibles médicamenteuses pour les thérapies pharmacologiques, montrant une représentation schématique du processus de la découverte de biomarqueurs à un médicament commercialisable.**



## *V. Conclusion*

Alors que les progrès technologiques ont accéléré l'identification des biomarqueurs de la susceptibilité, du traitement et de la progression des maladies, le défi consiste à transporter ces découvertes vers les soins cliniques. Si la promesse de la génomique doit être réalisée, nous devons tester / adopter des méthodes scientifiques pour documenter comment ces technologies pourraient améliorer les soins de santé et prévenir les maladies. Les voies par lesquelles les marqueurs génomiques doivent voyager pour aider les cliniciens comprennent quatre phases: 1) Découverte, élucidation des voies et réplication chez les populations indépendantes, 2) Validation de l'efficacité et utilité clinique chez les populations à risque, 3) Exigences réglementaires pour l'approbation de la dispensation clinique, et 4) Démonstration de l'impact, de l'efficacité et des aspects sanitaires l'intégrité individuelle et la protection contre la discrimination génomique, inégalement respectée dans de nombreuses juridictions.

L'établissement de relations entre de nombreux marqueurs génomiques et des phénotypes cliniques est très difficile. À grande échelle, les études cliniques prospectives sont difficiles mais très nécessaires pour démontrer l'utilité du génotypage comme base pour la médecine individualisée. Il convient toutefois de noter que les tests génétiques peuvent être utiles sans une connaissance complète de la fonction biologique des variants génétiques en question. Une stratégie innovante qui relie de nombreux résultats cliniques aux marqueurs génomiques pléiotropiques inclut le PheWAS récemment proposé, une étude d'association à l'échelle du phénomène utilisée pour découvrir de nouvelles associations de maladies géniques. Ce qui est essentiel, c'est un programme multidisciplinaire de recherche en traduction qui englobe des recherches

cliniques et axées sur la population et le développement d'une collaboration entre les intervenants pour faciliter la traduction fondée sur des données probantes et le remboursement des tests. Alors que la médecine personnalisée entre dans la pratique médicale traditionnelle, les médecins et autres prestataires de soins devront administrer ou donner des conseils sur l'application d'un nombre croissant de tests génomiques, de thérapies pharmacogénomiques et de décisions thérapeutiques basées sur des preuves prédictives et l'estimation des risques. Les insuffisances des programmes de génétique dans les écoles de médecine pourraient entraver l'adoption de nouvelles connaissances génétiques. Cependant, plusieurs programmes ont récemment vu le jour qui préparent la prochaine génération de médecins, d'infirmières, de pharmaciens et d'autres travailleurs de la santé à la vague de médecine génomique à venir.

La perception du public est également très importante, car l'adoption clinique de nombreuses interventions génomiques proposées dépend de la compréhension du public. La médecine participative est un nouveau scénario où les patients deviennent des participants actifs dans leurs propres soins médicaux. L'environnement, le mode de vie, l'alimentation, les antécédents familiaux et les observations des patients sur leurs propres symptômes doivent être combinés aux données des biomarqueurs pour fournir une vision globale et intégrée des facteurs influençant le risque, la progression et les résultats du traitement. Les individus peuvent maintenant obtenir leur information génomique personnelle par le biais de tests génétiques directement auprès des consommateurs, mais quel impact, le cas échéant, cela aurait-il sur leur mode de vie et leur santé par rapport à un effet délétère? Jusqu'à présent, quelques études ont été réalisées, avec des preuves encourageantes mais limitées suggérant que l'information



génétiq ue pourrait entraîner des changements dans les comportements liés à l'hygiène de vie, comme l'abandon du tabac. Des études plus approfondies de cette question sont nécessaires mais, surtout, la réalisation des soins de santé personnalisés, tout en s'appuyant sur la connaissance génomique de l'individu, requiert des informations environnementales tout aussi essentielles et la compréhension de la capacité de chaque individu à promouvoir sa santé. Tous ces éléments sont importants pour promouvoir les quatre «P» de l'avenir de la médecine: la médecine prédictive, préventive, personnalisée et participative.



## *VI. Résumé*

## Résumé

**Titre:** Le séquençage de l'ADN et médecine personnalisée

**Auteur:** Mohamed En-nouali

**Most-clés:** Génomique- NGS–Médecine individualisée-Médecine personnalisée

L'évolution technique du séquençage des ADN s'est faite avec une rapidité exponentielle au cours des dernières années. La diffusion dans les laboratoires hospitaliers des nouveaux appareils de séquençage à haut débit (ou NGS) permet de passer de l'étude ponctuelle d'un gène à une analyse globale portant sur des dizaines ou centaines de gènes ou sur l'ensemble de l'exome ; elle modifie de ce fait les concepts médicaux et les conditions d'exercice, notamment en génétique et en cancérologie. Cette possibilité d'une recherche simultanée de mutations dans la séquence d'un grand nombre de gènes trouve aujourd'hui ses applications dans le diagnostic, éventuellement néonatal, dans le dépistage systématique des hétérozygotes et le diagnostic préconceptionnel généralisé. En oncogénétique constitutionnelle, les NGS réalisent également l'analyse simultanée des gènes impliqués dans les formes « héréditaires » de cancers. Par ailleurs dans l'ensemble des affections malignes, les NGS identifient l'ensemble des anomalies génomiques affectant le tissu malade, en s'efforçant de faire la distinction entre mutations « drivers » importantes pour la progression tumorale et mutations « passagers » ou accessoires; en s'efforçant également d'identifier des mutations « druggable » susceptibles de faire l'objet de thérapies ciblées. Enfin le séquençage systématique de l'ensemble des gènes impliqués dans le métabolisme et la réponse aux médicaments est le support d'une pharmacogénétique individuelle. En définitive, le séquençage des génomes tumoraux et constitutionnels, l'identification des mutations somatiques, la détermination des variants pharmacogénétiques, sont les éléments déterminants d'une médecine dite personnalisée. Les premiers résultats de ces indications thérapeutiques ciblées montrent un gain en termes de durée de rémission et de survie; reste néanmoins posée la question du coût de ces traitements. Ces énormes capacités de séquençage des génomes soulèvent également un grand nombre de questions réglementaires et éthiques.

## Summary

**Title:** DNA sequencing and personalized medicine

**Author:** Mohamed En-nouali

**Keywords:** Genomics–Next Generation Sequencing (NGS)-Individualized medicine–Personalized medicine.

DNA sequencing technologies have advanced at an exponential rate in recent years. More and more hospital laboratories are acquiring new high-throughput sequencing devices, (“ next-generation sequencers ”, NGS), allowing them to analyze tens or hundreds of genes, or even the entire exome. This is having a major impact on medical concepts and practices, especially with respect to genetics and oncology. This ability to search for mutations simultaneously in a large number of genes is finding applications in the diagnosis of Mendelian diseases (including at birth), routine screening for heterozygotes, and preconception diagnosis. Data on multifactorial diseases are still preliminary, but it should soon be possible to identify “strong ” factors of genetic predisposition that have so far been beyond the scope of genome-wide association studies (GWAS). In the field of constitutional oncogenetics, NGS can also be used for simultaneous analysis of genes involved in “hereditary ” cancers. More generally, NGS can identify all genomic abnormalities in a given malignant tissue (hemopathy or solid tumor), and has the potential to distinguish between important mutations (those that drive tumor progression) from “bystander ” or accessory mutations, and also to identify “ druggable ” mutations amenable to targeted therapies. Systematic sequencing of all the genes involved in drug metabolism and responsiveness will lead to individualized pharmacogenetics (warfarin therapy). Finally, sequencing of the tumoral and constitutional genomes, identification of somatic mutations, and detection of pharmacogenetic variants will open up the era of personalized medicine. The first results of these targeted therapeutic indications show a gain in the duration of remission and survival, although the cost-effectiveness of these approaches remains to be determined. Finally, this huge capacity for genome sequencing raises a number of regulatory and ethical issues.

# ملخص

العنوان: تسلسل الحمض النووي والطب الشخصي

من طرف: محمد النوالي

الكلمات الأساسية : علم الجينوم- الجيل الجديد التسلسل- الطب الفردي- الطب الشخصي

وقد تم التطور التقني لتسلسل الحمض النووي بسرعة هائلة في السنوات الأخيرة. المزيد والمزيد من مختبرات المستشفيات تكتسب جديد الأجهزة الإنتاجية عالية التسلسل ("الجيل القادم من التسلسل"، ج.ق.ت). مما يسمح لهم بتحليل عشرات أو مئات من الجينات، أو حتى إكسوم بأكمله. هذا له تأثير كبير على المفاهيم والممارسات الطبية، وخاصة فيما يتعلق علم الوراثة والأورام. هذه القدرة على البحث عن الطفرات في وقت واحد في عدد كبير من الجينات تم العثور لها على تطبيقات في تشخيص الأمراض المندلية، الفحص الروتيني للمتخالف، والتشخيص ما قبل الحمل. البيانات حول الأمراض متعددة العوامل لا تزال أولية، ولكن من الممكن قريباً تحديد العوامل "القوية" من الاستعداد الوراثي التي كانت حتى الآن خارج نطاق دراسات الارتباط على نطاق الجينوم (غواس). في مجال علم الوراثة السرطانية، ويمكن أيضاً أن تستخدم ج.ق.ت للتحليل في وقت واحد من الجينات التي تشارك في السرطان "الوراثي" (21 الجينات سرطان الثدي، 6 القولون وجينات السرطان، وما إلى ذلك). وبصورة أعم، يمكن أن يحدد ج.ق.ت جميع الشذوذ الجيني (الحذف، نقل المواقع، الطفرات) في نسيج خبيث معين (اعتلال المعدة أو الورم الصلب)، ولديه القدرة على التمييز بين الطفرات الهامة (تلك التي تدفع تطور الورم) من "المارة" أو الطفرات التبعية، وأيضاً لتحديد الطفرات "قابلة للتلف" قابلة للعلاجات المستهدفة (تراستوزوماب، وجيفيتينيب...). التسلسل المنهجي لجميع الجينات المشاركة في استقلاب الدواء والاستجابة يؤدي إلى علم الوراثة الدوائي الفردي. وأخيراً، فإن تسلسل الجينوم الورمي، وتحديد الطفرات الجسدية، والكشف عن فارماكوجينيتيك فتح عصر الطب الشخصي. النتائج الأولى لهذه المؤشرات العلاجية المستهدفة تظهر مكسباً في مدة الشفاء من المرض والبقاء على قيد الحياة، على الرغم من أن فعالية وتكلفة هذا النهج لا يزالان يتعين تحديده. وأخيراً، فإن هذه القدرة الضخمة على تسلسل الجينوم تثير عدداً من القضايا التنظيمية والأخلاقية.



## *VII. Bibliographie*

- [1] "deoxyribonucleic acid". Merriam-Webster Dictionary.
- [2] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2014). Molecular Biology of the Cell (6th ed.). Garland. p. Chapter 4: DNA, Chromosomes and Genomes. ISBN 9780815344322.
- [3] Purcell A. "DNA". Basic Biology.
- [4] Nuwer R (18 July 2015). "Counting All the DNA on Earth". The New York Times. New York: The New York Times Company. ISSN 0362-4331. Retrieved 2015-07-18.
- [5] "The Biosphere: Diversity of Life". Aspen Global Change Institute. Basalt, CO. Retrieved 2015-07-19.
- [6] Russell P (2001). iGenetics. New York: Benjamin Cummings. ISBN 0-8053-4553-1.
- [7] Mashaghi A, Katan A (2013). "A physicist's view of DNA". *De Physicus*. 24e (3): 59–61. Bibcode:2013arXiv1311.2545M. arXiv:1311.2545v1 [3](#).
- [8] Saenger W (1984). Principles of Nucleic Acid Structure. New York: Springer-Verlag. ISBN 0-387-90762-9.
- [9] <sup>a</sup> <sup>b</sup> Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Peter W (2002). Molecular Biology of the Cell (Fourth ed.). New York and London: Garland Science. ISBN 0-8153-3218-1. OCLC 145080076.

- [10] Irobalieva RN, Fogg JM, Catanese DJ, Catanese DJ, Sutthibutpong T, Chen M, Barker AK, Ludtke SJ, Harris SA, Schmid MF, Chiu W, Zechiedrich L (October 2015). "Structural diversity of supercoiled DNA". *Nature Communications*. 6: 8440. Bibcode: 2015NatCo...6E8440I. PMC 4608029 . PMID 26455586. doi: 10.1038/ncomms9440.
- [11] <sup>a b c</sup> Watson JD, Crick FH (April 1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid" (PDF). *Nature*. 171 (4356): 737–8. Bibcode:1953Natur.171..737W. PMID 13054692. doi:10.1038/171737a0.
- [12] Mandelkern M, Elias JG, Eden D, Crothers DM (October 1981). "The dimensions of DNA in solution". *Journal of Molecular Biology*. 152 (1): 153–61. PMID 7338906. doi:10.1016/0022-2836(81)90099-1.
- [13] Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, et al. (May 2006). "The DNA sequence and biological annotation of human chromosome 1". *Nature*. 441(7091): 315–21. Bibcode:2006Natur.441..315G. PMID 16710414. doi:10.1038/nature04727.
- [14] Watson JD, Crick FH (April 1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid" (PDF). *Nature*. 171 (4356): 737–8. Bibcode:1953Natur.171..737W. PMID 13054692. doi:10.1038/171737a0.
- [15] <sup>a b c</sup> Berg J., Tymoczko J. and Stryer L. (2002) *Biochemistry*. W. H. Freeman and Company ISBN 0-7167-4955-6



- [16] Abbreviations and Symbols for Nucleic Acids, Polynucleotides and their Constituents IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Retrieved 3 January 2006.
- [17] <sup>a b</sup> Ghosh A, Bansal M (April 2003). "A glossary of DNA structures from A to Z". *Acta Crystallographica. Section D, Biological Crystallography*. 59 (Pt 4): 620–6. PMID 12657780. doi:10.1107/S0907444903003251.
- [18] Created from PDB 1D65
- [19] Yakovchuk P, Protozanova E, Frank-Kamenetskii MD (2006). "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix". *Nucleic Acids Research*. 34 (2): 564–74. PMC 1360284 . PMID 16449200. doi:10.1093/nar/gkj454.
- [20] Burton E. Tropp – "Molecular Biology"- Jones and Barlett Learning, ISBN 978-0-7637-8663-2
- [21] "Watson-Crick Structure of DNA – 1953". Steven Carr. Memorial University of Newfoundland. Retrieved 13 July 2016.
- [22] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. (February 2001). "The sequence of the human genome". *Science*. 291 (5507): 1304–51. Bibcode:2001Sci...291.1304V. PMID 11181995. doi:10.1126/science.1058040.
- [23] Thanbichler M, Wang SC, Shapiro L (October 2005). "The bacterial nucleoid: a highly organized and dynamic structure". *Journal of Cellular Biochemistry*. 96 (3): 506–21. PMID 15988757. doi:10.1002/jcb.20519.

- [24] Wolfsberg TG, McEntyre J, Schuler GD (February 2001). "Guide to the draft human genome". *Nature*. 409 (6822): 824–6. Bibcode:2001Natur.409..824W. PMID 11236998. doi:10.1038/35057000.
- [25] Gregory TR (January 2005). "The C-value enigma in plants and animals: a review of parallels and an appeal for partnership". *Annals of Botany*. 95 (1): 133–46. PMID 15596463. doi:10.1093/aob/mci009.
- [26] Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, et al. (June 2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project". *Nature*. 447 (7146): 799–816. Bibcode:2007Natur.447..799B. PMC 2212820 . PMID 17571346. doi:10.1038/nature05874.
- [27] Created from PDB 1MSW
- [28] Pidoux AL, Allshire RC (March 2005). "The role of heterochromatin in centromere function". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 360 (1455): 569–79. PMC 1569473 . PMID 15905142. doi:10.1098/rstb.2004.1611.
- [29] Harrison PM, Hegyi H, Balasubramanian S, Luscombe NM, Bertone P, Echols N, Johnson T, Gerstein M (February 2002). "Molecular fossils in the human genome: identification and analysis of the pseudogenes in chromosomes 21 and 22". *Genome Research*. 12 (2): 272–80. PMC 155275 . PMID 11827946. doi:10.1101/gr.207102.

- [30] Harrison PM, Gerstein M (May 2002). "Studying genomes through the aeons: protein families, pseudogenes and proteome evolution". *Journal of Molecular Biology*. 318 (5): 1155–74. doi:10.1016/S0022-2836(02)00109-2. PMID 12083509.
- [31] Albà M (2001). "Replicative DNA polymerases". *Genome Biology*. 2 (1): REVIEWS3002. doi:10.1186/gb-2001-2-1-reviews3002. PMC 150442 . PMID 11178285.
- [32] Miescher, Friedrich (1871) "Ueber die chemische Zusammensetzung der Eiterzellen"(On the chemical composition of pus cells), *Medicinisch-chemische Untersuchungen*, 4 : 441–460. From p. 456: "Ich habe mich daher später mit meinen Versuchen an die ganzen Kerne gehalten, die Trennung der Körper, die ich einstweilen ohne weiteres Präjudiz als lösliches und unlösliches Nuclein bezeichnen will, einem günstigeren Material überlassend."(Therefore, in my experiments I subsequently limited myself to the whole nucleus, leaving to a more favorable material the separation of the substances, that for the present, without further prejudice, I will designate as soluble and insoluble nuclear material ("Nuclein").)
- [33] Dahm R (January 2008). "Discovering DNA: Friedrich Miescher and the early years of nucleic acid research". *Human Genetics*. 122 (6): 565–81. doi:10.1007/s00439-007-0433-0. PMID 17901982.
- [34] See:
- Albrecht Kossel (1879) "Ueber Nuclein der Hefe" (On nuclein in yeast) *Zeitschrift für physiologische Chemie*, 3 : 284–291.

- Albrect Kossel (1880) "Ueber Nuclein der Hefe II" (On nuclein in yeast, Part 2) *Zeitschrift für physiologische Chemie*, 4 : 290–295.
- Albrect Kossel (1881) "Ueber die Verbreitung des Hypoxanthins im Thier- und Pflanzenreich" (On the distribution of hypoxanthins in the animal and plant kingdoms) *Zeitschrift für physiologische Chemie*, 5 : 267–271.
- Albrect Kossel, *Untersuchungen über die Nucleine und ihre Spaltungsprodukte*[Investigations into nuclein and its cleavage products] (Strassburg, Germany: K.J. Trübner, 1881), 19 pages.
- Albrect Kossel (1882) "Ueber Xanthin und Hypoxanthin" (On xanthin and hypoxanthin), *Zeitschrift für physiologische Chemie*, 6 : 422–431.
- Albrect Kossel (1883) "Zur Chemie des Zellkerns" (On the chemistry of the cell nucleus), *Zeitschrift für physiologische Chemie*, 7 : 7–22.
- Albrect Kossel (1886) "Weitere Beiträge zur Chemie des Zellkerns" (Further contributions to the chemistry of the cell nucleus), *Zeitschrift für Physiologische Chemie*, 10 : 248–264. Available online at: Max Planck Institute for the History of Science, Berlin, Germany. On p. 264, Kossel remarked presciently: "Der Erforschung der quantitativen Verhältnisse der vier stickstoffreichen Basen, der Abhängigkeit ihrer Menge von den physiologischen Zuständen der Zelle, verspricht wichtige Aufschlüsse über die elementaren physiologisch-chemischen Vorgänge." (The study of the quantitative relations of the four nitrogenous bases — [and] of the dependence of their quantity on the physiological states of the cell — promises important insights into the fundamental physiological-chemical processes.)

- [35] Jones ME (September 1953). "Albrecht Kossel, a biographical sketch". The Yale Journal of Biology and Medicine. National Center for Biotechnology Information. 26 (1): 80–97. PMC 2599350 [3](#). PMID 13103145.
- [36] Levene P (1 December 1919). "The structure of yeast nucleic acid". J Biol Chem. 40 (2): 415–24.
- [37] See:
1. W. T. Astbury and Florence O. Bell (1938) "Some recent developments in the X-ray study of proteins and related structures," Cold Spring Harbor Symposia on Quantitative Biology, 6 : 109–121. Available on-line at: University of Leeds.
  2. Astbury, W. T., (1947) "X-ray studies of nucleic acids," Symposia of the Society for Experimental Biology, 1 : 66–76. Available on-line at: Oregon State University.
- [38] Koltsov proposed that a cell's genetic information was encoded in a long chain of amino acids. See:
1. Н. К. Кольцов, "Физико-химические основы морфологии" (The physical-chemical basis of morphology) – speech given at the 3rd All-Union Meeting of Zoologist, Anatomists, and Histologists at Leningrad, U.S.S.R., December 12, 1927.
  2. Reprinted in: Успехи экспериментальной биологии (Advances in Experimental Biology), series B, 7 (1) : ?-? (1928).
  3. Reprinted in German as: Nikolaj K. Koltzoff (1928) "Physikalisch-chemische Grundlagen der Morphologie" (The physical-chemical basis of morphology), Biologisches Zentralblatt, 48 (6) : 345–369.

4. In 1934, Koltsov contended that the proteins that contain a cell's genetic information replicate. See: N. K. Koltzoff (October 5, 1934) "The structure of the chromosomes in the salivary glands of *Drosophila*," *Science*, 80 (2075) : 312–313. From page 313: "I think that the size of the chromosomes in the salivary glands [of *Drosophila*] is determined through the multiplication of genonemes. By this term I designate the axial thread of the chromosome, in which the geneticists locate the linear combination of genes; ... In the normal chromosome there is usually only one genoneme; before cell-division this genoneme has become divided into two strands."
- [39] Soyfer VN (September 2001). "The consequences of political dictatorship for Russian science". *Nature Reviews. Genetics*. 2 (9): 723–9. doi:10.1038/35088598. PMID 11533721.
- [40] Griffith F (January 1928). "The Significance of Pneumococcal Types". *The Journal of Hygiene*. 27 (2): 113–59. doi:10.1017/S0022172400031879. PMC 2167760 . PMID 20474956.
- [41] Lorenz MG, Wackernagel W (September 1994). "Bacterial gene transfer by natural genetic transformation in the environment". *Microbiological Reviews*. 58 (3): 563–602. PMC 372978 . PMID 7968924.
- [42] Avery OT, Macleod CM, McCarty M (February 1944). "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type III". *The Journal of Experimental Medicine*. 79 (2): 137–58. doi:10.1084/jem.79.2.137. PMC 2135445 . PMID 19871359.

- [43] Hershey AD, Chase M (May 1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage". *The Journal of General Physiology*. 36 (1): 39–56. doi:10.1085/jgp.36.1.39. PMC 2147348 . PMID 12981234.
- [44] The B-DNA X-ray pattern on the right of this linked image was obtained by Rosalind Franklin and Raymond Gosling in May 1952 at high hydration levels of DNA and it has been labeled as "Photo 51"
- [45] Regis, Ed (2009) *What Is Life?: investigating the nature of life in the age of synthetic biology*. Oxford: Oxford University Press ISBN 0-19-538341-9; p. 52
- [46] Nature Archives Double Helix of DNA: 50 Years
- [47] "Original X-ray diffraction image". [Osulibrary.oregonstate.edu](http://osulibrary.oregonstate.edu). Retrieved 6 February 2011.
- [48] The Nobel Prize in Physiology or Medicine 1962 [Nobelprize .org](http://Nobelprize.org) Accessed 22 December 06
- [49] Maddox B (January 2003). "The double helix and the 'wronged heroine'" (PDF). *Nature*. 421 (6921): 407–8. Bibcode:2003Natur.421..407M. doi:10.1038/nature01399. PMID 12540909.
- [50] Crick, F.H.C. On degenerate templates and the adaptor hypothesis (PDF). Archived 1 October 2008 at the Wayback Machine. [genome.wellcome.ac.uk](http://genome.wellcome.ac.uk) (Lecture, 1955). Retrieved 22 December 2006.

- [51] Meselson M, Stahl FW (July 1958). "THE REPLICATION OF DNA IN ESCHERICHIA COLI". *Proceedings of the National Academy of Sciences of the United States of America*. 44 (7): 671–82. Bibcode:1958PNAS...44..671M. doi:10.1073/pnas.44.7.671. PMC 528642 . PMID 16590258.
- [52] The Nobel Prize in Physiology or Medicine 1968 Nobelprize.org Accessed 22 December 06
- [53] Pray, L. (2008) Discovery of DNA structure and function: Watson and Crick. *Nature Education* 1(1):100
- [54] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 1977;74(12):5463–7.
- [55] Kircher M, Kelso J. High-throughput DNA sequencing – concepts and limitations. *Bioessays* 2010;32(6):524–36.
- [56] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822): 860–921.
- [57] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science* 2001;291(5507):1304–51.
- [58] Finishing the euchromatic sequence of the human genome. *Nature* 2004; 431(7011):931–45.



- [59] Maxam AM, Gilbert W. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America* 1977;74(2):560–4.
- [60] Metzker ML. Sequencing technologies – the next generation. *Nature Reviews Genetics* 2010;11(1):31–46.
- [61] Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. *Human Molecular Genetics* 2010;19(R2):R227–40.
- [62] Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. *Science* 2009;323(5910):133–8.
- [63] Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, et al. The challenges of sequencing by synthesis. *Nature Biotechnology* 2009;27(11):
- [64] Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437(7057):376–80.
- [65] Ronaghi M, Uhlen M, Nyren P. A sequencing method based on real-time pyrophosphate. *Science* 1998;281(5375):363–5.
- [66] Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P. Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry* 1996;242(1):84–9.
- [67] Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008;452(7189):872–6.

- [68] Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 2008;456(7218):53–9.
- [69] Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Research* 2008;18(7):1051–63.
- [70] Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, et al. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 2005;309(5741):1728–32.
- [71] Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 2011;475(7356):348–52.
- [72] Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, et al. Singlemolecule DNA sequencing of a viral genome. *Science* 2008;320(5872):106–9.
- [73] Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 2010;327(5961):78–81.
- [74] Kasianowicz JJ, Brandin E, Branton D, Deamer DW. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences of the United States of America* 1996;93(24): 13770–73.

- [75] Bayley H. Sequencing single molecules of DNA. *Current Opinion in Chemical Biology* 2006;10(6):628–37.
- [76] Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, et al. The potential and challenges of nanopore sequencing. *Nature Biotechnology* 2008;26(10):1146–53.
- [77] Derrington IM, Butler TZ, Collins MD, Manrao E, Pavlenok M, Niederweis M, et al. Nanopore DNA sequencing with MspA. *Proceedings of the National Academy of Sciences of the United States of America* 2010;107(37):16060–65.
- [78] Schneider GF, Dekker C. DNA sequencing with nanopores. *Nature Biotechnology* 2012;30(4):326–8.
- [79] Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology* 2009;4(4):265–70.
- [80] Wu HC, Astier Y, Maglia G, Mikhailova E, Bayley H. Protein nanopores with covalently attached molecular adapters. *Journal of the American Chemical Society* 2007;129(51):16142–48.
- [81] Astier Y, Braha O, Bayley H. Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *Journal of the American Chemical Society* 2006;128(5):1705–10.

- [82] Stoddart D, Heron AJ, Klingelhofer J, Mikhailova E, Maglia G, Bayley H. Nucleobase recognition in ssDNA at the central constriction of the  $\alpha$ -hemolysin pore. *Nano Letters* 2010;10(9):3633–7.
- [83] Manrao EA, Derrington IM, Laszlo AH, Langford KW, Hopper MK, Gillgren N, et al. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nature Biotechnology* 2012;30(4):349–53.
- [84] Cherf GM, Lieberman KR, Rashid H, Lam CE, Karplus K, Akeson M. Automated forward and reverse ratcheting of DNA in a nanopore at 5-A precision. *Nature Biotechnology* 2012;30(4):344–8.
- [85] Garaj S, Hubbard W, Reina A, Kong J, Branton D, Golovchenko JA. Graphene as a subnanometre trans-electrode membrane. *Nature* 2010;467(7312):190–3.
- [86] Dekker C. Solid-state nanopores. *Nature Nanotechnology* 2007;2(4):209–15.
- [87] Banerjee A, Mikhailova E, Cheley S, Gu LQ, Montoya M, Nagaoka Y, et al. Molecular bases of cyclodextrin adapter interactions with engineered protein nanopores. *Proceedings of the National Academy of Sciences of the United States of America* 2010;107(18):8165–70.
- [88] Larsson C, Grundberg I, Söderberg O, Nilsson M. In situ detection and genotyping of individual mRNA molecules. *Nature Methods* 2010;7(5):395–7.

- [89] Aston C, Mishra B, Schwartz DC. Optical mapping and its potential for largescale sequencing projects. *Trends in Biotechnology* 1999;17(7):297–302.
- [90] Westerlund F, et al. Fluorescence enhancement of single DNA molecules con-fined in Si/SiO<sub>2</sub> nanochannels. *Lab Chip* 2010;10(16):2049–51.
- [91] Guttmacher AE, Collins FS, Carmona RH. 2004. The family history: more important than ever.
- [92] Rich E, BurkeW, Heaton C, Haga S, Pinsky L, et al. 2004. Reconsidering the family history in primary care. *J. Gen. Intern. Med.* 19:273–80
- [93] Ginsburg GS, Willard HF. 2009. Genomic and personalized medicine: foundations and applications. *Transl. Res.* 154:277–87
- [94] KannelW, Dawber T, Kagan A, Revotskie N, Stokes J III. 1961. Factors of risk in the development of coronary heart disease: six year follow-up experience: the Framingham Study. *Ann. Intern.Med.* 55:33–50
- [95] Gail M, Greene M. 2000. Gail model and breast cancer. *Lancet* 355:1017
- [96] Liu J, Wyatt J, Altman D. 2006. Decision tools in health care: focus on the problem, not the solution. *BMC Med. Inform. Decis. Mak.* 6:4
- [97] Osheroff JA, Teich JM, Middleton B, Steen EB, Wright A, Detmer DE. 2007. A roadmap for national action on clinical decision support. *J. Am. Med. Inform. Assoc.* 14:141–45

- [98] Bennett JW, Glasziou PP. 2003. Computerised reminders and feedback in medication management: a systematic review of randomised controlled trials. *Med. J. Aust.* 178:217–22
- [99] Shea S, DuMouchel W, Bahamonde L. 1996. A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting. *J. Am. Med. Inform. Assoc.* 3:399–409
- [100] Walton R, Dovey S, Harvey E, Freemantle N. 1999. Computer support for determining drug dose: systematic review and meta-analysis. *BMJ* 318:984–90
- [101] Collins FS. 1999. Shattuck lecture: medical and societal consequences of the Human Genome Project. *N. Engl. J. Med.* 341:28–37
- [102] Abrahams E, Silver M. 2009. The case for personalized medicine. *J. Diabetes Sci. Technol.* 3:680–84
- [103] Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, et al. 2009. Repeatability of published microarray gene expression analyses. *Nat. Genet.* 41:149–55
- [104] Secr. Advis. Comm. Genet. Health Soc. 2008. U.S. system of oversight of genetic testing: a response to the charge of the secretary of health and human services. Rep., Dep. Health Hum. Serv., Washington, DC
- [105] Ginsburg GS. 2011. Application of molecular technologies to clinical medicine. In *Cecil Medicine*, ed. L Goldman, AI Schafer, pp. 199–203. Philadelphia: Elsevier. 24th ed.

- [106] Willard H, Angrist M, Ginsburg G. 2005. Genomic medicine: genetic variation and its impact on the future of health care. *Philos. Trans. R. Soc. B* 360:1543–50
- [107] Weinshilboum R, Wang L. 2004. Pharmacogenomics: bench to bedside. *Nat. Rev. Drug Discov.* 3:739–48
- [108] Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, et al. 2009. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N. Engl. J. Med.* 360:753–64
- [109] Feero WG, Guttmacher AE, Collins FS. 2010. Genomic medicine: an updated primer. *N. Engl. J. Med.* 362:2001–11
- [110] Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, et al. 2006. Global variation in copy number in the human genome. *Nature* 444:444–54
- [111] Bonetta L. 2010. Whole-genome sequencing breaks the cost barrier. *Cell* 141:917–19
- [112] Bentley DR. 2006. Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* 16:545–52
- [113] Service RF. 2006. The race for the \$1000 genome. *Science* 311:1544–46
- [114] Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, et al. 2007. The diploid genome sequence of an individual human. *PLoS Biol.* 5:e254
- [115] Pushkarev D, Neff NF, Quake SR. 2009. Single-molecule sequencing of an individual human genome *Nat. Biotechnol.* 27:847–50

- [116] Wang J, Wang W, Li R, Li Y, Tian G, et al. 2008. The diploid genome sequence of an Asian individual. *Nature* 456:60–65
- [117] Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, et al. 2008. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452:872–76
- [118] Bolze A, Byun M, McDonald D, Morgan NV, Abhyankar A, et al. 2010. Whole-exome-sequencing-based discovery of human FADD deficiency. *Am. J. Hum. Genet.* 87:873–81
- [119] Dahl F, Stenberg J, Fredriksson S, Welch K, Zhang M, et al. 2007. Multigene amplification and massively
- [120] Hornsey M, Loman N, Wareham DW, Ellington MJ, Pallen MJ, et al. 2011. Whole-genome comparison of two *Acinetobacter baumannii* isolates from a single patient, where resistance developed during tigecycline therapy. *J. Antimicrob. Chemother.* 66:1499–503
- [121] Kennemann L, Didelot X, Aebischer T, Kuhn S, Drescher B, et al. 2011. *Helicobacter pylori* genome evolution during human infection. *Proc. Natl. Acad. Sci. USA* 108:5033–38
- [122] Lupski JR, Reid JG, Gonzaga-Jauregui C, Rio Deiros D, Chen DC, et al. 2010. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N. Engl. J. Med.* 362:1181–91
- [123] Tsurusaki Y, Osaka H, Hamanoue H, Shimbo H, Tsuji M, et al. 2011. Rapid detection of a mutation causing X-linked leucoencephalopathy by exome sequencing. *J. Med. Genet.* In press



- [124] Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, et al. 2010. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328:636–39
- [125] Welch JS, Westervelt P, Ding L, Larson DE, Klco JM, et al. 2011. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA* 305:1577–84
- [126] Worthey EA, Mayer AN, Syverson GD, Helbling D, Bonacci BB, et al. 2011. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet. Med.* 13:255–62
- [127] Deloukas P, Bentley D. 2004. The HapMap project and its application to genetic studies of drug response. *Pharmacogenomics J.* 4:88–90
- [128] Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, et al. 2002. The structure of haplotype blocks in the human genome. *Science* 296:2225–29
- [129] Weiss KM, Clark AG. 2002. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* 18:19–24
- [130] Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TO. 2011. A catalog of published genome-wide association studies. <http://www.genome.gov/gwastudies>
- [131] Johnson AD, O'Donnell CJ. 2009. An open access database of genome-wide association results. *BMC Med. Genet.* 10:6

- [132] Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, et al. 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* 40:955–62
- [133] McCarthy MI. 2010. Genomics, type 2 diabetes, and obesity. *N. Engl. J. Med.* 363:2339–50
- [134] Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445:881–85
- [135] Helgadóttir A, Thorleifsson G, Manolescu A, Gretarsdóttir S, Blondal T, et al. 2007. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 316:1491–93
- [136] McCarthy JJ, Parker A, Salem R, Moliterno DJ, Wang Q, et al. 2004. Large scale association analysis for identification of genes underlying premature coronary heart disease: cumulative perspective from analysis of 111 candidate genes. *J. Med. Genet.* 41:334–41
- [137] . McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, et al. 2007. A common allele on chromosome 9 associated with coronary heart disease. *Science* 316:1488–91
- [138] Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, et al. 2007. Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* 357:443–53
- [139] Marshall E. 2009. Lawsuit challenges legal basis for patenting human genes. *Science* 324:1000–1

- [140] Hannon GJ, Beach D. 1994. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature* 371:257–61
- [141] Pasmant E, Laurendeau I, Heron D, Vidaud M, Vidaud D, Bieche I. 2007. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.* 67:3963–69
- [142] . Visel A, Zhu Y, May D, Afzal V, Gong E, et al. 2010. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* 464:409–12
- [143] . Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, et al. 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* 38:662–74
- [144] Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, et al. 2011. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature* 470:264–68
- [145] Huang Y, Qiao F, Atkinson C, Holers VM, Tomlinson S. 2008. A novel targeted inhibitor of the alternative pathway of complement and its therapeutic application in ischemia/reperfusion injury. *J. Immunol.* 181:8068–76
- [146] Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, et al. 2005. Complement factor H polymorphism in age-related macular degeneration. *Science* 308:385–89

- [147] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, et al. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 461:399–401
- [148] Thompson AJ, Muir AJ, Sulkowski MS, Ge D, Fellay J, et al. 2010. Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. *Gastroenterology* 139:120–29
- [149] Link E, Parish S, Armitage J, Bowman L, Heath S, et al. 2008. SLCO1B1 variants and statin-induced myopathy: a genomewide study. *N. Engl. J. Med.* 359:789–99
- [150] Simon R. 2003. Diagnostic and prognostic prediction using gene expression profiles in high-dimensional microarray data. *Br. J. Cancer* 89:1599–604
- [151] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, et al. 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* 27:1160–67
- [152] Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, et al. 2002. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 346:1937–47
- [153] Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, et al. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503–11

- [154] Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, et al. 2001. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl. Acad. Sci. USA* 98:13790–95
- [155] 17. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, et al. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406:536–40 18. 34
- [156] Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–36
- [157] Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531–37
- [158] Bray NJ. 2008. Gene expression in the etiology of schizophrenia. *Schizophr. Bull.* 34:412–18
- [159] Comabella M, Martin R. 2007. Genomics in multiple sclerosis: current state and future directions. *J. Neuroimmunol.* 187:1–8
- [160] Goes FS, Sanders LL, Potash JB. 2008. The genetics of psychotic bipolar disorder. *Curr. Psychiatry Rep.* 10:178–89
- [161] Kittleson MM, Hare JM. 2005. Molecular signature analysis: the potential of gene-expression analysis in cardiomyopathy. *Future Cardiol.* 1:793–808

- [162] Mehta D, Menke A, Binder EB. 2010. Gene expression studies in major depression. *Curr. Psychiatry Rep.* 12:135–44
- [163] Van Baarsen LG, Bos CL, van der Pouw Kraan TC, Verweij CL. 2009. Transcription profiling of rheumatic diseases. *Arthritis Res. Ther.* 11:207
- [164] Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–97
- [165] Lee RC, Feinbaum RL, Ambros V. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75:843–54
- [166] Feng J, Sun G, Yan J, Noltner K, Li W, et al. 2009. Evidence for X-chromosomal schizophrenia associated with microRNA alterations. *PLoS One* 4:e6121
- [167] Iorio MV, Croce CM. 2009. MicroRNAs in cancer: small molecules with a huge impact. *J. Clin. Oncol.* 27:5848–56
- [168] Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. 2005. Modulation of hepatitis C virus RNA abundance by a liver-specific microRNA. *Science* 309:1577–81
- [169] O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. 2005. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 435:839–43
- [170] Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, et al. 2007. Requirement of *bic/microRNA-155* for normal immune function. *Science* 316:608–11

- [171] Tatsuguchi M, Seok HY, Callis TE, Thomson JM, Chen JF, et al. 2007. Expression of microRNAs is dynamically regulated during cardiomyocyte hypertrophy. *J. Mol. Cell. Cardiol.* 42:1137–41
- [172] Budhu A, Ji J, Wang XW. 2010. The clinical potential of microRNAs. *J. Hematol. Oncol.* 3:37
- [173] Calin GA, Croce CM. 2006. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* 6:857–66
- [174] Kartha RV, Subramanian S. 2010. MicroRNAs in cardiovascular diseases: biology and potential clinical applications. *J. Cardiovasc. Transl. Res.* 3:256–70
- [175] Xiao J, Yang B, Lin H, Lu Y, Luo X, Wang Z. 2007. Novel approaches for gene-specific interference via manipulating actions of microRNAs: examination on the pacemaker channel genes HCN2 and HCN4. *J. Cell. Physiol.* 212:285–92
- [176] Davis ME, Zuckerman JE, Choi CH, Seligson D, Tolcher A, et al. 2010. Evidence of RNAi in humans from systemically administered siRNA via targeted nanoparticles. *Nature* 464:1067–70
- [177] Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. 2003. Prediction of mammalian microRNA targets. *Cell* 115:787–98
- [178] Thomas M, Lieberman J, Lal A. 2010. Desperately seeking microRNA targets. *Nat. Struct. Mol. Biol.* 17:1169–74
- [179] Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322:1845–48

- [180] He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. 2008. The antisense transcriptomes of human cells. *Science* 322:1855–57
- [181] Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40:1413–15
- [182] Sugarbaker DJ, Richards WG, Gordon GJ, Dong L, De Rienzo A, et al. 2008. Transcriptome sequencing of malignant pleural mesothelioma tumors. *Proc. Natl. Acad. Sci. USA* 105:3521–26
- [183] Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, et al. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321:956–60
- [184] Guffanti A, Iacono M, Pelucchi P, Kim N, Solda G, et al. 2009. A transcriptional sketch of a primary human breast cancer by 454 deep sequencing. *BMC Genomics* 10:163
- [185] Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, et al. 2009. Transcriptome sequencing to detect gene fusions in cancer. *Nature* 458:97–101
- [186] Palanisamy N, Ateeq B, Kalyana-Sundaram S, Pflueger D, Ramnarayanan K, et al. 2010. Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma. *Nat. Med.* 16:793–98



- [187] Zhao Q, Caballero OL, Levy S, Stevenson BJ, Iseli C, et al. 2009. Transcriptome-guided characterization of genomic rearrangements in a breast cancer cell line. *Proc. Natl. Acad. Sci. USA* 106:1886–91
- [188] Du X, Callister SJ, Manes NP, Adkins JN, Alexandridis RA, et al. 2008. A computational strategy to analyze label-free temporal bottom-up proteomics data. *J. Proteome Res.* 7:2595–604
- [189] Issaq HJ, Veenstra TD. 2008. Would you prefer multiple reaction monitoring or antibodies with your biomarker validation? *Expert Rev. Proteomics* 5:761–63
- [190] Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. 2009. Metabolomics applied to diabetes research: moving from information to knowledge. *Diabetes* 58:2429–43
- [191] Griffin JL, Shockcor JP. 2004. Metabolic profiles of cancer cells. *Nat. Rev. Cancer* 4:551–61
- [192] Kaddurah-Daouk R, McEvoy J, Baillie RA, Lee D, Yao JK, et al. 2007. Metabolomic mapping of atypical antipsychotic effects in schizophrenia. *Mol. Psychiatry* 12:934–45
- [193] Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, et al. 2009. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* 9:311–26
- [194] Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, et al. 2005. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation* 112:3868–75

- [195] Shah SH, Hauser ER, Bain JR, Muehlbauer MJ, Haynes C, et al. 2009. High heritability of metabolomics profiles in families burdened with premature cardiovascular disease. *Mol. Syst. Biol.* 5:258
- [196] Ebbels TM, Keun HC, Beckonert OP, Bollard ME, Lindon JC, et al. 2007. Prediction and classification of drug toxicity using probabilistic modeling of temporal metabolic data: the consortium on metabonomic
- [197] Keun HC. 2006. Metabonomic modeling of drug toxicity. *Pharmacol. Ther.* 109:92–106
- [198] Nicholson JK, Connelly J, Lindon JC, Holmes E. 2002. Metabonomics: a platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* 1:153–61
- [199] . Sharma S, Kelly TK, Jones PA. 2010. Epigenetics in cancer. *Carcinogenesis* 31:27–36
- [200] Bird AP. 1986. CpG-rich islands and the function of DNA methylation. *Nature* 321:209–13
- [201] Prendergast GC, Ziff EB. 1991. Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science* 251:186–89
- [202] Archey WB, McEachern KA, Robson M, Offit K, Vaziri SA, et al. 2002. Increased CpG methylation of the estrogen receptor gene in BRCA1-linked estrogen receptor-negative breast cancers. *Oncogene* 21:7034–41
- [203] Feinberg AP, Vogelstein B. 1983. Hypomethylation of ras oncogenes in primary human cancers. *Biochem. Biophys. Res. Commun.* 111:47–54

- [204] Rosty C, Ueki T, Argani P, Jansen M, Yeo CJ, et al. 2002. Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation. *Am. J. Pathol.* 160:45–50
- [205] Corwin EJ. 2004. The concept of epigenetics and its role in the development of cardiovascular disease: commentary on “New and emerging theories of cardiovascular disease.” *Biol. Res. Nurs.* 6:11–16, 21–23
- [206] Devaskar SU, Thamocharan M. 2007. Metabolic programming in the pathogenesis of insulin resistance. *Rev. Endocr. Metab. Disord.* 8:105–13
- [207] Ji Q, Hao X, Meng Y, Zhang M, Desano J, et al. 2008. Restoration of tumor suppressor miR-34 inhibit human p53-mutant gastric cancer tumorspheres. *BMC Cancer* 8:266
- [208] . Lu Q, Kaplan M, Ray D, Zacharek S, Gutsch D, Richardson B. 2002. Demethylation of ITGAL(CD11a) regulatory sequences in systemic lupus erythematosus. *Arthritis Rheum.* 46:1282–91
- [209] Am. Assoc. Cancer Res. Hum. Epigenome Task Force, Eur. Union Netw. Excell. Sci. Advis. Board. 2008. Moving AHEAD with an international human epigenome project. *Nature* 454:711–15
- [210] Bernstein BE, Meissner A, Lander ES. 2007. The mammalian epigenome. *Cell* 128:669–81

- [211] Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, et al. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* 38:1378–85
- [212] Willard HF, Ginsburg GS. 2009. *Genomic and Personalized Medicine*. Amsterdam: Elsevier. 1478 pp.
- [213] Gonzalez-Angulo AM. 2010. Reply to D.H. Roukos and D.A.X. Schinagl et al. *J. Clin. Oncol.* 28:e282–83
- [214] Gonzalez-Angulo AM, Hennessy BT, Mills GB. 2010. Future of personalized medicine in oncology: a systems biology approach. *J. Clin. Oncol.* 28:2777–83
- [215] Roukos DH. 2010. Beyond HER2 and trastuzumab: heterogeneity, systems biology, and cancer origin research may guide the future for personalized treatment of very early but aggressive breast cancer. *J. Clin. Oncol.* 28:e279–80
- [216] Wang K, Lee I, Carlson G, Hood L, Galas D. 2010. Systems biology and the discovery of diagnostic biomarkers. *Dis. Markers* 28:199–207
- [217] Bell C. J., Dinwiddie D C., Miller N. A. et al. — Carrier testing for severe childhood recessive diseases by next generation sequencing. *Sci. Transl. Med.*, 2011, 3, 65ra4.
- [218] Steffann J., Frydman N., Burlet Ph. et al. Le diagnostic pré-implantatoire couplé au typage HLA: l'expérience française. *Bull. Acad. Natle Med.*, 2011, 195(4-5), 1015-22.
- [219] Fan H. C., Gu W., Wang J. et al. Non-invasive prenatal measurement of the fetal genome. *Nature*. 2012, 487, 320-24.

- [220] Snyder M.W., Simmons L.E., Kitzman J.O. et al.—Noninvasive fetal genome sequencing: a primer. *Prenatal diagnosis*. 2013, 33, 547-54.
- [221] Dugoff L.—Application of genomic technology in prenatal diagnosis. *N. Engl. J. Med.*, 2012, 367, 2249-50.
- [222] Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M. 2010. Genetic susceptibility to breast cancer. *Mol. Oncol.* 4:174–91
- [223] Trainer AH, Lewis CR, Tucker K, Meiser B, Friedlander M, Ward RL. 2010. The role of BRCA mutation testing in determining breast cancer therapy. *Nat. Rev. Clin. Oncol.* 7:708–17
- [224] Vilar E, Gruber SB. 2010. Microsatellite instability in colorectal cancer: the stable evidence. *Nat. Rev. Clin. Oncol.* 7:153–62
- [225] Hedley PL, Jorgensen P, Schlamowitz S, Wangari R, Moolman-Smook J, et al. 2009. The genetic basis of long QT and short QT syndromes: a mutation update. *Hum. Mutat.* 30:1486–511
- [226] Tester DJ, Ackerman MJ. 2011. Genetic testing for potentially lethal, highly treatable inherited cardiomyopathies/ channelopathies in clinical practice. *Circulation* 123:1021–37
- [227] Shimizu W. 2008. Clinical impact of genetic studies in lethal inherited cardiac arrhythmias. *Circ. J.* 72:1926–36
- [228] Lee W, Jiang Z, Liu J, Haverty PM, Guan Y, et al. 2010. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature* 465:473–77

- [229] Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, et al. 2008. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 456:66–72
- [230] Pleasance ED, Stephens PJ, O’Meara S, McBride DJ, Meynert A, et al. 2010. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 463:184–90
- [231] Garber K. 2005. Human Cancer Genome Project moving forward despite some doubts in community. *J. Natl. Cancer Inst.* 97:1322–24
- [232] Cancer Genome Atlas Res. Netw. 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455:1061–68
- [233] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, et al. 2010. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR,
- [234] Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531–37
- [235] Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, et al. 2001. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl. Acad. Sci. USA* 98:13790–95

- [236] Huang JZ, Sanger WG, Greiner TC, Staudt LM, Weisenburger DD, et al. 2002. The t(14;18) defines a unique subset of diffuse large B-cell lymphoma with a germinal center B-cell gene expression profile. *Blood* 99:2285–90
- [237] Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, et al. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* 98:10869–74
- [238] Mateos-Caceres PJ, Garcia-Mendez A, Lopez Farre A, Macaya C, Nunez A, et al. 2004. Proteomic analysis of plasma from patients during an acute coronary syndrome. *J. Am. Coll. Cardiol.* 44:1578–83
- [239] Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, et al. 2005. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation* 112:3868–75
- [240] . Fernandez Parguina A, Grigorian-Shamajian L, Agra RM, Teijeira-Fernandez E, Rosa I, et al. 2010. Proteins involved in platelet signaling are differentially regulated in acute coronary syndrome: a proteomic study. *PLoS One* 5:e13404
- [241] Rosenberg S, Elashoff MR, Beineke P, Daniels SE, Wingrove JA, et al. 2010. Multicenter validation of the diagnostic accuracy of a blood-based gene expression test for assessing obstructive coronary artery disease in nondiabetic patients. *Ann. Intern. Med.* 153:425–34
- [242] Wingrove JA, Daniels SE, Sehnert AJ, Tingley W, Elashoff MR, et al. 2008. Correlation of peripheral blood gene expression with the extent of coronary artery stenosis. *Circ. Cardiovasc. Genet.* 1:31–38

- [243] Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, et al. 2008. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N. Engl. J. Med.* 359:1995–2004
- [244] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, et al. 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* 27:1160–67
- [245] Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, et al. 2002. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 346:1937–47
- [246] Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–36
- [247] Cobleigh MA, Tabesh B, Bitterman P, Baker J, Cronin M, et al. 2005. Tumor gene expression and prognosis in breast cancer patients with 10 or more positive lymph nodes. *Clin. Cancer Res.* 11:8623–31
- [248] Cronin M, Pho M, Dutta D, Stephens JC, Shak S, et al. 2004. Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am. J. Pathol.* 164:35–42
- [249] Sparano JA, Paik S. 2008. Development of the 21-gene assay and its application in clinical practice and clinical trials. *J. Clin. Oncol.* 26:721–28



- [250] Kelly CM, Krishnamurthy S, Bianchini G, Litton JK, Gonzalez-Angulo AM, et al. 2010. Utility of oncotype DX risk estimates in clinically intermediate risk hormone receptor-positive, HER2-normal, grade II, lymph node-negative breast cancers. *Cancer* 116:5161–67
- [251] Paik S, Shak S, Tang G, Kim C, Baker J, et al. 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* 351:2817–26
- [252] Van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, et al. 2002. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* 347:1999–2009
- [253] . Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, et al. 2006. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J. Natl. Cancer Inst.* 98:1183–92
- [254] Glas AM, Floore A, Delahaye LJ, Witteveen AT, Pover RC, et al. 2006. Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics* 7:278
- [255] Hornberger J, Chien R. 2010. Meta-analysis of the decision impact of the 21-gene breast cancer Recurrence Score in clinical practice. Presented at Annu. San Antonio Breast Cancer Symp., 33rd, San Antonio, TX
- [256] Newell KA, Asare A, Kirk AD, Gisler TD, Bourcier K, et al. 2010. Identification of a B cell signature associated with renal transplant tolerance in humans. *J. Clin. Investig.* 120:1836–47

- [257] Holleman A, Cheok MH, den Boer ML, Yang W, Veerman AJ, et al. 2004. Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. *N. Engl. J. Med.* 351:533–42
- [258] Lugthart S, Cheok MH, den Boer ML, Yang W, Holleman A, et al. 2005. Identification of genes associated with chemotherapy crossresistance and treatment response in childhood acute lymphoblastic leukemia. *Cancer Cell* 7:375–86
- [259] King CR, Kraus MH, Aaronson SA. 1985. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. *Science* 229:974–76
- [260] Press MF, Pike MC, Chazin VR, Hung G, Udove JA, et al. 1993. Her-2/neu expression in node-negative breast cancer: direct tissue quantitation by computerized image analysis and association of overexpression with increased risk of recurrent disease. *Cancer Res.* 53:4960–70
- [261] Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, et al. 1999. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J. Clin. Oncol*
- [262] Maughan KL, Lutterbie MA, Ham PS. 2010. Treatment of breast cancer. *Am. Fam. Physician* 81:1339–46

- [263] Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, et al. 2001. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med.* 344:783–92
- [264] Herbst RS, Maddox AM, Rothenberg ML, Small EJ, Rubin EH, et al. 2002. Selective oral epidermal growth factor receptor tyrosine kinase inhibitor ZD1839 is generally well-tolerated and has activity in non-small-cell lung cancer and other solid tumors: results of a phase I trial. *J. Clin. Oncol.* 20:3815–25
- [265] Masui H, Kawamoto T, Sato JD, Wolf B, Sato G, Mendelsohn J. 1984. Growth inhibition of human tumor cells in athymic mice by anti-epidermal growth factor receptor monoclonal antibodies. *Cancer Res.* 44:1002–7
- [266] Ranson M, Hammond LA, Ferry D, Kris M, Tullo A, et al. 2002. ZD1839, a selective oral epidermal growth factor receptor-tyrosine kinase inhibitor, is well tolerated and active in patients with solid, malignant tumors: results of a phase I trial. *J. Clin. Oncol.* 20:2240–50
- [267] Thatcher N, Chang A, Parikh P, Rodrigues Pereira J, Ciuleanu T, et al. 2005. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* 366:1527–37

- [268] Sequist LV, Bell DW, Lynch TJ, Haber DA. 2007. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J. Clin. Oncol.* 25:587–95
- [269] Voora D, McLeod H, Eby C, Gage B. 2005. The pharmacogenetics of coumarin therapy. *Pharmacogenomics* 6:503–13
- [270] Higashi MK, Veenstra DL, Kondo LM, Wittkowsky AK, Srinouanprachanh SL, et al. 2002. Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. *JAMA* 287:1690–98
- [271] Sconce EA, Khan TI, Wynne HA, Avery P, Monkhouse L, et al. 2005. The impact of CYP2C9 and VKORC1 genetic polymorphism and patient characteristics upon warfarin dose requirements: proposal for a new dosing regimen. *Blood* 106:2329–33
- [272] Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, et al. 2009. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N. Engl. J. Med.* 360:753–64
- [273] Driscoll C, Cashion A, Hathaway D, Thompson C, Conley Y, et al. 2006. Blood gene expression profiling in liver transplant recipients with hepatitis C virus and posttransplantation diabetes mellitus. *Transplant. Proc.* 38:3646–48
- [274] Martínez-Llordella M, Puig-Pey I, Orlando G, Ramoni M, Tisone G, et al. 2007. Multiparameter immune profiling of operational tolerance in liver transplantation. *Am. J. Transplant.* 7:309–19

- [275] Starling R, Pham M, Valantine H, Miller L, Eisen H, et al. 2006. Molecular testing in the management of cardiac transplant recipients: initial clinical experience. *J. Heart Lung Transplant.* 25:1389–95
- [276] Deng M, Eisen H, Mehra M, Billingham M, Marboe C, et al. 2006. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. *Am. J. Transplant.* 6:150–60
- [277] Evans R, Williams G, Baron H, Deng M, Eisen H, et al. 2005. The economic implications of noninvasive molecular testing for cardiac allograft rejection. *Am. J. Transplant.* 5:1553–58
- [278] Pham MX, Teuteberg JJ, Kfoury AG, Starling RC, Deng MC, et al. 2010. Gene-expression profiling for rejection surveillance after cardiac transplantation. *N. Engl. J. Med.* 362:1890–900
- [279] Chen B, Chen M, Paisley J, Zaas A, Woods C, et al. 2010. Bayesian inference of the number of factors in gene-expression analysis: application to human virus challenge studies. *BMC Bioinforma.* 11:552
- [280] Zaas AK, Chen M, Varkey J, Veldman T, Hero AO III, et al. 2009. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host Microbe* 6:207–17
- [281] Polonikov AV, Ivanov VP, Solodilova MA. 2009. Genetic variation of genes for xenobiotic-metabolizing enzymes and risk of bronchial asthma: the importance of gene-gene and gene-environment interactions for disease susceptibility. *J. Hum. Genet.* 54:440–49
- [282] Seibold MA, Schwartz DA. 2011. The lung: the natural boundary between nature and nurture. *Annu. Rev. Physiol.* 73:457–78

- [283] Simpson A, John SL, Jury F, Niven R, Woodcock A, et al. 2006. Endotoxin exposure, CD14, and allergic disease: an interaction between genes and the environment. *Am. J. Respir. Crit. CareMed.* 174:386–92
- [284] Fleischmann RD, AdamsMD, White O, Clayton RA, Kirkness EF, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512
- [285] Seib KL, Dougan G, Rappuoli R. 2009. The key role of genomics in modern vaccine and drug design for emerging infectious diseases. *PLoS Genet.* 5:e1000612
- [286] Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, et al. 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 190:6881–93
- [287] Bechah Y, El Karkouri K, Mediannikov O, Leroy Q, Pelletier N, et al. 2010. Genomic, proteomic, and transcriptomic analysis of virulent and avirulent *Rickettsia prowazekii* reveals its adaptive mutation capabilities. *Genome Res.* 20:655–63
- [288] Giuliani MM, Adu-Bobie J, Comanducci M, Arico B, Savino S, et al. 2006. A universal vaccine for serogroup B meningococcus. *Proc. Natl. Acad. Sci. USA* 103:10834–39

- [289] Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, et al. 2007. Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc. Natl. Acad. Sci. USA* 104:9451–56
- [290] Davies J. 2001. In a map for human life, count the microbes, too. *Science* 291:2316
- [291] Savage DC. 1977. Microbial ecology of the gastrointestinal tract. *Annu. Rev. Microbiol.* 31:107–33
- [292] Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, et al. 2006. Metagenomic analysis of the human distal gut microbiome. *Science* 312:1355–59
- [293] Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. 2007. The human microbiome project. *Nature* 449:804–10
- [294] Francois P, Scherl A, Hochstrasser D, Schrenzel J. 2007. Proteomic approach to investigate MRSA. *Methods Mol. Biol.* 391:179–99
- [295] Long WH, Xiao HS, Gu XM, Zhang QH, Yang HJ, et al. 2004. A universal microarray for detection of SARS coronavirus. *J. Virol. Methods* 121:57–63
- [296] Ley RE, Turnbaugh PJ, Klein S, Gordon JI. 2006. Microbial ecology: human gut microbes associated with obesity. *Nature* 444:1022–23
- [297] Peterson DA, Frank DN, Pace NR, Gordon JI. 2008. Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. *Cell Host Microbe* 3:417–27

- [298] Peterson J, Garges S, Giovanni M, McInnes P, Wang L, et al. 2009. The NIH human microbiome project. *Genome Res.* 19:2317–23
- [299] Douglas PS, Ginsburg GS. 2008. Clinical genomic testing: getting it right. *J. Cardiovasc. Transl. Res.* 1:17–20
- [300] Scheuner MT, Sieverding P, Shekelle PG. 2008. Delivery of genomic medicine for common chronic adult diseases: a systematic review. *JAMA* 299:1320–34
- [301] Lapham EV, Kozma C, Weiss JO, Benkendorf JL, Wilson MA. 2000. The gap between practice and genetics education of health professionals: HuGEM survey results. *Genet. Med.* 2:226–31
- [302] Metcalfe S, Hurworth R, Newstead J, Robins R. 2002. Needs assessment study of genetics education for general practitioners in Australia. *Genet. Med.* 4:71–77
- [303] Wilkins-Haug L, Hill LD, Power ML, Holzman GB, Schulkin J. 2000. Gynecologists' training, knowledge, and experiences in genetics: a survey. *Obstet. Gynecol.* 95:421–24
- [304] Goddard KA, Moore C, Ottman D, Szegda KL, Bradley L, Khoury MJ. 2007. Awareness and use of direct-to-consumer nutrigenomic tests, United States, 2006. *Genet. Med.* 9:510–17
- [305] Gramling R, Nash J, Siren K, Eaton C, Culpepper L. 2004. Family physician self-efficacy with screening for inherited cancer risk. *Ann. Fam. Med.* 2:130–32



- [306] FruehFW, Gurwitz D. 2004. From pharmacogenetics to personalized medicine: a vital need for educating health professionals and the community. *Pharmacogenomics* 5:571–79
- [307] Eisenstein EL, Lobach DF, Kawamoto K, Edwards R, Willis JM, et al. 2009. A randomized clinical trial of clinical decision support in a rural community health network serving lower income individuals: study design and baseline characteristics. *Stud. Health Technol. Inform.* 143:220–26
- [308] Kawamoto K, Houlihan CA, Balas EA, Lobach DF. 2005. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 330:765
- [309] Kawamoto K, Lobach DF, Willard HF, Ginsburg GS. 2009. A national clinical decision support infrastructure to enable the widespread and consistent practice of genomic and personalized medicine. *BMC Med. Inform. Decis. Mak.* 9:17
- [310] Apse KA, Biesecker BB, Giardiello FM, Fuller BP, Bernhardt BA. 2004. Perceptions of genetic discrimination among at-risk relatives of colorectal cancer patients. *Genet. Med.* 6:510–166. Archey WB, McEachern KA, Robson M, Offit K, Vaziri SA, et al. 2002. Increased CpG methylation of the estrogen receptor gene in BRCA1-linked estrogen receptor-negative breast cancers. *Oncogene* 21:7034–41
- [311] Hall MA, McEwen JE, Barton JC, Walker AP, Howe EG, et al. 2005. Concerns in a primary care population about genetic discrimination by insurers. *Genet. Med.* 7:311–16

- [312] Klitzman R. 2010. Views of discrimination among individuals confronting genetic disease. *J. Genet. Couns.* 19:68–83
- [313] Haga SB, Willard HF. 2006. Defining the spectrum of genome policy. *Nat. Rev. Genet.* 7:966–72
- [314] Vastag B. 2006. New clinical trials policy at FDA. *Nat. Biotechnol.* 24:1043
- [315] Ginsburg GS, Willard HF. 2009. Genomic and personalized medicine: foundations and applications. *Transl. Res.* 154:277–87
- [316] Fed. Trade Comm. 2006. At-home genetic tests: A healthy dose of skepticism may be the best prescription. <http://www.ftc.gov/bcp/edu/pubs/consumer/health/hea02.shtm>
- [317] Hunter DJ, Khoury MJ, Drazen JM. 2008. Letting the genome out of the bottle: Will we get our wish? *N. Engl. J. Med.* 358:105–7
- [318] McGuire AL, Evans BJ, Caulfield T, Burke W. 2010. Regulating direct-to-consumer personal genome testing. *Science* 330:181–82
- [319] Ginsburg GS, Burke TW, Febbo P. 2008. Centralized biorepositories for genetic and genomic research. *JAMA* 299:1359–61
- [320] Friede A. 2003. National Biospecimen Network Blueprint. Durham, NC: Constella
- [321] Int. Soc. Biol. Environ. Repos. 2005. International Society for Biological and Environmental Repositories: 2004 Annual Meeting. Philadelphia: Taylor & Francis

- [322] HanssonMG, Dillner J, Bartram CR, Carlson JA, Helgesson G. 2006. Should donors be allowed to give broad consent to future biobank research? *Lancet Oncol.* 7:266–69
- [323] Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, et al. 2009. Repeatability of published microarray gene expression analyses. *Nat. Genet.* 41:149–55
- [324] Inst. Med. Comm. Comp. Eff. Res. Prioritization. 2009. Initial National Priorities for Comparative Effectiveness Research. Washington, DC: National Academies. 227 pp
- [325] . Umscheid CA. 2010. Maximizing the clinical utility of comparative effectiveness research. *Clin. Pharmacol. Ther.* 88:876–79
- [326] Garber AM, Tunis SR. 2009. Does comparative-effectiveness research threaten personalized medicine? *N. Engl. J. Med.* 360:1925–27
- [327] Woodward C. 2009. Obama sparks an ideological donnybrook with his push to compare medical treatments. *CMAJ* 180:807
- [328] Marshall E. 2009. Lawsuit challenges legal basis for patenting human genes. *Science* 324:1000–1
- [329] Marshall E. 2010. Cancer gene patents ruled invalid. *Science* 328:153
- [330] Dinan MA, Simmons LA, Snyderman R. 2010. Commentary: Personalized health planning and the Patient Protection and Affordable Care Act: an opportunity for academic medicine to lead health care reform. *Acad. Med.* 85:1665–68

- [331] Whellan DJ, Gaulden L, GattisWA, Granger B, Russell SD, et al. 2001. The benefit of implementing a heart failure disease management program. *Arch. Intern. Med.* 161:2223–28
- [332] Lee SS, Mudaliar A. 2009. Racing forward: the Genomics and PersonalizedMedicine Act. *Science* 323:342
- [333] Goodman J, Walsh V. 2001. The Story of Taxol: Nature and Politics in the Pursuit of an Anti-Cancer Drug. Cambridge: Camb. Univ. Press. 282 pp.
- [334] Burns DK. 2010. Developing pharmacogenetic evidence throughout clinical development. *Clin. Pharmacol. Ther.* 88:867–70
- [335] Westfall JM, Mold J, Fagnan L. 2007. Practice-based research: “Blue Highways” on the NIH roadmap. *JAMA* 297:403–6

## *Serment d'Hippocrate*

*Au moment d'être admis à devenir membre de la profession médicale, je m'engage solennellement à consacrer ma vie au service de l'humanité.*

- *Je traiterai mes maîtres avec le respect et la reconnaissance qui leur sont dus.*
- *Je pratiquerai ma profession avec conscience et dignité. La santé de mes malades sera mon premier but.*
- *Je ne trahirai pas les secrets qui me seront confiés.*
- *Je maintiendrai par tous les moyens en mon pouvoir l'honneur et les nobles traditions de la profession médicale.*
- *Les médecins seront mes frères.*
- *Aucune considération de religion, de nationalité, de race, aucune considération politique et sociale ne s'interposera entre mon devoir et mon patient.*
- *Je maintiendrai le respect de la vie humaine dès la conception.*
- *Même sous la menace, je n'userai pas de mes connaissances médicales d'une façon contraire aux lois de l'humanité.*
- *Je m'y engage librement et sur mon honneur.*

# قسم أبقراط

بسم الله الرحمان الرحيم

أقسم بالله العظيم

في هذه اللحظة التي يتم فيها قبولي عضوا في المهنة الطبية أتعهد علانية:

- بأن أكرس حياتي لخدمة الإنسانية .
- وأن أحترم أساتذتي وأعترف لهم بالجميل الذي يستحقونه .
- وأن أمارس مهنتي بوانزع من ضميري وشر في جاعلا صحة مريض هدي في الأول .
- وأن لا أفشي الأسرار المعهودة إلي .
- وأن أحافظ بكل ما لدي من وسائل على الشرف والتقاليد النبيلة لمهنة الطب .
- وأن أعتبر سائر الأطباء إخوة لي .
- وأن أقوم بواجبي نحو مرضاي بدون أي اعتبار ديني أو وطني أو عرقي أو سياسي أو اجتماعي .
- وأن أحافظ بكل حزم على احترام الحياة الإنسانية منذ نشأتها .
- وأن لا أستعمل معلوماتي الطبية بطريق يضر بحقوق الإنسان مهما لاقيت من تهديد .
- بكل هذا أتعهد عن كامل اختيار ومقسما بالله .

والله على ما أقول شهيد .

جامعة محمد الخامس – الرباط  
كلية الطب والصيدلة بالرباط

أطروحة رقم: 434

سنة : 2017

## تسلسل الحمض النووي والطب الشخصي

### أطروحة

قدمت ونوقشت علانية يوم : .....

من طرف

**السيد: محمد النوالي**

المزدداد في: 12 نونبر 1992 بمراكش

### لنيل شهادة الدكتوراه في الطب

الكلمات الأساسية: علم الجينوم – تسلسل الجيل الجديد – الطب الفردي – الطب الشخصي.

### تحت إشراف اللجنة المكونة من الأساتذة

رئيس

مشرف

أعضاء

السيد: توفيق ضاكة

أستاذ في علم وظائف الأعضاء

السيد: عمر شقيري

أستاذ في علم الأنسجة والأجنة

السيد: فؤاد زويدية

أستاذ في علم التشريح الدقيق

السيد: يونس الرحالي

أستاذ في الصيدلة الغالبية

*UNIVERSITE MOHAMMED V - RABAT*  
*FACULTE DE MEDECINE ET DE PHARMACIE - RABAT-*

*ANNEE: 2017*

*THESE N°: 434*

**LE SEQUENÇAGE DE L'ADN  
ET MEDECINE PERSONNALISEE**

**THÈSE**

*Présentée et soutenue publiquement le :.....*

**PAR**

**Mr. Mohamed EN-NOUALI**  
*Né le 12 Novembre 1992 à Marrakech*

**Pour l'Obtention du Doctorat en Médecine**

**MOTS CLES** : Génomique – Séquençage de nouvelle génération – Médecine individualisée –  
Médecine Personnalisée.

**JURY**

**Mr. T. DAKKA**

Professeur de Physiologie

**PRESIDENT**

**Mr. O. CHOKAIRI**

Professeur d'Histologie Embryologie

**RAPPORTEUR**

**Mr. F. ZOUAIDIA**

Professeur d'Anatomie Pathologie

**Mr. Y. RAHALI**

Professeur de Pharmacie Galénique

**JUGES**



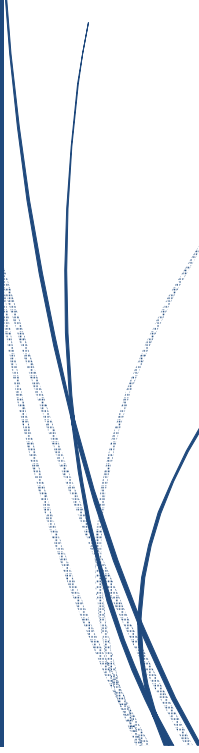


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

سبحانك لا علم لنا إلا ما علمتنا

إنك أنت العليم الحكيم

سورة البقرة: الآية: 31



صَلَّى  
الْعَظِيمِ



**UNIVERSITE MOHAMMED V DE RABAT**  
**FACULTE DE MEDECINE ET DE PHARMACIE - RABAT**

**DOYENS HONORAIRES :**

1962 – 1969 : Professeur Abdelmalek FARAJ  
1969 – 1974 : Professeur Abdellatif BERBICH  
1974 – 1981 : Professeur Bachir LAZRAK  
1981 – 1989 : Professeur Taieb CHKILI  
1989 – 1997 : Professeur Mohamed Tahar ALAOUI  
1997 – 2003 : Professeur Abdelmajid BELMAHI  
2003 – 2013 : Professeur Najia HAJJAJ - HASSOUNI



**ADMINISTRATION :**

**Doyen** : Professeur Mohamed ADNAOUI  
**Vice Doyen chargé des Affaires Académiques et étudiantes**  
Professeur Mohammed AHALLAT  
**Vice Doyen chargé de la Recherche et de la Coopération**  
Professeur Taoufiq DAKKA  
**Vice Doyen chargé des Affaires Spécifiques à la Pharmacie**  
Professeur Jamal TAOUFIK  
**Secrétaire Général** : Mr. Mohamed KARRA

**1- ENSEIGNANTS-CHERCHEURS MEDECINS**

**ET  
PHARMACIENS**

**PROFESSEURS :**

**Décembre 1984**

Pr. MAAOUNI Abdelaziz  
Pr. MAAZOUZI Ahmed Wajdi  
Pr. SETTAF Abdellatif

Médecine Interne – **Clinique Royale**  
Anesthésie -Réanimation  
pathologie Chirurgicale

**Novembre et Décembre 1985**

Pr. BENSAID Younes

Pathologie Chirurgicale

**Janvier, Février et Décembre 1987**

Pr. CHAHED OUZZANI Houria  
Pr. LACHKAR Hassan  
Pr. YAHYAOUY Mohamed

Gastro-Entérologie  
Médecine Interne  
Neurologie

**Décembre 1988**

Pr. BENHAMAMOUCH Mohamed Najib  
Pr. DAFIRI Rachida

Chirurgie Pédiatrique  
Radiologie

### **Décembre 1989**

Pr. ADNABOU Mohamed  
Pr. CHAD Bouziane  
Pr. OUZZANI Taïbi Mohamed Réda

### **Janvier et Novembre 1990**

Pr. CHKOFF Rachid  
Pr. HACHIM Mohammed\*  
Pr. KHARBACH Aïcha  
Pr. MANSOURI Fatima  
Pr. TAZI Saoud Anas

### **Février Avril Juillet et Décembre 1991**

Pr. AL HAMANY Zaïtounia  
Pr. AZZOUZI Abderrahim  
Pr. BAYAHIA Rabéa  
Pr. BELKOUCHI Abdelkader  
Pr. BENCHEKROUN Belabbes Abdellatif  
Pr. BENSOUDA Yahia  
Pr. BERRAHO Amina  
Pr. BEZZAD Rachid  
Pr. CHABRAOUI Layachi  
Pr. CHERRAH Yahia  
Pr. CHOKAIRI Omar  
Pr. KHATTAB Mohamed  
Pr. SOULAYMANI Rachida  
Pr. TAOUFIK Jamal

### **Décembre 1992**

Pr. AHALLAT Mohamed  
Pr. BENSOUDA Adil  
Pr. BOUJIDA Mohamed Najib  
Pr. CHAHED OUZZANI Laaziza  
Pr. CHRAIBI Chafiq  
Pr. DEHAYNI Mohamed\*  
Pr. EL OUAHABI Abdessamad  
Pr. FELLAT Rokaya  
Pr. GHAFIR Driss\*  
Pr. JIDDANE Mohamed  
Pr. TAGHY Ahmed  
Pr. ZOUHDI Mimoun

### **Mars 1994**

Pr. BENJAAFAR Noureddine  
Pr. BEN RAIS Nozha  
Pr. CAOUI Malika  
Pr. CHRAIBI Abdelmjid  
  
Pr. EL AMRANI Sabah

Médecine Interne – **Doyen de la FMPR**  
Pathologie Chirurgicale  
Neurologie

Pathologie Chirurgicale  
Médecine-Interne  
Gynécologie -Obstétrique  
Anatomie-Pathologique  
Anesthésie Réanimation

Anatomie-Pathologique  
Anesthésie Réanimation – **Doyen de la FMPO**  
Néphrologie  
Chirurgie Générale  
Chirurgie Générale  
Pharmacie galénique  
Ophtalmologie  
Gynécologie Obstétrique  
Biochimie et Chimie  
Pharmacologie  
Histologie Embryologie  
Pédiatrie  
Pharmacologie – **Dir. du Centre National PV**  
Chimie thérapeutique **V.D à la pharmacie+Dir du CEDOC**

Chirurgie Générale V.D Aff. Acad. et Estud  
Anesthésie Réanimation  
Radiologie  
Gastro-Entérologie  
Gynécologie Obstétrique  
Gynécologie Obstétrique  
Neurochirurgie  
Cardiologie  
Médecine Interne  
Anatomie  
Chirurgie Générale  
Microbiologie



Radiothérapie  
Biophysique  
Biophysique  
Endocrinologie et Maladies Métaboliques **Doyen de la FMPA**  
Gynécologie Obstétrique

Pr. EL BARDOUNI Ahmed  
 Pr. EL HASSANI My Rachid  
 Pr. ERROUGANI Abdelkader  
 Pr. ESSAKALI Malika  
 Pr. ETTAYEBI Fouad  
 Pr. HADRI Larbi\*  
 Pr. HASSAM Badredine  
 Pr. IFRINE Lahssan  
 Pr. JELTHI Ahmed  
 Pr. MAHFOUD Mustapha  
 Pr. RHRAB Brahim  
 Pr. SENOUCI Karima

### **Mars 1994**

Pr. ABBAR Mohamed\*  
 Pr. ABDELHAK M'barek  
 Pr. BELAIDI Halima  
 Pr. BENTAHILA Abdelali  
 Pr. BENYAHIA Mohammed Ali  
 Pr. BERRADA Mohamed Saleh  
 Pr. CHAMI Ilham  
 Pr. CHERKAOUI Lalla Ouafae  
 Pr. JALIL Abdelouahed  
 Pr. LAKHDAR Amina  
 Pr. MOUANE Nezha

### **Mars 1995**

Pr. ABOUQUAL Redouane  
 Pr. AMRAOUI Mohamed  
 Pr. BAIDADA Abdelaziz  
 Pr. BARGACH Samir  
 Pr. CHAARI Jilali\*  
 Pr. DIMOU M'barek\*  
 Pr. DRISSI KAMILI Med Nordine\*  
 Pr. EL MESNAOUI Abbes  
 Pr. ESSAKALI HOUSSYNI Leila  
 Pr. HDA Abdelhamid\*  
 Pr. IBEN ATTYA ANDALOUSSI Ahmed  
 Pr. OUAZZANI CHAHDI Bahia  
 Pr. SEFIANI Abdelaziz  
 Pr. ZEGGWAGH Amine Ali

### **Décembre 1996**

Pr. AMIL Touriya\*  
 Pr. BELKACEM Rachid  
 Pr. BOULANOUAR Abdelkrim  
 Pr. EL ALAMI EL FARICHA EL Hassan  
 Pr. GAOUZI Ahmed  
 Pr. MAHFOUDI M'barek\*  
 Pr. OUADGHIRI Mohamed  
 Pr. OUZEDDOUN Naima  
 Pr. ZBIR EL Mehdi\*

Traumato-Orthopédie  
 Radiologie  
 Chirurgie Générale- **Directeur CHIS**  
 Immunologie  
 Chirurgie Pédiatrique  
 Médecine Interne  
 Dermatologie  
 Chirurgie Générale  
 Anatomie Pathologique  
 Traumatologie – Orthopédie  
 Gynécologie –Obstétrique  
 Dermatologie

Urologie  
 Chirurgie – Pédiatrique  
 Neurologie  
 Pédiatrie  
 Gynécologie – Obstétrique  
 Traumatologie – Orthopédie  
 Radiologie  
 Ophtalmologie  
 Chirurgie Générale  
 Gynécologie Obstétrique  
 Pédiatrie

Réanimation Médicale  
 Chirurgie Générale  
 Gynécologie Obstétrique  
 Gynécologie Obstétrique  
 Médecine Interne  
 Anesthésie Réanimation  
 Anesthésie Réanimation  
 Chirurgie Générale  
 Oto-Rhino-Laryngologie  
 Cardiologie - **Directeur HMI Med V**  
 Urologie  
 Ophtalmologie  
 Génétique  
 Réanimation Médicale

Radiologie  
 Chirurgie Pédiatrie  
 Ophtalmologie  
 Chirurgie Générale  
 Pédiatrie  
 Radiologie  
 Traumatologie-Orthopédie  
 Néphrologie  
 Cardiologie



### **Novembre 1997**

Pr. ALAMI Mohamed Hassan  
Pr. BEN SLIMANE Lounis  
Pr. BIROUK Nazha  
Pr. ERREIMI Naima  
Pr. FELLAT Nadia  
Pr. HAIMEUR Charki\*  
Pr. KADDOURI Nouredine  
Pr. KOUTANI Abdellatif  
Pr. LAHLOU Mohamed Khalid  
Pr. MAHRAOUI CHAFIQ  
Pr. TAOUFIQ Jallal  
Pr. YOUSFI MALKI Mounia

Gynécologie-Obstétrique  
Urologie  
Neurologie  
Pédiatrie  
Cardiologie  
Anesthésie Réanimation  
Chirurgie Pédiatrique  
Urologie  
Chirurgie Générale  
Pédiatrie  
Psychiatrie  
Gynécologie Obstétrique

### **Novembre 1998**

Pr. AFIFI RAJAA  
Pr. BENOMAR ALI  
Pr. BOUGTAB Abdesslam  
Pr. ER RIHANI Hassan  
Pr. BENKIRANE Majid\*  
Pr. KHATOURI ALI\*

Gastro-Entérologie  
Neurologie – *Doyen de la FMP Abulcassis*  
Chirurgie Générale  
Oncologie Médicale  
Hématologie  
Cardiologie

### **Janvier 2000**

Pr. ABID Ahmed\*  
Pr. AIT OUMAR Hassan  
Pr. BENJELLOUN Dakhama Badr.Sououd  
Pr. BOURKADI Jamal-Eddine  
Pr. CHARIF CHEFCHAOUNI Al Montacer  
Pr. ECHARRAB El Mahjoub  
Pr. EL FTOUH Mustapha  
Pr. EL MOSTARCHID Brahim\*  
Pr. ISMAILI Hassane\*  
Pr. MAHMOUDI Abdelkrim\*  
Pr. TACHINANTE Rajae  
Pr. TAZI MEZALEK Zoubida

Pneumophtisiologie  
Pédiatrie  
Pédiatrie  
Pneumo-phtisiologie  
Chirurgie Générale  
Chirurgie Générale  
Pneumo-phtisiologie  
Neurochirurgie  
Traumatologie Orthopédie- *Dir. Hop. Av. Marr.*  
Anesthésie-Réanimation *Inspecteur du SSM*  
Anesthésie-Réanimation  
Médecine Interne



### **Novembre 2000**

Pr. AIDI Saadia  
Pr. AJANA Fatima Zohra  
Pr. BENAMR Said  
Pr. CHERTI Mohammed  
Pr. ECH-CHERIF EL KETTANI Selma  
Pr. EL HASSANI Amine  
Pr. EL KHADER Khalid  
Pr. EL MAGHRAOUI Abdellah\*  
Pr. GHARBI Mohamed El Hassan  
Pr. MAHASSINI Najat  
Pr. MDAGHRI ALAOUI Asmae  
Pr. ROUIMI Abdelhadi\*

Neurologie  
Gastro-Entérologie  
Chirurgie Générale  
Cardiologie  
Anesthésie-Réanimation  
Pédiatrie *Directeur Hop. Chekikh Zaied*  
Urologie  
Rhumatologie  
Endocrinologie et Maladies Métaboliques  
Anatomie Pathologique  
Pédiatrie  
Neurologie

## Décembre 2000

Pr. ZOHAIR ABDELAH\*

ORL

## Décembre 2001

Pr. BALKHI Hicham\*  
Pr. BENABDELJLIL Maria  
Pr. BENAMAR Loubna  
Pr. BENAMOR Jouda  
Pr. BENELBARHDADI Imane  
Pr. BENNANI Rajae  
Pr. BENOUECHANE Thami  
Pr. BEZZA Ahmed\*  
Pr. BOUCHIKHI IDRISSE Med Larbi  
Pr. BOUMDIN El Hassane\*  
Pr. CHAT Latifa  
Pr. DAALI Mustapha\*  
Pr. DRISSI Sidi Mourad\*  
Pr. EL HIJRI Ahmed  
Pr. EL MAAQILI Moulay Rachid  
Pr. EL MADHI Tarik  
Pr. EL OUNANI Mohamed  
Pr. ETTAIR Said  
Pr. GAZZAZ Miloudi\*  
Pr. HRORA Abdelmalek  
Pr. KABBAJ Saad  
Pr. KABIRI EL Hassane\*  
Pr. LAMRANI Moulay Omar  
Pr. LEKEHAL Brahim  
Pr. MAHASSIN Fattouma\*  
Pr. MEDARHRI Jalil  
Pr. MIKDAME Mohammed\*  
Pr. MOHSINE Raouf  
Pr. NOUINI Yassine  
Pr. SABBABH Farid  
Pr. SEFIANI Yasser  
Pr. TAOUFIQ BENCHEKROUN Soumia

Anesthésie-Réanimation  
Neurologie  
Néphrologie  
Pneumo-phtisiologie  
Gastro-Entérologie  
Cardiologie  
Pédiatrie  
Rhumatologie  
Anatomie  
Radiologie  
Radiologie  
Chirurgie Générale  
Radiologie  
Anesthésie-Réanimation  
Neuro-Chirurgie  
Chirurgie-Pédiatrique  
Chirurgie Générale  
Pédiatrie **Directeur. Hop.d'Enfants**  
Neuro-Chirurgie  
Chirurgie Générale  
Anesthésie-Réanimation  
Chirurgie Thoracique  
Traumatologie Orthopédie  
Chirurgie Vasculaire Périphérique  
Médecine Interne  
Chirurgie Générale  
Hématologie Clinique  
Chirurgie Générale  
Urologie **Directeur Hôpital Ibn Sina**  
Chirurgie Générale  
Chirurgie Vasculaire Périphérique  
Pédiatrie



## Décembre 2002

Pr. AL BOUZIDI Abderrahmane\*  
Pr. AMEUR Ahmed \*  
Pr. AMRI Rachida  
Pr. AOURARH Aziz\*  
Pr. BAMOU Youssef \*  
Pr. BELMEJDOUB Ghizlene\*  
Pr. BENZEKRI Laila  
Pr. BENZZOUBEIR Nadia  
Pr. BERNOUSSI Zakiya  
Pr. BICHRA Mohamed Zakariya\*  
Pr. CHOHO Abdelkrim \*

Anatomie Pathologique  
Urologie  
Cardiologie  
Gastro-Entérologie  
Biochimie-Chimie  
Endocrinologie et Maladies Métaboliques  
Dermatologie  
Gastro-Entérologie  
Anatomie Pathologique  
Psychiatrie  
Chirurgie Générale

Pr. CHKIRATE Bouchra  
 Pr. EL ALAMI EL FELLOUS Sidi Zouhair  
 Pr. EL HAOURI Mohamed \*  
 Pr. FILALI ADIB Abdelhai  
 Pr. HAJJI Zakia  
 Pr. IKEN Ali  
 Pr. JAAFAR Abdeloihab\*  
 Pr. KRIOUILE Yamina  
 Pr. LAGHMARI Mina  
 Pr. MABROUK Hfid\*  
 Pr. MOUSSAOUI RAHALI Driss\*  
 Pr. OUJILAL Abdelilah  
 Pr. RACHID Khalid \*  
 Pr. RAISS Mohamed  
 Pr. RGUIBI IDRISSE Sidi Mustapha\*  
 Pr. RHOU Hakima  
 Pr. SIAH Samir \*  
 Pr. THIMOU Amal  
 Pr. ZENTAR Aziz\*

#### **Janvier 2004**

Pr. ABDELLAH El Hassan  
 Pr. AMRANI Mariam  
 Pr. BENBOUZID Mohammed Anas  
 Pr. BENKIRANE Ahmed\*  
 Pr. BOUGHALEM Mohamed\*  
 Pr. BOULAADAS Malik  
 Pr. BOURAZZA Ahmed\*  
 Pr. CHAGAR Belkacem\*  
 Pr. CHERRADI Nadia  
 Pr. EL FENNI Jamal\*  
 Pr. EL HANCHI ZAKI  
 Pr. EL KHORASSANI Mohamed  
 Pr. EL YOUNASSI Badreddine\*  
 Pr. HACHI Hafid  
 Pr. JABOUIRIK Fatima  
 Pr. KHARMAZ Mohamed  
 Pr. MOUGHIL Said  
 Pr. OUBAAZ Abdelbarre\*  
 Pr. TARIB Abdelilah\*  
 Pr. TIJAMI Fouad  
 Pr. ZARZUR Jamila

#### **Janvier 2005**

Pr. ABBASSI Abdellah  
 Pr. AL KANDRY Sif Eddine\*  
 Pr. ALLALI Fadoua  
 Pr. AMAZOUZI Abdellah  
 Pr. AZIZ Nouredine\*  
 Pr. BAHIRI Rachid

Pédiatrie  
 Chirurgie Pédiatrique  
 Dermatologie  
 Gynécologie Obstétrique  
 Ophtalmologie  
 Urologie  
 Traumatologie Orthopédie  
 Pédiatrie  
 Ophtalmologie  
 Traumatologie Orthopédie  
 Gynécologie Obstétrique  
 Oto-Rhino-Laryngologie  
 Traumatologie Orthopédie  
 Chirurgie Générale  
 Pneumophtisiologie  
 Néphrologie  
 Anesthésie Réanimation  
 Pédiatrie  
 Chirurgie Générale

Ophtalmologie  
 Anatomie Pathologique  
 Oto-Rhino-Laryngologie  
 Gastro-Entérologie  
 Anesthésie Réanimation  
 Stomatologie et Chirurgie Maxillo-faciale  
 Neurologie  
 Traumatologie Orthopédie  
 Anatomie Pathologique  
 Radiologie  
 Gynécologie Obstétrique  
 Pédiatrie  
 Cardiologie  
 Chirurgie Générale  
 Pédiatrie  
 Traumatologie Orthopédie  
 Chirurgie Cardio-Vasculaire  
 Ophtalmologie  
 Pharmacie Clinique  
 Chirurgie Générale  
 Cardiologie

Chirurgie Réparatrice et Plastique  
 Chirurgie Générale  
 Rhumatologie  
 Ophtalmologie  
 Radiologie  
 Rhumatologie



Pr. BARKAT Amina  
 Pr. BENYASS Aatif  
 Pr. BERNOUSSI Abdelghani  
 Pr. DOUDOUH Abderrahim\*  
 Pr. EL HAMZAOUI Sakina\*  
 Pr. HAJJI Leila  
 Pr. HESSISSEN Leila  
 Pr. JIDAL Mohamed\*  
 Pr. LAAROUSSI Mohamed  
 Pr. LYAGOUBI Mohammed  
 Pr. NIAMANE Radouane\*  
 Pr. RAGALA Abdelhak  
 Pr. SBIHI Souad  
 Pr. ZERAIDI Najia

### **Décembre 2005**

Pr. CHANI Mohamed

### **Avril 2006**

Pr. ACHEMLAL Lahsen\*  
 Pr. AKJOUJ Said\*  
 Pr. BELMEKKI Abdelkader\*  
 Pr. BENCHEIKH Razika  
 Pr. BIYI Abdelhamid\*  
 Pr. BOUHAFS Mohamed El Amine  
 Pr. BOULAHYA Abdellatif\*  
 Pr. CHENGUETI ANSARI Anas  
 Pr. DOGHMI Nawal  
 Pr. FELLAT Ibtissam  
 Pr. FAROUDY Mamoun  
 Pr. HARMOUCHE Hicham  
 Pr. HANAFI Sidi Mohamed\*  
 Pr. IDRIS LAHLOU Amine\*  
 Pr. JROUNDI Laila  
 Pr. KARMOUNI Tariq  
 Pr. KILI Amina  
 Pr. KISRA Hassan  
 Pr. KISRA Mounir  
 Pr. LAATIRIS Abdelkader\*  
 Pr. LMIMOUNI Badreddine\*  
 Pr. MANSOURI Hamid\*  
 Pr. OUANASS Abderrazzak  
 Pr. SAFI Soumaya\*  
 Pr. SEKKAT Fatima Zahra  
 Pr. SOUALHI Mouna  
 Pr. TELLAL Saida\*  
 Pr. ZAHRAOUI Rachida

### **Octobre 2007**

Pr. ABIDI Khalid  
 Pr. ACHACHI Leila

Pédiatrie  
 Cardiologie  
 Ophtalmologie  
 Biophysique  
 Microbiologie  
 Cardiologie (mise en disponibilité)  
 Pédiatrie  
 Radiologie  
 Chirurgie Cardio-vasculaire  
 Parasitologie  
 Rhumatologie  
 Gynécologie Obstétrique  
 Histo-Embryologie Cytogénétique  
 Gynécologie Obstétrique

Anesthésie Réanimation

Rhumatologie  
 Radiologie  
 Hématologie  
 O.R.L  
 Biophysique  
 Chirurgie - Pédiatrique  
 Chirurgie Cardio – Vasculaire  
 Gynécologie Obstétrique  
 Cardiologie  
 Cardiologie  
 Anesthésie Réanimation  
 Médecine Interne  
 Anesthésie Réanimation  
 Microbiologie  
 Radiologie  
 Urologie  
 Pédiatrie  
 Psychiatrie  
 Chirurgie – Pédiatrique  
 Pharmacie Galénique  
 Parasitologie  
 Radiothérapie  
 Psychiatrie  
 Endocrinologie  
 Psychiatrie  
 Pneumo – Phtisiologie  
 Biochimie  
 Pneumo – Phtisiologie





Pr. ACHOUR Abdessamad\*  
 Pr. AIT HOUSSA Mahdi\*  
 Pr. AMHAJJI Larbi\*  
 Pr. AOUI Sarra  
 Pr. BAITE Abdelouahed\*  
 Pr. BALOUCH Lhoussaine\*  
 Pr. BENZIANE Hamid\*  
 Pr. BOUTIMZINE Nourdine  
 Pr. CHARKAOUI Naoual\*  
 Pr. EHIRCHIOU Abdelkader\*  
 Pr. ELABSI Mohamed  
 Pr. EL MOUSSAOUI Rachid  
 Pr. EL OMARI Fatima  
 Pr. GHARIB Noureddine  
 Pr. HADADI Khalid\*  
 Pr. ICHOU Mohamed\*  
 Pr. ISMAILI Nadia  
 Pr. KEBDANI Tayeb  
 Pr. LALAOU SALIM Jaafar\*  
 Pr. LOUZI Lhoussain\*  
 Pr. MADANI Naoufel  
 Pr. MAHI Mohamed\*  
 Pr. MARC Karima  
 Pr. MASRAR Azlarab  
 Pr. MRABET Mustapha\*  
 Pr. MRANI Saad\*  
 Pr. OUZZIF Ez zohra\*  
 Pr. RABHI Monsef\*  
 Pr. RADOUANE Bouchaib\*  
 Pr. SEFFAR Myriame  
 Pr. SEKHSOKH Yessine\*  
 Pr. SIFAT Hassan\*  
 Pr. TABERKANET Mustafa\*  
 Pr. TACHFOUTI Samira  
 Pr. TAJDINE Mohammed Tariq\*  
 Pr. TANANE Mansour\*  
 Pr. TLIGUI Houssain  
 Pr. TOUATI Zakia

### **Décembre 2007**

Pr. DOUHAL ABDERRAHMAN

### **Décembre 2008**

Pr ZOUBIR Mohamed\*  
 Pr TAHIRI My El Hassan\*

Chirurgie générale  
 Chirurgie cardio vasculaire  
 Traumatologie orthopédie  
 Parasitologie  
 Anesthésie réanimation **Directeur ERSM**  
 Biochimie-chimie  
 Pharmacie clinique  
 Ophtalmologie  
 Pharmacie galénique  
 Chirurgie générale  
 Chirurgie générale  
 Anesthésie réanimation  
 Psychiatrie  
 Chirurgie plastique et réparatrice  
 Radiothérapie  
 Oncologie médicale  
 Dermatologie  
 Radiothérapie  
 Anesthésie réanimation  
 Microbiologie  
 Réanimation médicale  
 Radiologie  
 Pneumo phtisiologie  
 Hématologique  
 Médecine préventive santé publique et hygiène  
 Virologie  
 Biochimie-chimie  
 Médecine interne  
 Radiologie  
 Microbiologie  
 Microbiologie  
 Radiothérapie  
 Chirurgie vasculaire périphérique  
 Ophtalmologie  
 Chirurgie générale  
 Traumatologie orthopédie  
 Parasitologie  
 Cardiologie



Ophtalmologie

Anesthésie Réanimation  
 Chirurgie Générale

## **Mars 2009**

Pr. ABOUZAHIR Ali\*  
Pr. AGDR Aomar\*  
Pr. AIT ALI Abdelmounaim\*  
Pr. AIT BENHADDOU El hachmia  
Pr. AKHADDAR Ali\*  
Pr. ALLALI Nazik  
Pr. AMINE Bouchra  
Pr. ARKHA Yassir  
Pr. BELYAMANI Lahcen\*  
Pr. BJIJOU Younes  
Pr. BOUHSAIN Sanae\*  
Pr. BOUI Mohammed\*  
Pr. BOUNAIM Ahmed\*  
Pr. BOUSSOUGA Mostapha\*  
Pr. CHAKOUR Mohammed \*  
Pr. CHTATA Hassan Toufik\*  
Pr. DOGHMI Kamal\*  
Pr. EL MALKI Hadj Omar  
Pr. EL OUENNASS Mostapha\*  
Pr. ENNIBI Khalid\*  
Pr. FATHI Khalid  
Pr. HASSIKOU Hasna \*  
Pr. KABBAJ Nawal  
Pr. KABIRI Meryem  
Pr. KARBOUBI Lamya  
Pr. L'KASSIMI Hachemi\*  
Pr. LAMSAOURI Jamal\*  
Pr. MARMADE Lahcen  
Pr. MESKINI Toufik  
Pr. MESSAOUDI Nezha \*  
Pr. MSSROURI Rahal  
Pr. NASSAR Ittimade  
Pr. OUKERRAJ Latifa  
Pr. RHORFI Ismail Abderrahmani \*

## **PROFESSEURS AGREGES :**

### **Octobre 2010**

Pr. ALILOU Mustapha  
Pr. AMEZIANE Taoufiq\*  
Pr. BELAGUID Abdelaziz  
Pr. BOUAITY Brahim\*  
Pr. CHADLI Mariama\*  
Pr. CHEMSI Mohamed\*  
Pr. DAMI Abdellah\*  
Pr. DARBI Abdellatif\*  
Pr. DENDANE Mohammed Anouar  
Pr. EL HAFIDI Naima  
Pr. EL KHARRAS Abdennasser\*

Médecine interne  
Pédiatre  
Chirurgie Générale  
Neurologie  
Neuro-chirurgie  
Radiologie  
Rhumatologie  
Neuro-chirurgie  
Anesthésie Réanimation  
Anatomie  
Biochimie-chimie  
Dermatologie  
Chirurgie Générale  
Traumatologie orthopédique  
Hématologie biologique  
Chirurgie vasculaire périphérique  
Hématologie clinique  
Chirurgie Générale  
Microbiologie  
Médecine interne  
Gynécologie obstétrique  
Rhumatologie  
Gastro-entérologie  
Pédiatrie  
Pédiatrie  
Microbiologie ***Directeur Hôpital My Ismail***  
Chimie Thérapeutique  
Chirurgie Cardio-vasculaire  
Pédiatrie  
Hématologie biologique  
Chirurgie Générale  
Radiologie  
Cardiologie  
Pneumo-phtisiologie



Anesthésie réanimation  
Médecine interne  
Physiologie  
ORL  
Microbiologie  
Médecine aéronautique  
Biochimie chimie  
Radiologie  
Chirurgie pédiatrique  
Pédiatrie  
Radiologie

Pr. EL MAZOUZ Samir  
 Pr. EL SAYEGH Hachem  
 Pr. ERRABIH Ikram  
 Pr. LAMALMI Najat  
 Pr. MOSADIK Ahlam  
 Pr. MOUJAHID Mountassir\*  
 Pr. NAZIH Mouna\*  
 Pr. ZOUAIDIA Fouad

Chirurgie plastique et réparatrice  
 Urologie  
 Gastro entérologie  
 Anatomie pathologique  
 Anesthésie Réanimation  
 Chirurgie générale  
 Hématologie  
 Anatomie pathologique

### **Mai 2012**

Pr. AMRANI Abdelouahed  
 Pr. ABOUELALAA Khalil\*  
 Pr. BELAIZI Mohamed\*  
 Pr. BENCHEBBA Driss\*  
 Pr. DRISSI Mohamed\*  
 Pr. EL ALAOUI MHAMDI Mouna  
 Pr. EL KHATTABI Abdessadek\*  
 Pr. EL OUAZZANI Hanane\*  
 Pr. ER-RAJI Mounir  
 Pr. JAHID Ahmed  
 Pr. MEHSSANI Jamal\*  
 Pr. RAISSOUNI Maha\*

Chirurgie Pédiatrique  
 Anesthésie Réanimation  
 Psychiatrie  
 Traumatologie Orthopédique  
 Anesthésie Réanimation  
 Chirurgie Générale  
 Médecine Interne  
 Pneumophtisiologie  
 Chirurgie Pédiatrique  
 Anatomie pathologique  
 Psychiatrie  
 Cardiologie

### **Février 2013**

Pr. AHID Samir  
 Pr. AIT EL CADI Mina  
 Pr. AMRANI HANCHI Laila  
 Pr. AMOUR Mourad  
 Pr. AWAB Almahdi  
 Pr. BELAYACHI Jihane  
 Pr. BELKHADIR Zakaria Houssain  
 Pr. BENCHEKROUN Laila  
 Pr. BENKIRANE Souad  
 Pr. BENNANA Ahmed\*  
 0.  
 Pr. BENSGHIR Mustapha\*  
 Pr. BENYAHIA Mohammed\*  
 Pr. BOUATIA Mustapha  
 Pr. BOUABID Ahmed Salim\*  
 Pr. BOUTARBOUCH Mahjouba  
 Pr. CHAIB Ali\*  
 Pr. DENDANE Tarek  
 Pr. DINI Nouzha\*  
 Pr. ECH-CHERIF EL KETTANI Mohamed Ali  
 Pr. ECH-CHERIF EL KETTANI Najwa  
 Pr. ELFATEMI Nizare  
 Pr. EL GUERROUJ Hasnae  
 Pr. EL HARTI Jaouad

Pharmacologie – Chimie  
 Toxicologie  
 Gastro-Entérologie  
 Anesthésie Réanimation  
 Anesthésie Réanimation  
 Réanimation Médicale  
 Anesthésie Réanimation  
 Biochimie-Chimie  
 Hématologie  
 Informatique Pharmaceutique

Anesthésie Réanimation  
 Néphrologie  
 Chimie Analytique  
 Traumatologie Orthopédie  
 Anatomie  
 Cardiologie  
 Réanimation Médicale  
 Pédiatrie  
 Anesthésie Réanimation  
 Radiologie  
 Neuro-Chirurgie  
 Médecine Nucléaire  
 Chimie Thérapeutique



Pr. EL JOUDI Rachid\*  
 Pr. EL KABABRI Maria  
 Pr. EL KHANNOUSSI Basma  
 Pr. EL KHLOUFI Samir  
 Pr. EL KORAICHI Alae  
 Pr. EN-NOUALI Hassane\*  
 Pr. ERRGUIG Laila  
 Pr. FIKRI Meryim  
 Pr. GHFIR Imade  
 Pr. IMANE Zineb  
 Pr. IRAQI Hind  
 Pr. KABBAJ Hakima  
 Pr. KADIRI Mohamed\*  
 Pr. LATIB Rachida  
 Pr. MAAMAR Mouna Fatima Zahra  
 Pr. MEDDAH Bouchra  
 Pr. MELHAOUI Adyl  
 Pr. MRABTI Hind  
 Pr. NEJJARI Rachid  
 Pr. OUBEJJA Houda  
 Pr. OUKABLI Mohamed\*  
 Pr. RAHALI Younes  
 Pr. RATBI Ilham  
 Pr. RAHMANI Mounia  
 Pr. REDA Karim\*  
 Pr. REGRAGUI Wafa  
 Pr. RKAIN Hanan  
 Pr. ROSTOM Samira  
 Pr. ROUAS Lamiaa  
 Pr. ROUIBAA Fedoua\*  
 Pr. SALIHOUN Mouna  
 Pr. SAYAH Rochde  
 Pr. SEDDIK Hassan\*  
 Pr. ZERHOUNI Hicham  
 Pr. ZINE Ali\*

Toxicologie  
 Pédiatrie  
 Anatomie Pathologie  
 Anatomie  
 Anesthésie Réanimation  
 Radiologie  
 Physiologie  
 Radiologie  
 Médecine Nucléaire  
 Pédiatrie  
 Endocrinologie et maladies métaboliques  
 Microbiologie  
 Psychiatrie  
 Radiologie  
 Médecine Interne  
 Pharmacologie  
 Neuro-chirurgie  
 Oncologie Médicale  
 Pharmacognosie  
 Chirurgie Pédiatrique  
 Anatomie Pathologique  
 Pharmacie Galénique  
 Génétique  
 Neurologie  
 Ophtalmologie  
 Neurologie  
 Physiologie  
 Rhumatologie  
 Anatomie Pathologique  
 Gastro-Entérologie  
 Gastro-Entérologie  
 Chirurgie Cardio-Vasculaire  
 Gastro-Entérologie  
 Chirurgie Pédiatrique  
 Traumatologie Orthopédie

### **Avril 2013**

Pr. EL KHATIB Mohamed Karim\*  
 Pr. GHOUNDALE Omar\*  
 Pr. ZYANI Mohammad\*

Stomatologie et Chirurgie Maxillo-faciale  
 Urologie  
 Médecine Interne

**\*Enseignants Militaires**



## **MARS 2014**

ACHIR ABDELLAH  
BENCHAKROUN MOHAMMED  
BOUCHIKH MOHAMMED  
EL KABBAJ DRISS  
EL MACHTANI IDRISSE SAMIRA  
HARDIZI HOUYAM  
HASSANI AMALE  
HERRAK LAILA  
JANANE ABDELLA TIF  
JEAIDI ANASS  
KOUACH JAOUAD  
LEMNOUER ABDELHAY  
MAKRAM SANAA  
OULAHYANE RACHID  
RHISSASSI MOHAMED JMFAR  
SABRY MOHAMED  
SEKKACH YOUSSEF  
TAZL MOUKBA. :LA.KLA.

**\*Enseignants Militaires**

Chirurgie Thoracique  
Traumatologie- Orthopédie  
Chirurgie Thoracique  
Néphrologie  
Biochimie-Chimie  
Histologie- Embryologie-Cytogénétique  
Pédiatrie  
Pneumologie  
Urologie  
Hématologie Biologique  
Génécologie-Obstétrique  
Microbiologie  
Pharmacologie  
Chirurgie Pédiatrique  
CCV  
Cardiologie  
Médecine Interne  
Génécologie-Obstétrique

## **DECEMBRE 2014**

ABILKACEM RACHID'  
AIT BOUGHIMA FADILA  
BEKKALI HICHAM  
BENAZZOU SALMA  
BOUABDELLAH MOUNYA  
BOUCHRIK MOURAD  
DERRAJI SOUFIANE  
DOBLALI TAOUFIK  
EL AYOUBI EL IDRISSE ALI  
EL GHADBANE ABDEDAIM HATIM  
EL MARJANY MOHAMMED  
FEJJAL NAWFAL  
JAHIDI MOHAMED  
LAKHAL ZOUHAIR  
OUDGHIRI NEZHA  
Rami Mohamed  
SABIR MARIA  
SBAI IDRISSE KARIM

**\*Enseignants Militaires**

Pédiatrie  
Médecine Légale  
Anesthésie-Réanimation  
Chirurgie Maxillo-Faciale  
Biochimie-Chimie  
Parasitologie  
Pharmacie Clinique  
Microbiologie  
Anatomie  
Anesthésie-Réanimation  
Radiothérapie  
Chirurgie Réparatrice et Plastique  
O.R.L  
Cardiologie  
Anesthésie-Réanimation  
Chirurgie Pédiatrique  
Psychiatrie  
Médecine préventive, santé publique et Hyg.



## **AOUT 2015**

Meziane meryem  
Tahri latifa

Dermatologie  
Rhumatologie

## **JANVIER 2016**

BENKABBOU AMINE  
EL ASRI FOUAD  
ERRAMI NOUREDDINE  
NITASSI SOPHIA

Chirurgie Générale  
Ophtalmologie  
O.R.L  
O.R.L

## **2- ENSEIGNANTS – CHERCHEURS SCIENTIFIQUES**

### **PROFESSEURS / PRs. HABILITES**

Pr. ABOUDRAR Saadia  
Pr. ALAMI OUHABI Naima  
Pr. ALAOUI KATIM  
Pr. ALAOUI SLIMANI Lalla Naïma  
Pr. ANSAR M'hammed  
Pr. BOUHOUCHE Ahmed  
Pr. BOUKLOUZE Abdelaziz  
Pr. BOURJOUANE Mohamed  
Pr. CHAHED OUAZZANI Lalla Chadia  
Pr. DAKKA Taoufiq  
Pr. DRAOUI Mustapha  
Pr. EL GUESSABI Lahcen  
Pr. ETTAIB Abdelkader  
Pr. FAOUZI Moulay El Abbes  
Pr. HAMZAOUI Laila  
Pr. HMAMOUCHE Mohamed  
Pr. IBRAHIMI Azeddine  
Pr. KHANFRI Jamal Eddine  
Pr. OULAD BOUYAHYA IDRISSE Med  
Pr. REDHA Ahlam  
Pr. TOUATI Driss  
Pr. ZAHIDI Ahmed  
Pr. ZELLOU Amina

Physiologie  
Biochimie – chimie  
Pharmacologie  
Histologie-Embryologie  
Chimie Organique et Pharmacie Chimique  
Génétique Humaine  
Applications Pharmaceutiques  
Microbiologie  
Biochimie – chimie  
Physiologie  
Chimie Analytique  
Pharmacognosie  
Zootechnie  
Pharmacologie  
Biophysique  
Chimie Organique  
Biologie moléculaire  
Biologie  
Chimie Organique  
Chimie  
Pharmacognosie  
Pharmacologie  
Chimie Organique

*Mise à jour le 14/12/2016 par le  
Service des Ressources Humaines*





# *Dédicaces*

*JE DEDIE CETTE THESE ...*



*A mon père et à ma mère*

*Aucune dédicace ne saurait exprimer tout le respect, la reconnaissance et  
l'amour que je vous porte.*

*A mon grand frère Amine*

*Je te souhaite une vie pleine de bonheur, de santé et de prospérité.*

*A tous les membres de ma famille petits et grands*

*Veillez trouver dans ce modeste travail l'expression de mon affection la  
plus sincère.*

*A tous mes amis et mes collègues*

*En guise de reconnaissance et d'amitié pure, je vous dédie ce modeste  
travail.*

*A tous ceux qui me sont chers*



# *Remerciements*

*Notre maître et président de thèse*

*Mr le professeur DAKKA T.*

*Professeur de physiologie*

*Nous sommes très touchés par l'honneur que vous nous faites en acceptant  
de présider le jury de cette thèse.*

*C'est avec grande joie que nous avons accueilli votre accord.*

*Que ce travail soit pour nous une occasion de vous exprimer notre  
admiration et notre profond respect.*

*A notre maître et rapporteur de thèse*

*Mr le professeur CHOKAIRI O.*

*Professeur d'Histologie et Embryologie.*

*Nous vous remercions de nous faire avoir fait l'honneur de nous confier ce travail.*

*Acceptez, cher maître, l'hommage de notre gratitude qui, si grande qu'elle puisse être, ne sera jamais à la hauteur de votre dévouement.*

*A notre maître et juge de thèse*

*Mr. le professeur ZOUIDIA F.*

*Professeur d'Anatomie et Pathologie*

*Vous avez accepté de siéger parmi le jury de notre thèse. Ce geste dénote non seulement de votre gentillesse mais surtout de votre souci du devoir envers vos étudiants.*

*Veillez accepter Monsieur le Professeur, ma profonde reconnaissance et mes remerciements les plus sincères.*

*Soyez assuré que c'est une fierté pour nous de vous compter parmi les membres de notre jury.*

*A notre maître et juge de thèse*

*Mr. le professeur RAHALI Y.*

*Professeur de Pharmacie Galénique*

*C'est pour nous un immense privilège de vous voir accepter de juger ce  
travail.*

*Veillez croire cher maître à notre très haute considération et notre profond  
respect.*

*Enfin qu'il nous soit permis d'exprimer globalement nos remerciements à tous ceux et à toutes celles qui ont manifesté leur appui et ont facilité notre travail.*

*Merci*

*Mohamed En-nouali*



## Liste des abréviations

|              |                                                                                                                                          |
|--------------|------------------------------------------------------------------------------------------------------------------------------------------|
| ADN          | : Acide désoxyribonucléique                                                                                                              |
| a-hémolysine | : Alpha-Hémolysine                                                                                                                       |
| ARN          | : Acide ribonucléique                                                                                                                    |
| CAD          | : Coronary artery disease                                                                                                                |
| CARGO        | : The Cardiac Allograft Rejection Gene Expression Observation (L'observation de l'expression du gène de rejet de l'allogreffe cardiaque) |
| CD           | : Cyclodextrin                                                                                                                           |
| ERS          | : Evaluation standard des risques pour la santé                                                                                          |
| FDA          | : The US Food and Drug Administration                                                                                                    |
| GWAS         | : Genome-wide association studies (études d'association à l'échelle du génome)                                                           |
| HGP          | : Human Genome Project (projet du génome humain)                                                                                         |
| HSF          | : Historique de santé familiale                                                                                                          |
| IMAGE        | : Invasive Monitoring Attenuation Through Gene Expression                                                                                |
| miRNAs       | : micrornas                                                                                                                              |
| MsPA         | : Mycobacterium smegmatis porin A                                                                                                        |
| NGS          | : Next generation sequencing (séquenceur de nouvelle génération)                                                                         |
| ONT          | : Oxford Nanopore Technologies                                                                                                           |

|        |                                                                                  |
|--------|----------------------------------------------------------------------------------|
| PBMC   | : Peripheral blood mononuclear cell (Cellule mononucléaire du sang périphérique) |
| PheWAS | : Phenome-wide association studies (Etudes d'association à l'échelle du phenome) |
| phi29  | : Bacillus phage phi29                                                           |
| READNA | : The REvolutionary Approaches and Devices for Nucleic Acid analysis             |
| REC    | : Recherche sur l'efficacité comparative                                         |
| SCA    | : Syndromes coronariens aigus                                                    |
| siRNAs | : Small interfering RNAs                                                         |
| SNP    | : Single-nucleotide polymorphisms                                                |
| SSD    | : Systèmes de soutien à la décision clinique                                     |
| UE     | : Union européenne                                                               |
| WGS    | : Whole genome sequencing (le séquençage complet du génome)                      |



# *Sommaire*

|                                                                                    |           |
|------------------------------------------------------------------------------------|-----------|
| <b>I. Introduction .....</b>                                                       | <b>2</b>  |
| <b>II. Rappel.....</b>                                                             | <b>6</b>  |
| 1. Définition.....                                                                 | 7         |
| 2. Propriétés .....                                                                | 10        |
| 3. Fonctions biologiques .....                                                     | 13        |
| 4. Gènes et génomes.....                                                           | 14        |
| 5. Transcription et traduction .....                                               | 15        |
| 6. Réplication .....                                                               | 17        |
| 7. Histoire de la recherche sur l'ADN .....                                        | 17        |
| <b>III. Le séquençage de l'ADN.....</b>                                            | <b>21</b> |
| 1. Caractéristiques des générations de séquençage de l'ADN .....                   | 24        |
| 1.1. Séquençage de 1ère génération.....                                            | 24        |
| 1.2. Séquençage de 2ème génération.....                                            | 25        |
| 1.3. Séquençage de 2.5 génération.....                                             | 26        |
| 1.4. Séquençage de 3ème génération.....                                            | 27        |
| 1.5. Séquençage de 4ème génération.....                                            | 27        |
| 2. Dispositifs de séquençage d'ADN de 2ème et 2.5 génération .....                 | 28        |
| 2.1. Dispositifs de séquençage d'ADN de 2ème génération qui ont réussi ..          | 28        |
| 2.2. Dispositifs de séquençage d'ADN de 2e génération qui n'ont jamais réussi..... | 29        |

|                                                                                                                                                             |           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 2.3. Méthodes de détection innovantes .....                                                                                                                 | 29        |
| 2.4. L'analyse des données .....                                                                                                                            | 31        |
| 3. Séquençage de 3ème génération – nano pores .....                                                                                                         | 32        |
| 3.1. Introduction aux technologies nano pores :.....                                                                                                        | 33        |
| 3.2. Acteurs commerciaux clés du séquençage nano poreux.....                                                                                                | 36        |
| 4. READNA et séquençage d'ADN .....                                                                                                                         | 37        |
| 4.1. Aperçu de READNA.....                                                                                                                                  | 37        |
| 4.2. Séquençage de nano pores d'Oxford (séquençage d'exonucléases nano poreuses et séquençage de brins de nano pores) - Séquençage de 3ème génération ..... | 38        |
| 4.2.1. Séquençage d'exonucléase de nanopores .....                                                                                                          | 38        |
| 4.2.2. Séquençage de brin de nano pores .....                                                                                                               | 40        |
| 4.3. Séquençage in situ (séquençage des cellules individuelles) - Séquençage de 4ème génération .....                                                       | 41        |
| 4.4. Manipulation de chromosomes complets, techniques de cartographie optique .....                                                                         | 41        |
| <b>IV. La médecine personnalisée .....</b>                                                                                                                  | <b>44</b> |
| 1. Qu'est-ce que la médecine personnalisée ?.....                                                                                                           | 45        |
| 1.1. La Médecine Personnalisée et l'Histoire de santé familiale .....                                                                                       | 45        |
| 1.2. Évaluation du risque de maladie chronique .....                                                                                                        | 47        |
| 1.3. Système de soutien à la décision clinique .....                                                                                                        | 48        |

|                                                                   |    |
|-------------------------------------------------------------------|----|
| 1.4. La médecine personnalisée et le génome humain .....          | 48 |
| 2. La boîte à outils du génome humain .....                       | 50 |
| 2.1. Variation à l'échelle du génome .....                        | 51 |
| 2.2. Transcriptomique .....                                       | 54 |
| 2.3. Protéomique .....                                            | 56 |
| 2.4. Métabolomique .....                                          | 57 |
| 2.5. Épigénomique .....                                           | 57 |
| 2.6. Biologie des systèmes et médecine des systèmes.....          | 58 |
| 3. De l'information de génome à la médecine personnalisée .....   | 59 |
| 3.1. Tests préconceptionnels et dépistage des hétérozygotes ..... | 59 |
| 3.2. Diagnostic prénatal et préimplantatoire .....                | 61 |
| 3.3. Prédiction du risque.....                                    | 63 |
| 3.4. Diagnostic de la maladie et caractérisation moléculaire..... | 64 |
| 3.5. Pronostic de la maladie.....                                 | 66 |
| 3.6. Pharmacogénomique .....                                      | 68 |
| 3.7. Surveillance de la réponse des maladies à la thérapie .....  | 70 |
| 4. Génomes interactives et microbiomes .....                      | 71 |
| 4.1. Génomes pathogènes.....                                      | 72 |
| 4.2. Le microbiome humain .....                                   | 73 |
| 5. Intégration de la génomique dans la pratique clinique .....    | 76 |
| 5.1. Education .....                                              | 77 |

|                                                            |           |
|------------------------------------------------------------|-----------|
| 5.2. Système de soutien à la décision clinique .....       | 77        |
| 5.3. Vie privée.....                                       | 78        |
| 5.4. Politique réglementaire.....                          | 79        |
| 5.5. Normes.....                                           | 79        |
| 5.6. Recherche sur l'efficacité comparative .....          | 81        |
| 5.7. Propriété intellectuelle.....                         | 81        |
| 5.8. Remboursement .....                                   | 82        |
| 6. Quand la "médecine personnalisée" est "médecine "?..... | 83        |
| <b>V. Conclusion.....</b>                                  | <b>86</b> |
| <b>VI. Résumé .....</b>                                    | <b>90</b> |
| <b>VII. bibliographie .....</b>                            | <b>94</b> |

" Le fait que j'ai un risque génétique pour la maladie d'Alzheimer et la cécité n'est pas une bonne nouvelle. Mais la réalité est que chacun d'entre nous aura des dizaines de ces risques, et ce que nous devons apprendre est de savoir comment les traiter. "

**Craig Venter**

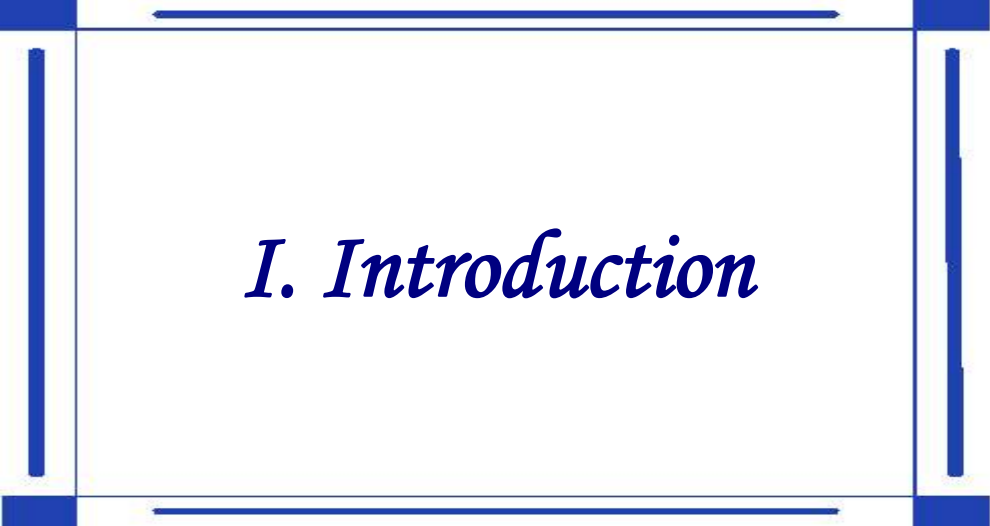
"Les individus ne sont pas des choses stables, ils sont éphémères. Les chromosomes aussi sont mélangés dans l'oubli, comme des mains de cartes peu après qu'ils soient distribués. Mais les cartes elles-mêmes survivent au brassage. Les cartes sont les gènes. Les gènes ne sont pas détruits par la traversée, ils changent simplement de partenaires et continuent leur chemin. Bien sûr, ils continuent. C'est leur affaire. Ils sont les répliqueurs et nous sommes leurs machines de survie. Quand nous avons servi notre but, nous sommes mis de côté. Mais les gènes sont des habitants du temps géologique: les gènes sont éternels. "

**Richard Dawkins, Le gène égoïste**

" Au cours de la dernière décennie, nous avons débloqué de nombreux mystères concernant l'ADN et l'ARN. Cette connaissance ne se limite pas aux livres sur les étagères et ne se limite pas aux établis des laboratoires. Nous avons utilisé ces résultats de recherche pour identifier les causes de nombreuses maladies. De plus, les scientifiques ont traduit ces connaissances génétiques en plusieurs traitements et thérapies, ce qui a permis de créer un pont entre le laboratoire et le patient. ”

**Le président Barack Obama** Sur la Loi sur la génomique et la médecine personnalisée (S.976), le 23 mars 2007





# *I. Introduction*

La génomique est l'étude du matériel génétique complet, y compris les gènes et leurs fonctions, de l'organisme. Il se concentre sur l'obtention et l'analyse des données du génome, c'est-à-dire, les données au niveau de l'ADN en essayant de découvrir les modèles spécifiques d'expressions génétiques ou de molécules biologiques les dits biomarqueurs (molécule de signature et marqueur moléculaire) qui différencient l'état maladie de l'état normal et également déterminer la variabilité dans la réponse des patients aux médicaments .La génomique transforme la pratique médicale par l'individualisation des traitements médicaux en disposant des données génétiques des patients. La médecine est individuelle ; Nous ne sommes pas tous les mêmes et nos variations génétique détermine comment une personne réagira à la thérapie. Ces variations dans la séquence d'un gène peuvent entraîner des différences dans les enzymes qui métabolisent les médicaments. C'est ce qui explique pourquoi les enzymes apparaissent sous différentes formes chez les individus. C'est aussi pourquoi des personnes différentes métabolisent un ou le même médicament différemment. Pourquoi une drogue fonctionne efficacement pour certaines personnes, mais pas pour d'autres ? Pourquoi quelqu'un requiert-il deux fois la quantité régulièrement prescrite des mêmes médicaments pour fonctionner efficacement ? Pourquoi certaines personnes ressentent des effets secondaires des médicaments, mais d'autres ne le font pas ? Pourquoi certaines personnes vont être affectés par des cancers /n'importe quelle maladie, mais d'autres pas. Voici les raisons pour lesquelles la médecine personnalisée et individualisée est importante pour tous.

Le séquençage des acides nucléiques est le pilier de la recherche biologique. Il existe plusieurs générations de technologies de séquençage de l'ADN qui peuvent être bien caractérisées par leur nature et le type de production qu'elles fournissent. Le séquençage des terminateurs didésoxy développé par Sanger a dominé pendant 30 ans et a été le cheval de bataille du projet sur le génome humain. En 2005, le premier séquenceur de 2e génération a été présenté avec des ordres de sortie beaucoup plus élevés que le séquençage de Sanger et une réduction spectaculaire des coûts. Nous sommes maintenant à l'aube de la 3ème génération avec des systèmes nano pores qui sont en cours de développement pour le séquençage de l'ADN.

Au cours de la dernière décennie, de hauts fonctionnaires, des dirigeants de l'industrie, des fournisseurs de soins de santé et le public ont adopté la médecine génomique et personnalisée. La médecine génomique, qui est l'utilisation de l'information provenant des génomes et de leurs dérivés (ARN, protéines et métabolites) pour guider la prise de décision médicale, est un élément clé de la médecine personnalisée, qui est un domaine de soins de santé en évolution rapide, qui s'appuie sur les informations cliniques, génétiques, génomiques et environnementales uniques à chaque personne. Comme la médecine commence à utiliser des outils génomiques qui permettent une prédiction et un traitement plus précis, incluant l'interrogation du génome entier sur la variation de séquence, la transcription, les protéines et les métabolites, les principes de la médecine génomique et personnalisée nécessiteront le développement, la standardisation, et l'intégration de plusieurs outils importants dans les systèmes de santé et les flux de travail cliniques. Ces outils comprennent l'évaluation des risques pour la santé, les antécédents médicaux de la famille et l'aide à la

décision clinique pour les risques complexes et les informations prédictives. Conjugués à l'information génomique, ces outils permettront un changement de paradigme vers une approche globale qui identifiera les risques individuels et guidera la prise en charge clinique et la prise de décision, qui formeront la base d'une approche plus efficace et éclairée des soins aux patients. L'évaluation des risques liés à l'ADN pour la maladie complexe commune, les signatures moléculaires pour le diagnostic et le pronostic du cancer, la thérapie guidée par le génome et la sélection des doses sont parmi les exemples les plus importants. En outre, les informations provenant de génomes individuels, qui constituent un domaine en évolution rapide dans le domaine du développement technologique, engendrent une révolution sociale et de l'information parmi les consommateurs qui affectera sans aucun doute la prise de décision en matière de soins de santé. Bien que ces découvertes scientifiques et d'autres fassent leur chemin du génome à la clinique, L'application intégrale de la médecine génomique et personnalisée dans les soins de santé exigera des changements radicaux dans les politiques de réglementation et de remboursement, ainsi que des protections législatives pour la protection de la vie privée à l'échelle du système. Ainsi, il existe des défis à la fois scientifiques et politiques pour les soins de santé personnalisés; Cependant, ils seront confrontés et résolus avec la certitude que la science derrière la médecine génomique est solide et la pratique de la médecine qu'elle informe est basée sur des preuves.

L'objectif principal de cette étude est d'examiner les applications et le rôle de la génomique dans la médecine personnalisée.



## *II. Rappel*

## 1. Définition

L'acide désoxyribonucléique (l'ADN) (1) est une molécule qui porte les instructions génétiques utilisées dans la croissance, le développement, le fonctionnement et la reproduction de tous les organismes vivants connus et de nombreux virus. L'ADN et l'acide ribonucléique (ARN) sont des acides nucléiques ; aux côtés des protéines, des lipides et des glucides complexes (polysaccharides), ils constituent l'un des quatre principaux types de macromolécules essentiels à toutes les formes de vie connues. La plupart des molécules d'ADN sont constituées de deux brins de bio polymère enroulés l'un autour de l'autre pour former une double hélice.

Les deux brins d'ADN sont appelés poly nucléotides car ils sont composés d'unités monomères plus simples appelées nucléotides (2,3). Chaque nucléotide est composé de l'un des quatre nucléobases contenant de l'azote - cytosine (C), guanine (G), adénine (A) ou thymine (T) - un sucre appelé désoxyribose et un groupe phosphate. Les nucléotides sont joints les uns aux autres dans une chaîne par des liaisons covalentes entre le sucre d'un nucléotide et le phosphate du suivant, résultant en une alternance épine dorsale de sucre-phosphate. Les bases azotées des deux brins poly nucléotidiques séparés sont liées ensemble, selon les règles d'appariement des bases (A avec T et C avec G), avec des liaisons hydrogène pour former de l'ADN double brin. La quantité totale de paires de bases d'ADN apparentées sur Terre est estimée à  $5,0 \times 10^{37}$  et pèse 50 milliards de tonnes (4). En comparaison, la masse totale de la biosphère a été estimée à 4 billions de tonnes de carbone (TtC) (5).

L'ADN stocke des informations biologiques. Le squelette de l'ADN est résistant au clivage, et les deux brins de la structure à double brin stockent la même information biologique. Cette information est répliquée au fur et à mesure que les deux brins se séparent. Une grande partie de l'ADN (plus de 98% pour les humains) est non codante, ce qui signifie que ces sections ne servent pas de modèles pour les séquences protéiques.

Les deux brins d'ADN vont dans des directions opposées l'un à l'autre et sont donc antiparallèles. Attaché à chaque sucre est l'un des quatre types de nucléobases (informellement, *bases*). C'est la séquence de ces quatre nucléobases le long du squelette qui code l'information biologique. Les brins d'ARN sont créés en utilisant des brins d'ADN comme matrice dans un processus appelé transcription. Sous le code génétique, ces brins d'ARN sont traduits pour spécifier la séquence des acides aminés dans les protéines dans un processus appelé traduction.

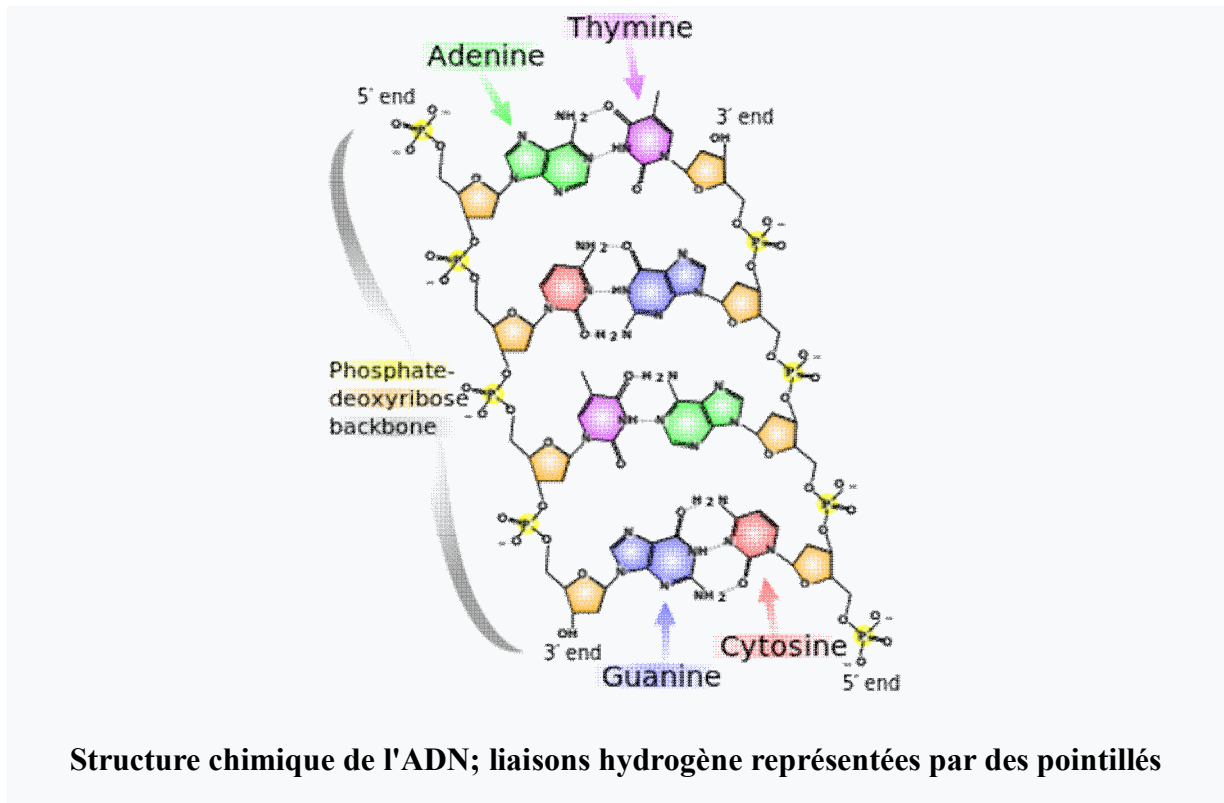
Au sein des cellules eucaryotes, l'ADN est organisé en longues structures appelées chromosomes. Au cours de la division cellulaire, ces chromosomes sont dupliqués dans le processus de réplication de l'ADN, fournissant à chaque cellule son propre ensemble complet de chromosomes. Les organismes eucaryotes (animaux, plantes, champignons et protistes) stockent la plus grande partie de leur ADN à l'intérieur du noyau cellulaire et une partie de leur ADN dans des organites, tels que les mitochondries ou les chloroplastes (6). En revanche, les procaryotes (bactéries et archaea) ne stockent leur ADN que dans le cytoplasme. Au sein des chromosomes eucaryotes, les protéines de la chromatine telles que les histones se concentrent et organisent l'ADN. Ces

structures compactes guident les interactions entre l'ADN et les autres protéines, aidant à contrôler quelles parties de l'ADN sont transcrites.

L'ADN a été isolé par Friedrich Miescher en 1869. Sa structure moléculaire a été identifiée par James Watson et Francis Crick au Laboratoire Cavendish de l' Université de Cambridge en 1953, dont les efforts de modélisation ont été guidés par les données de diffraction des rayons X acquises par Raymond Gosling , qui était un étudiant de troisième cycle de Rosalind Franklin . L'ADN est utilisé par les chercheurs comme un outil moléculaire pour explorer les lois physiques et les théories, telles que le théorème ergodique et la théorie de l'élasticité. Les propriétés matérielles uniques de l'ADN en ont fait une molécule attrayante pour les scientifiques et les ingénieurs intéressés par la micro et la nano fabrication. Parmi les avancées notables dans ce domaine, citons l'ADN origami et les matériaux hybrides à base d'ADN (7).



## 2. Propriétés



L'ADN est un long polymère fabriqué à partir d'unités répétées appelées nucléotides (8,9). La structure de l'ADN est dynamique le long de sa longueur, étant capable de s'enrouler dans des boucles serrées, et d'autres formes (10). Dans toutes les espèces il est composé de deux chaînes hélicoïdales, liées les unes aux autres par des liaisons d'hydrogène. Les deux chaînes sont enroulées autour du même axe et ont la même hauteur de 34 angströms (3,4 nanomètres). La paire de chaînes a un rayon de 10 angströms (1,0 nanomètre) (11). Selon une autre étude, lorsqu'elle est mesurée dans une solution différente, la chaîne d'ADN mesure de 22 à 26 nm de large (2,2 à 2,6 nanomètres), et une unité nucléotidique mesure 3,3 Å (0,33 nm) de longueur (12). Bien que chaque unité

de répétition de nucléotide individuelle soit très petite, les polymères d'ADN peuvent être de très grandes molécules contenant des millions à des centaines de millions de nucléotides. Par exemple, l'ADN du plus grand chromosome humain, le chromosome numéro 1, comprend environ 220 millions de paires de bases et aurait une longueur de 85 mm s'il était redressé (13).

Dans les organismes vivants, l'ADN n'existe généralement pas comme une seule molécule, mais plutôt comme une paire de molécules qui sont maintenues étroitement ensemble (14,15). Ces deux longs brins s'entrelacent comme des vignes, dans la forme d'une double hélice. Le nucléotide contient à la fois un segment du squelette de la molécule (qui maintient la chaîne ensemble) et une nucléobase (qui interagit avec l'autre brin d'ADN dans l'hélice). Une nucléobase liée à un sucre est appelée un nucléoside et une base liée à un sucre et un ou plusieurs groupes de phosphate est appelé un nucléotide. Un polymère comprenant plusieurs nucléotides liés (comme dans l'ADN) est appelé polynucléotide (16).

Le squelette du brin d'ADN est fabriqué à partir de résidus de phosphate et de sucre alternés (17). Le sucre dans l'ADN est le 2- désoxyribose, qui est un pentose (cinq- carbone). Les sucres sont réunis par des groupes phosphate qui forment des liaisons phosphodiester entre les troisième et cinquième atomes de carbone des anneaux de sucre adjacents. Ces liaisons asymétriques signifient qu'un brin d'ADN a une direction. Dans une double hélice, la direction des nucléotides dans un brin est opposée à leur direction dans l'autre brin: les brins sont *antiparallèles*. Les extrémités asymétriques des brins d'ADN sont dites avoir une directionnalité de amorce *cinq* (5') et amorce *trois* (3'), avec l'extrémité 5' ayant un groupe phosphate terminal et l'extrémité 3' un groupe hydroxyle

terminal. Une différence majeure entre l'ADN et l'ARN est le sucre, avec le 2-désoxyribose dans l'ADN étant remplacé par l'alternative pentose sucre ribose dans l'ARN (15).

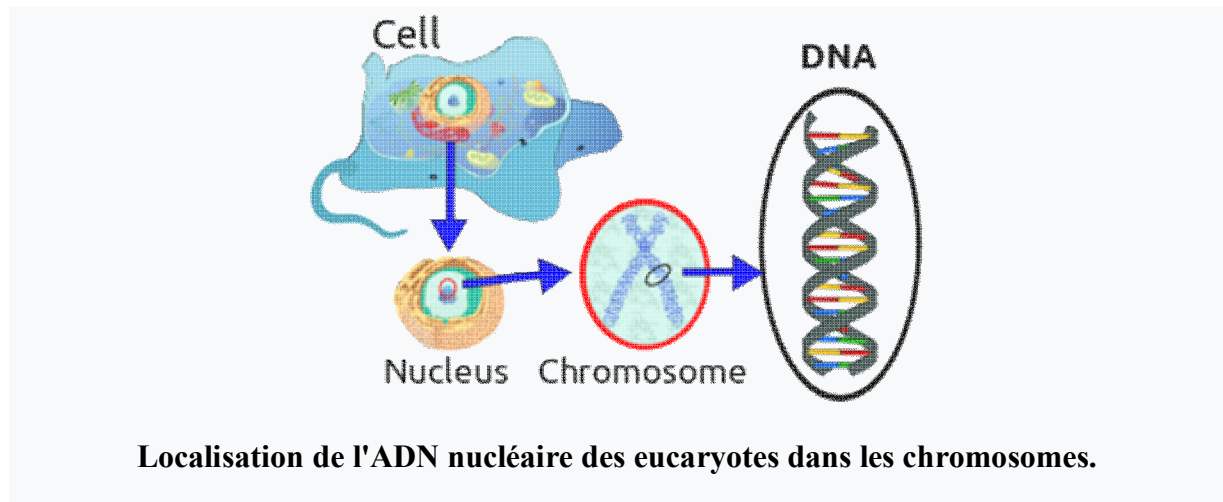


**Une section d'ADN. Les bases se trouvent horizontalement entre les deux brins en spirale (18).**

La double hélice d'ADN est stabilisée principalement par deux forces: les liaisons hydrogène entre les nucléotides et les interactions d'empilement de bases entre les nucléobases aromatiques (19). Dans l'environnement aqueux de la cellule, les liaisons  $\pi$  conjuguées des bases nucléotidiques s'alignent perpendiculairement à l'axe de la molécule d'ADN, minimisant leur interaction avec l'enveloppe de solvation. Les quatre bases trouvées dans l'ADN sont l'adénine (A), la cytosine (C), la guanine (G) et la thymine (T). Ces quatre bases sont attachées au sucre-phosphate pour former le nucléotide complet, comme indiqué pour l'adénosine mono phosphate. Des paires adénine avec de la

thymine et des paires guanine avec de la cytosine . Il était représenté par des paires de bases A-T et des paires de bases G-C (20,21).

### 3. Fonctions biologiques



L'ADN se présente généralement sous forme de chromosomes linéaires chez les eucaryotes et de chromosomes circulaires chez les procaryotes . L'ensemble des chromosomes dans une cellule constitue son génome ; le génome humain a environ 3 milliards de paires de bases d'ADN disposées en 46 chromosomes (22). L'information portée par l'ADN est contenue dans la séquence de morceaux d'ADN appelés gènes .La transmission de l'information génétique dans les gènes est réalisée via l'appariement de bases complémentaires. Par exemple, en transcription, lorsqu'une cellule utilise l'information dans un gène, la séquence d'ADN est copiée dans une séquence d'ARN complémentaire par l'intermédiaire de l'attraction entre l'ADN et les nucléotides d'ARN corrects. Habituellement, cette copie d'ARN est ensuite utilisée pour construire une séquence de protéine correspondante dans un processus appelé traduction , qui dépend de la même interaction entre les

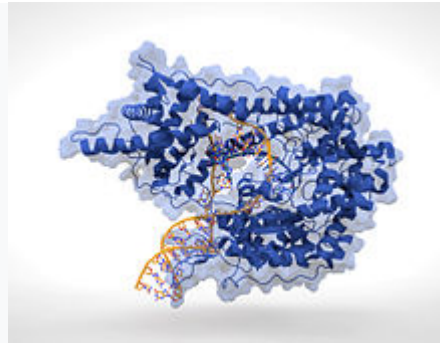
nucléotides d'ARN. En variante, une cellule peut simplement copier son information génétique dans un processus appelé réplication de l'ADN; Ici, l'accent est mis sur les interactions entre l'ADN et les autres molécules qui interviennent dans la fonction du génome.

#### **4. Gènes et génomes**

L'ADN génomique est étroitement et soigneusement emballé dans le processus appelé condensation de l'ADN , pour s'adapter aux petits volumes disponibles de la cellule. Chez les eucaryotes, l'ADN est situé dans le noyau de la cellule , avec de petites quantités dans les mitochondries et les chloroplastes . Chez les procaryotes, l'ADN est maintenu dans un corps de forme irrégulière dans le cytoplasme appelé nucléoïde (23) .L'information génétique dans un génome est contenue dans des gènes, et l'ensemble complet de cette information dans un organisme est appelé son génotype . Un gène est une unité de l'hérédité et est une région de l'ADN qui influence une caractéristique particulière dans un organisme. Les gènes contiennent un cadre de lecture ouvert qui peut être transcrit, et des séquences régulatrices telles que des promoteurs et des amplificateurs , qui contrôlent la transcription du cadre de lecture ouvert.

Dans de nombreuses espèces , seule une petite fraction de la séquence totale du génome code pour la protéine. Par exemple, seulement environ 1,5% du génome humain est constitué d'exons codant pour des protéines, avec plus de 50% d'ADN humain constitué de séquences répétitives non codantes (24). Les raisons de la présence de tant d' ADN non codant dans les génomes eucaryotes et les différences extraordinaires dans la taille du génome , ou *valeur-C* , entre les espèces, représentent un puzzle de longue date connu sous le nom « énigme valeur C » (25). Cependant, certaines séquences d'ADN qui ne codent pas les

protéines peuvent encore coder pour des molécules ARN fonctionnels non codants, qui sont impliquées dans la régulation de l'expression des gènes (26).



**T7 ARN polymérase (bleu) produisant un ARNm (vert) à partir d'une matrice d'ADN (orange) (27).**

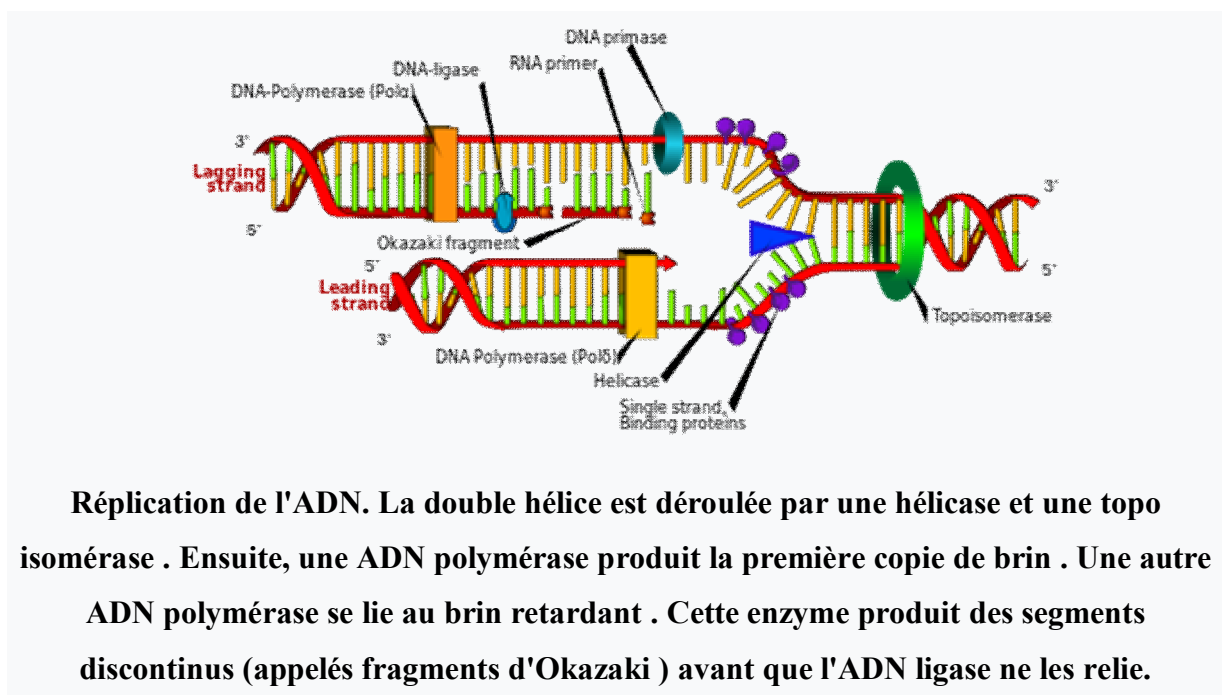
Certaines séquences d'ADN non codantes jouent un rôle structurel dans les chromosomes. Les télomères et les centromères contiennent généralement peu de gènes mais sont importants pour la fonction et la stabilité des chromosomes (28). Une forme abondante d'ADN non codant chez les humains est des pseudos gènes , qui sont des copies de gènes qui ont été désactivés par mutation (29). Ces séquences ne sont généralement que des fossiles moléculaires , bien qu'elles puissent parfois servir de matériel génétique brut pour la création de nouveaux gènes à travers le processus de duplication et de divergence des gènes (30) .

## **5. Transcription et traduction**

Un gène est une séquence d'ADN qui contient des informations génétiques et peut influencer le phénotype d'un organisme. Au sein d'un gène, la séquence de bases le long d'un brin d'ADN définit une séquence d'ARN messenger , qui définit alors une ou plusieurs séquences protéiques. La relation entre les séquences nucléotidiques des gènes et les séquences d'acides aminés des

protéines est déterminée par les règles de traduction , connues collectivement sous le nom de code génétique . Le code génétique consiste en des «mots» à trois lettres appelés *codons* formés à partir d'une séquence de trois nucléotides (par exemple, ACT, CAG, TTT).

En transcription, les codons d'un gène sont copiés dans l'ARN messenger par l'ARN polymérase . Cette copie d'ARN est ensuite décodée par un ribosome qui lit la séquence d'ARN en appariant l'ARN messenger par transfert de base , ce qui porte des acides aminés. Comme il y a 4 bases en combinaisons de 3 lettres, il y a 64 codons possibles ( $4^3$  combinaisons). Ceux-ci codent pour les vingt acides aminés standards , donnant à la plupart des acides aminés plus d'un codon possible. Il existe également trois codons «stop» ou «non-sens» signifiant la fin de la région codante; ce sont les codons TAA, TGA et TAG.



## 6. Réplication

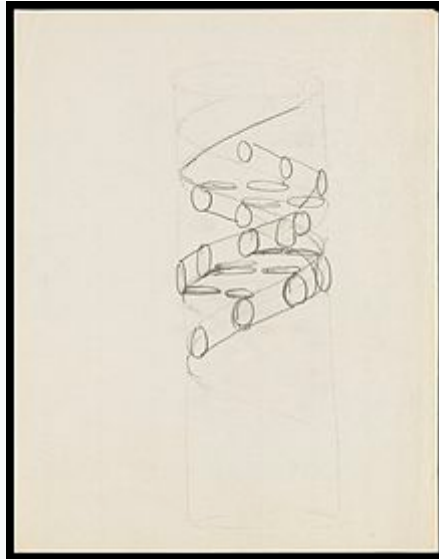
La division cellulaire est essentielle à la croissance d'un organisme, mais, quand une cellule se divise, elle doit répliquer l'ADN dans son génome de sorte que les deux cellules filles aient la même information génétique que leur parent. La structure double-brin de l'ADN fournit un mécanisme simple pour la réplication de l'ADN . Ici, les deux brins sont séparés puis la séquence d'ADN complémentaire de chaque brin est recrée par une enzyme appelée ADN polymérase . Cette enzyme construit le brin complémentaire en trouvant la base correcte par appariement de bases complémentaires et en le liant sur le brin d'origine. Comme les ADN polymérases peuvent seulement étendre un brin d'ADN dans une direction 5 'vers 3', différents mécanismes sont utilisés pour copier les brins antiparallèles de la double hélice (31). De cette façon, la base sur le vieux brin dicte quelle base apparaît sur le nouveau brin, et la cellule finit avec une copie parfaite de son ADN.

## 7. Histoire de la recherche sur l'ADN



**James Watson et Francis Crick (à droite), co-auteurs du modèle en double hélice, avec Maclyn McCarty (à gauche).**





**Croquis au crayon de la double hélice d'ADN par Francis Crick en 1953**

L'ADN a été isolé pour la première fois par le physicien suisse Friedrich Miescher qui, en 1869, a découvert une substance microscopique dans le pus des bandages chirurgicaux jetés. Comme il résidait dans les noyaux des cellules, il l'appelait "nucléine" (32,33). En 1878, Albrecht Kossel a isolé le composant non-protéine de "nucléine", acide nucléique, et a isolé plus tard ses cinq nucléobases primaires (34,35). En 1919, Phoebus Levene a identifié la base, le sucre, et le phosphate de l'unité de nucléotide (36). Levene a suggéré que l'ADN consistait en une chaîne d'unités nucléotidiques reliées entre elles par les groupes phosphate. Levene pensait que la chaîne était courte et les bases répétées dans un ordre fixe. En 1937, William Astbury a produit les premiers modèles de diffraction des rayons X qui ont montré que l'ADN avait une structure régulière (37).

En 1927, Nikolaï Koltsov a proposé que les traits hérités soient hérités par une «molécule héréditaire géante» composée de «deux brins miroirs qui se répliqueraient de manière semi-conservatrice en utilisant chaque brin comme modèle» (38,39). En 1928, l'expérience de Frederick Griffith dans a découvert que les traits de la forme "lisse" de *Pneumocoques* pourraient être transférés à la forme "rugueuse" des mêmes bactéries en mélangeant des bactéries "lisses" tuées avec la forme "rugueux" vivant ". Ce système a fourni la première indication claire que l'ADN porte l'information génétique - l'expérience Avery-MacLeod-McCarty - quand Oswald Avery, avec ses collègues Colin MacLeod et Maclyn McCarty, a identifié l'ADN comme le principe de transformation en 1943 (40). Le rôle de l'ADN dans l' hérédité a été confirmé en 1952 lorsque Alfred Hershey et Martha Chase dans l' expérience Hershey-Chase ont montré que l'ADN est le matériau génétique du phage T2 (41) .

À la fin de 1951, Francis Crick a commencé à travailler avec James Watson au Laboratoire Cavendish de l'Université de Cambridge . En 1953, Watson et Crick ont suggéré ce qui est maintenant accepté comme le premier modèle correct de structure d'ADN en double hélice dans la revue *Nature*(11). Leur modèle moléculaire d'ADN en double hélice était alors basé sur une image de diffraction des rayons X (appelée " Photo 51 ") (42) prise par Rosalind Franklin et Raymond Gosling en mai 1952, et l'information que les bases d'ADN sont appariées. Le 28 février 1953, Crick interrompit l'heure du déjeuner de son patron au pub The Eagle à Cambridge pour annoncer que lui et Watson avaient "découvert le secret de la vie" (43).

Des preuves expérimentales soutenant le modèle de Watson et Crick ont été publiées dans une série de cinq articles dans le même numéro de *Nature* (44). Parmi ceux-ci, l'article de Franklin et Gosling fut la première publication de leurs propres données de diffraction des rayons X et de leur méthode d'analyse originale qui soutenait en partie le modèle de Watson et Crick (45,46) ; ce numéro contenait également un article sur la structure d'ADN par Maurice Wilkins et deux de ses collègues, dont l'analyse et les modèles de rayons X de l'ADN B *in vivo* ont également soutenu la présence *in vivo* des configurations d'ADN double hélice proposé par Crick et Watson pour leur modèle moléculaire double-hélice de l'ADN dans les deux pages précédentes de *Nature* (47) . En 1962, après la mort de Franklin, Watson, Crick et Wilkins ont reçu conjointement le prix Nobel de physiologie ou de médecine (48).les prix Nobel sont attribués uniquement aux bénéficiaires vivants. Un débat se poursuit sur qui devrait être crédité de la découverte (49).

Dans une présentation influente en 1957, Crick exposait le dogme central de la biologie moléculaire , qui prédisait la relation entre l'ADN, l'ARN et les protéines, et articulait «l'hypothèse de l'adaptateur» (50). La confirmation finale du mécanisme de réplication impliqué par la structure en double hélice suivie en 1958 par l'expérience Meselson-Stahl (51). D'autres travaux de Crick et ses collègues ont montré que le code génétique était basé sur des triplets de bases non-chevauchants, appelés codons, permettant à Har Gobind Khorana , Robert W. Holley et Marshall Warren Nirenberg de déchiffrer le code génétique (52). Ces résultats représentent la naissance de la biologie moléculaire (53).



### *III. Le séquençage de l'ADN*

Un document historique publié par Frederick Sanger en 1977 a présenté une technique de séquençage de l'ADN qui allait devenir l'étalon-or pour les 30 prochaines années (54). La méthode originale de séquençage de l'ADN du terminateur didéoxy a été mise en œuvre avec une électrophorèse sur gel automatisée et une chimie de terminaison fluorescente, initialement sur gel en plaque et plus tard sur des systèmes à base de gel capillaire (55). La plupart des améliorations technologiques de cette méthode ont eu lieu lors du projet du génome humain (HGP) qui a été lancé en 1990 et au cours des 13 prochaines années, dans un effort international a séquencé le premier génome humain complet à un coût proche de trois milliards de dollars (56). Simultanément, une entreprise privée concurrente a également séquencé un génome humain. En 2001, les premiers résultats de cet effort gargantuesque ont été publiés (57), et en 2004 la séquence finie a été publiée (58). Le projet du génome humain a prouvé que le séquençage complet du génome (WGS) pouvait être réalisé, mais à un coût prohibitif à effectuer de façon régulière. Un effet du HGP a été une révolution technologique et l'arrivée de ce qu'on appelle le séquençage de prochaine génération (NGS). Le but de ces technologies était de réduire le temps, l'effort et le coût de WGS à un niveau où le génome humain pourrait être séquencé de façon routinière pour la recherche et les applications cliniques. Aujourd'hui, nous pouvons distinguer au moins quatre générations avec des caractéristiques distinctes. La transition du séquençage d'ADN de première génération à la deuxième génération a été perturbante, avec une production augmentant de plus de 5 ordres de grandeur et une baisse des coûts de plus de 5 ordres de grandeur. Avec les systèmes de séquençage de 2ème génération actuels, le WGS d'un génome humain à 30\* de couverture peut être atteint en 10 jours pour un coût inférieur à 10 000 \$.

Le développement technologique continue de faire baisser le coût des WGS humains en dessous de 1000 \$. L'incitation supplémentaire de la Fondation X Prize qui s'est engagée à attribuer 10 millions de dollars à la première équipe qui séquence avec précision le génome entier de 100 sujets en 30 jours pour 1000 \$ ou moins par génome a accentué l'intensité des développements technologiques dans le domaine (<http://genomics.xprize.org/>). L'objectif central du séquençage de 2ème génération est de développer des systèmes très bon marché et très rapides avec pour objectif principal de générer de gros volumes de séquences. En raison de la nature de la chimie, une séquence du génome se compose de milliards de fragments courts entre 70 et 400 bases de long, ce qui est un inconvénient et est traité par calcul. La qualité et la quantité des données sont des problèmes majeurs car les systèmes produisent maintenant 60 G Bases de séquences par jour sur un seul séquenceur d'ADN. En conséquence, le calcul pour l'analyse des données est ce qui fait le plus peser sur le séquençage de 2ème génération. L'augmentation de la production de données ne fera qu'accentuer ce problème et de nouvelles méthodes de traitement de cette quantité de données sont en cours d'élaboration. On peut s'attendre à ce que les techniques qui seront disponibles à l'avenir fourniront une séquence avec des caractéristiques nettement différentes de celles de la 2e génération actuelle. Il est également probable que des outils supplémentaires et raffinés pour certaines applications de séquençage seront nécessaires pour compléter les données produites à l'aide de séquenceurs de deuxième génération.

Même si le séquençage de deuxième génération n'est pas une ancienne technologie, on annonce des systèmes de troisième génération qui sont à nouveau perturbateurs. L'Union européenne a réalisé que de nombreux

problèmes dans le domaine du séquençage de l'ADN doivent être résolus et un projet européen intitulé READNA (Approches et outils de conversion pour l'analyse des acides nucléiques) a été financé avec 12 millions d'euros pour développer de nouvelles technologies d'analyse des acides nucléiques. La revue présentera les technologies de 3ème et 4ème génération qui sont développées dans le cadre de ce projet innovant.

## **1. Caractéristiques des générations de séquençage de l'ADN**

Il existe au moins quatre générations de méthodes de séquençage de l'ADN qui peuvent être classées selon les principales caractéristiques distinctives (tableau 1). D'autres classifications sont utilisées, cependant, nous nous concentrons sur des caractéristiques particulières pour cette classification qui rendent la catégorisation claire.

### **1.1. Séquençage de 1ère génération**

Pour le séquençage de 1ère génération, des échelles de fragments de la séquence sont générées soit par extension enzymatique d'une amorce hybridée à un pool de molécules modèles et en introduisant des terminaisons T, C, G ou A spécifiques le long du modèle (54). Alternativement, des clivages spécifiques à la base ont été introduits dans une méthode conçue par Maxam et Gilbert (59). Les échelles à fragments sont séparées par électrophorèse sur gel. Les produits sont soit révélés par un marquage radioactif et une exposition à un film ou par l'introduction de colorants fluorescents dans la réaction de terminaison et l'imagerie par fluorescence. Ceci est maintenant réalisé au cours d'une séquence de séquençage en temps réel.



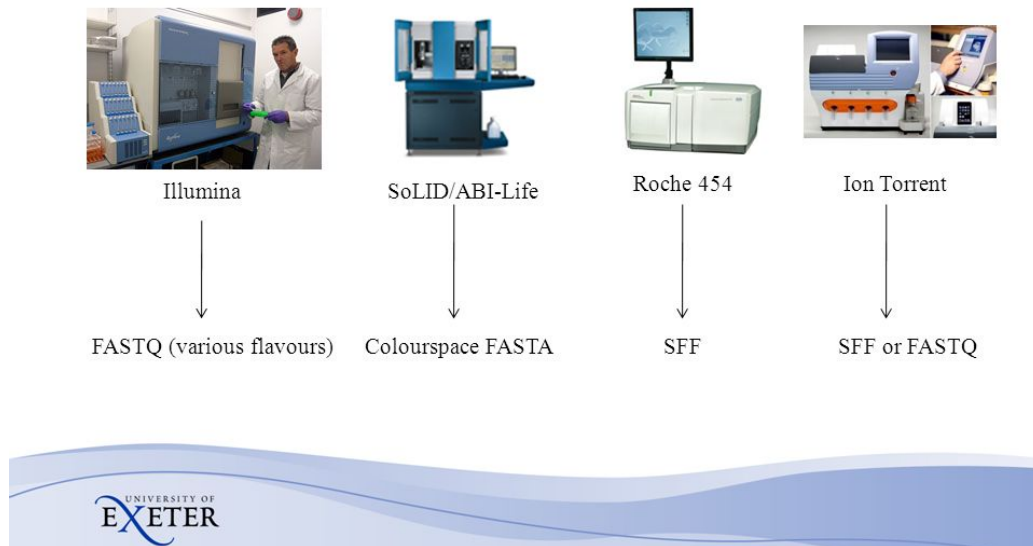
**Figure1 : Séquenceur de 1 ère génération**

## **1.2. Séquençage de 2ème génération**

Un changement de paradigme sous-tend le séquençage de 2ème génération (60). Des modèles d'ADN clonal individuels sont générés et tous séquencés en parallèle en utilisant des micros fluidiques. Les réactions de séquençage sont réalisées avec des cycles d'additions de bases et d'imagerie. Pour déterminer la séquence, des images consécutives sont superposées et des bases sont appelées aux positions de toutes les molécules modèles. Les additions de base peuvent être réalisées par polymérisation ou ligature de l'ADN et la visualisation se fait par chimiluminescence (Chimie de pyroséquençage utilisée dans l'instrument Roche FLX), fluorescence (Illumina et Life Technologies 5500) ou avec des changements de pH (dans l'instrument Ion Torrent). Les caractéristiques clés du séquençage de 2ème génération sont l'utilisation de nombreux modèles clonaux en parallèle et un processus de détermination de séquence utilisant la réplication enzymatique.



## 2<sup>nd</sup> generation sequencing output formats



**Figure2 : Séquenceurs de 2ème génération**

### 1.3. Séquençage de 2.5 génération

Le séquenceur SMRT de Pacific Biosciences est parfois appelé séquenceur de 3ème génération, mais il satisfait aux mêmes caractéristiques que les instruments de séquençage de 2ème génération en ce que les molécules clonales individuelles sont séquencées en utilisant un système de réplication de modèle enzymatique (61). Cependant, il rompt avec le concept de réseau aléatoire des systèmes de «deuxième génération» «traditionnels». Le système enzymatique polymérase est positionné au fond des puits des guides d'ondes en mode zéro (62). Les molécules d'ADN modèle sont capturées et l'incorporation de base est

surveillée par clivage du pyrophosphate lié au colorant fluorescent dans la fenêtre d'observation limitée en volume.



**Figure3 : Le séquenceur SMRT de Pacific Biosciences**

#### **1.4. Séquençage de 3ème génération**

La troisième génération est définie par la lecture directe des molécules individuelles. Aucun système enzymatique de réplication n'est utilisé pour identifier la séquence. Les séquenceurs nano pores (décrits ci-dessous) en sont des exemples.

#### **1.5. Séquençage de 4ème génération**

La quatrième génération est une méthode qui est actuellement encore très expérimentale. Il définit la réalisation d'une expérience de séquençage en contexte, telle que l'ADN de cellules individuelles dans une coupe histologique, donc dans leur contexte. Les applications de cette méthode seraient l'interrogation de séquences d'ADN définies telles que celles qui sont susceptibles d'avoir subi une mutation somatique, plutôt que la génération de séquences complètes de chaque cellule.

## **2. Dispositifs de séquençage d'ADN de 2ème et 2.5 génération (63)**

### **2.1. Dispositifs de séquençage d'ADN de 2ème génération qui ont réussi**

Le premier instrument à atteindre le marché était de 454 Corporation (64). Ce système utilise une PCR en émulsion avec des oligonucléotides attachés à des billes pour l'amplification clonale. La réaction de séquençage est réalisée en utilisant la chimie de pyroséquençage développée à l'origine par Ronaghi et al. (65-67). Pour cette méthode, les bases sont ajoutées une à la fois. La libération de pyrophosphate est convertie en luminescence. La quantité de lumière émise correspond au nombre de bases incorporées de ce type (T, C, G ou A). 454 Corporation a ensuite été achetée par Roche et le système a évolué vers la Roche GS FLX et une version miniaturisée appelée Roche GS FLX junior. Le deuxième système à atteindre le marché était le système 1G de Solexa (68). Ce système utilise la PCR en pont pour l'amplification clonale avec des oligonucléotides greffés en surface et l'extension d'une seule base d'amorce (également appelée SBE) pour le séquençage. Solexa a fusionné avec Illumina et le système a évolué vers l'Illumina HiSeq 2500, avec lequel un génome humain peut être séquençé à 30 \* en un peu plus d'une journée. Récemment, Illumina a sorti l'instrument de table MiSeq qui offre des temps d'exécution plus courts pour des volumes de données plus faibles. Le troisième système, le SOLiD a été publié par Applied Biosystems en 2007 (69). Dans ce système, la PCR en émulsion est combinée avec la ligature oligonucléotidique pour la détermination de la séquence. Après une fusion avec Invitrogen, le système est devenu le Life Technologies 5500. Life Technologies a récemment acquis Ion Torrent qui a

développé un système de séquençage de 2<sup>e</sup> génération avec un système de détection innovant (voir ci-dessous).

## **2.2. Dispositifs de séquençage d'ADN de 2<sup>e</sup> génération qui n'ont jamais réussi**

Il y a plusieurs systèmes de séquençage d'ADN de 2<sup>ème</sup> génération qui n'ont pas réussi. Le séquençage des polynômes a été implémenté dans un dispositif " open source " appelé le Polonator (70). Les concepts sous-jacents sont similaires à la ligation oligonucléotidique du système SOLiD utilisée pour la réaction de séquençage (71). L'Heliscope d'Helicos était un dispositif qui n'appliquait pas l'amplification clonale, mais capturait seulement les molécules d'ADN individuelles à la surface d'une lame de verre (72). Le séquençage a été réalisé par une extension d'amorce à base unique. Pour contourner les incorporations de base erronées qui entraînent des erreurs de séquençage si des molécules isolées sont analysées, le cycle de séquençage complet est répété, générant ainsi deux fois la séquence pour chaque molécule d'ADN matrice. Les différences entre les deux exécutions sont des erreurs de séquençage qui ne sont pas utilisées pour le traitement en aval. Tous ces systèmes utilisent la détection optique du processus de séquençage avec un système d'imagerie à haute résolution, suivi par un traitement d'image pour déterminer toutes les séquences individuelles d'une série.

## **2.3. Méthodes de détection innovantes**

Récemment, des systèmes de séquençage d'ADN de deuxième génération sont apparus utilisant des systèmes de détection alternatifs. IonTorrent utilise une incorporation de base enzymatique, une à la fois, similaire au système

Roche FLX. Les incorporations de base réussies sont déterminées par les changements de pH dus à la libération de  $H^+$  pendant l'incorporation de la base. Un microsystème avec plusieurs capteurs de pH en parallèle est utilisé pour détecter les bases. Complete Genomics utilise un schéma d'amplification aléatoire pour la préparation des échantillons. Les matrices de clones sont capturées avec un système sonde-ancree et les séquences sont déterminées en utilisant la ligature oligonucléotidique (73). Complete Genomics ne vend pas ses systèmes mais les utilise exclusivement pour effectuer le séquençage de service. Pacific Biosciences a développé une technologie qui va au-delà du séquençage de 2ème génération et de ce que nous avons appelé la génération 2.5. Dans cette technologie, une ADN polymérase est greffée au fond d'un micro puits. La détection se fait par le bas en tirant parti d'un principe physique quantique. Le diamètre du micro puits ne permet pas à la lumière de pénétrer profondément dans le puits (guides d'onde de mode zéro) et limite ainsi le volume d'irradiation (62). L'entrée de la lumière dans le puits est suffisamment profonde pour exciter le système de séquençage, mais pas les autres composants (fluorescents) qui sont en solution. Une molécule d'ADN modèle est capturée par l'ADN polymérase attachée au fond du puits. L'incorporation des nucléotides est suivie par la fluorescence suivante qui est liée au pyrophosphate qui est clivé du nucléotide dans la réaction de polymérisation. L'incorporation est suivie en temps réel. Cette technologie permet de séquencer de très longues bandes d'ADN. Des longueurs de lecture de plusieurs kb ont été rapportées. Il y a quelques années, Life Technologies a annoncé qu'elle développait une technologie similaire, sauf qu'elle n'utilisait pas les guides d'onde à mode zéro, mais qu'elle utilisait plutôt des nano cristaux et un système FRET. Cependant, ce système très intéressant n'a pas encore atteint le marché.

## 2.4. L'analyse des données

### Défis d'analyse de données associés à la 2e génération de séquençage

Les défis d'analyse de données associés aux résultats de séquençage de deuxième génération ont trois composantes principales: (1) Appels de base, (2) Alignement et (3) Appel de variantes. Pour la plupart des systèmes, l'appel de base est étroitement lié au système de séquençage et est effectué à l'aide d'un logiciel fourni par le fournisseur du dispositif de séquençage. L'alignement est une stratégie qui est choisie lorsque l'échantillon en cours d'examen a une séquence de référence. Les séquences produites sont comparées à la séquence de référence et la position la plus probable dans le génome pour une lecture donnée est déterminée. Les séquences qui se lient uniquement au génome et ne présentent pas de discordances par rapport à la référence sont les plus faciles à positionner. Évidemment, les séquences qui diffèrent d'un degré acceptable de la référence sont les plus intéressantes car elles sont utilisées pour identifier la variabilité. Il y a une batterie de différents outils qui ont été développés pour faire l'alignement, pour entrer dans les détails des avantages et des inconvénients peut exploser cette revue. Ce qu'il faut dire, c'est qu'il y a quatre critères d'équilibrage dans un aligneur, (a) est-ce qu'un aligneur est déterministe ou non, (b) est-il exhaustif, (c) quel pourcentage de lectures est-il capable de mapper à une référence? (d) quel effort de calcul est requis pour le faire. L'appel de variantes est intimement lié à l'alignement. Plus l'alignement est fiable, plus la tâche de l'appel de variante est facile. Encore une fois, il existe plusieurs variantes appelants disponibles et les détailler irait trop loin. De nos jours, la plupart des aligneurs parviennent assez bien à identifier des variants de nucléotides simples et de petits indels. Cependant, l'identification d'indels plus

importants et de variantes structurelles est plus exigeante et les logiciels pour ce faire sont encore soumis à de nombreux changements. Il est crucial, en termes d'effort de calcul, que l'étape d'alignement soit celle qui nécessite le plus de ressources de calcul.

Les stratégies d'alignement reposent sur l'hypothèse que la position la plus proche de la référence est l'endroit correct pour une séquence lue. Dans de nombreux cas, cela sera correct. Cependant, il existe une probabilité résiduelle que cela puisse être faux et que la position correcte soit à un endroit différent dans le génome où la correspondance est plus éloignée. Evidemment, cela influence les résultats. Une stratégie alternative à l'alignement est l'assemblage de séquence de novo. Plusieurs logiciels pour l'assemblage de novo à partir de lectures courtes de séquenceurs de 2ème génération ont été présentés. Cependant, ces approches nécessitent des architectures matérielles qui ne sont pas des produits et du temps de calcul, une sophistication requise et des efforts considérables. Beaucoup de développement est en cours dans ce domaine. Des stratégies utilisant une combinaison d'alignement et d'assemblage de novo pourraient également être bénéfiques. Dans tous les cas, lorsque la troisième génération deviendra disponible avec différentes caractéristiques de lecture, les stratégies d'analyse des données devront être adaptées. Ils ne seront pas nécessairement plus exigeants que pour le séquençage de 2ème génération.

### **3. Séquençage de 3ème génération – nano pores**

Les seuls systèmes notables répondant aux critères que nous avons définis ci-dessus pour le séquençage des acides nucléiques de 3ème génération sont les nano pores. Les méthodes montrées à ce jour mesurent directement la molécule

d'ADN sans avoir besoin de la soumettre à un système de réplication enzymatique.



**Figure 4 : Séquenceur ONT**

### **3.1. Introduction aux technologies nano pores :**

Le principe général du séquençage d'ADN à base de nano pores est extrêmement simple et a été démontré pour la première fois en 1996 (74). Essentiellement, le système consiste en un nano pore incorporé dans une membrane (biologique ou synthétique) immergée dans une solution saline. Un courant est appliqué au système qui entraîne les ions à travers le nano pore. L'arrivée d'une biomolécule chargée telle qu'une base d'ADN créera une résistance dans le flux d'ions qui conduira à des changements dans le courant électrique qui peut être mesuré. Des molécules de taille différente bloquent le nano pore par un degré différent et pour un temps différent (75). Si les quatre bases d'ADN (T, C, G et A) peuvent être distinguées lorsqu'elles traversent le nano pore, cela permettrait de mesurer une séquence d'ADN. En substance, le principe mis en évidence ci-dessus pointe vers une méthode capable de lectures d'ADN à lecture longue ( $> 2$  kb) (76). La clé du prix du séquençage des nano



pores est la taille nanométrique de ces pores, la possibilité de parallélisation et le faible coût des consommables (77).

Actuellement, deux voies principales sont poursuivies pour le séquençage des nano pores, la première est l'utilisation de nano pores biologiques et la seconde est l'utilisation de nano pores à l'état solide (78). Ci-dessous l'accent est mis sur les nano pores biologiques que les entreprises tirant parti des nano pores biologiques envisagent de commercialiser des plates-formes de séquençage nano pores en 2012 et 2013 (Oxford Nanopore Technologies (ONT) et Genia).

Actuellement, les nano pores biologiques ont la haute main car deux articles ont été publiés qui prouvent que le séquençage de l'ADN est possible en utilisant soit un pore  $\alpha$ -hémolysine (79) soit un pore MsPA (77). L'avantage des nano pores biologiques est qu'ils ont une taille bien définie, peuvent être modifiés très facilement et produits en masse restant homogène en taille et en structure. Dans les expériences de nano pores d' $\alpha$ -hémolysine, il faut noter que pour obtenir un système entièrement fonctionnel capable de discriminer chacune des quatre bases d'ADN, le nano pore doit être génétiquement modifié et en plus une molécule adaptatrice importante pour ralentir la translocation des bases d'ADN à travers le nano pore devait être fixé de manière non covalente dans le pore (80). Essentiellement, des expériences de mutagenèse ont conclu qu'une mutation spécifique au sein du bêta-baril était nécessaire pour obtenir une discrimination de base. Il décrit que l'attachement d'un dérivé de cyclo dextrine (CD) était essentiel pour la détection continue des bases circulant à travers le nano pore artificiel (81), en particulier la séparation optimale des bases a été obtenue en attachant le CD au résidu cystéine L135C dans le canon bêta du pore (82). Les conditions pour la construction d'hémolysine / CD modifiée ont été

optimisées pour les conditions requises par l'exonucléase pour être actives tout en maintenant la capacité de discrimination de base. En fin de compte, un équilibre a été recherché entre la tension appliquée et la concentration en sel utilisée pour s'assurer que l'exonucléase fonctionne correctement et que la majorité des bases d'ADN circulent à travers le pore (pour éviter les doubles lectures de bases) à une vitesse non trop rapide pour permettre la détection de molécules individuelles. Le papier MspA utilise les mêmes concepts généraux mais tire parti du plus petit site de constriction dans le pore MspA (diamètre 1,2 nm et longueur 0,5 nm) comparé à l' $\alpha$ -hémolysine (diamètre 1,5 nm et longueur 5 nm) qui montre un meilleur potentiel de discrimination entre les bases. Un inconvénient de la méthode est que les molécules d'ADN double brin sont utilisées pour bloquer la translocation des molécules d'ADN (par rapport à l'utilisation de la molécule CD dans le travail d' $\alpha$ -hémolysine), l'ADN doit donc être converti de façon à séparer les nucléotides monocaténaire doivent être mesurés, ce qui entraîne une forte préparation de pré-séquençage de l'ADN. Cependant, dans un développement récent, l'intégration de l'ADN polymérase de bactériophage (DNAP) phi29 dans le pore agit comme un moteur pour tirer le modèle monocaténaire à travers l'ouverture d'un nucléotide à la fois, permettant à un signal distinct d'être généré et chacun des nucléotides peut être mesuré (83). Un résultat similaire a été publié en utilisant le phi29 avec le pore  $\alpha$ -hémolysine et ce travail a été autorisé par l'ONT (84).

En ce qui concerne les nano pores à l'état solide, le graphène est un matériau qui pourrait être utilisé dans le futur (85), bien que dans un article crucial de Hagan Bayley, il reste encore du chemin à parcourir avant que les nano pores solides et en particulier le graphène une étape où ils peuvent être

utilisés pour séquencer régulièrement l'ADN. Il convient de mentionner que la technologie des nano pores à l'état solide produira des pores plus stables, dont la longueur et le diamètre peuvent être contrôlés, possèdent des propriétés de surface ajustables et présentent un plus grand potentiel d'intégration dans les dispositifs que des matrices biologiques (86).

À la veille de la commercialisation de deux systèmes nano pores (Oxford Nanopore Technologies et Genia Technologies, Inc.), il semble que de nombreux défis technologiques ont été résolus. Bien que les résultats attendus du séquençage de l'ADN soient la seule façon de décider si le défi des nano pores a été fissuré, il faudra encore quelques années pour avoir un système de séquençage de l'ADN qui offre le type de données requis par les chercheurs.

### **3.2. Acteurs commerciaux clés du séquençage nano poreux**

Avant de se concentrer sur READNA et son partenaire ONT, voici une courte compilation d'entreprises développant des plateformes de séquençage nano pores.

Genia Technologies Inc. \* - Séquençage de nanopores à l'aide de nano pores biologiques et contrôle électronique pour la détection du pyrophosphate marqué issu d'un nucléotide triphosphate lors de la réplication avec une polymérase fixée à l'entrée du nano pore (date de diffusion: 2013).

IBM / Roche \* - Le séquençage nano pore à l'aide de nano pores à l'état solide et le contrôle électronique déplace l'ADN à travers des nano pores intégrés dans une puce de silicium et la détection du courant de tunnel électronique (pas de date de diffusion).

NABsys Inc. \* - Le séquençage de nano pores utilisant des nano pores à l'état solide et le contrôle électronique déplace l'ADN avec des sondes oligonucléotidiques hybridées à travers des nano pores incorporés dans une puce de silicium et la détection du courant ionique (aucune date de diffusion).

NobleGen Biosciences Inc. \* - Plateforme " Optipore " - Nanopores Le séquençage à l'aide de nano pores à l'état solide et le contrôle électronique déplacent le gabarit d'ADN synthétique avec des balises oligonucléotidiques hybridées à travers des nano pores dans une puce de silicium, déplaçant les balises et détectant par fluorescence.

Oxford Nanopore Technologies \* - Le séquençage des nano pores à l'aide de nano pores biologiques et d'une enzyme de processus déplace l'ADN à travers des nano pores incorporés dans une bicouche lipidique et la détection du courant ionique (2012).

## **4. READNA et séquençage d'ADN**

### **4.1. Aperçu de READNA**

Le projet «Approches et dispositifs révolutionnaires pour l'analyse des acides nucléiques» READNA est un projet financé par le cadre de l'UE 7 qui a débuté le 1er juin 2008. Le consortium READNA comprend des projets visant à accélérer de nouvelles technologies de séquençage d'ADN révolutionnaires et des méthodes pour améliorer les méthodes d'analyse des acides nucléiques existantes. Des informations sur tous les développements de READNA peuvent être trouvées sur (<http://www.cng.fr/READNA>). Le but ultime sera de faire progresser les technologies permettant le séquençage d'un génome humain entier pour 1000 \$ en moins d'une journée. Ici, nous allons nous concentrer sur une technologie de séquençage d'ADN de 3ème génération qui a été développée par Oxford Nanopore Technologies au sein de READNA.

## **4.2. Séquençage de nano pores d'Oxford (séquençage d'exonucléases nano poreuses et séquençage de brins de nano pores) - Séquençage de 3ème génération**

Les technologies Nanopore sont développées par plusieurs partenaires READNA. Nous nous concentrerons ici sur les efforts de deux partenaires READNA, Oxford Nanopore Technologies (ONT) et l'Université d'Oxford (Professeur Hagan Bayley, également cofondateur de l'ONT).

Oxford Nanopore Technologies (ONT) poursuit deux types différents de technologie de l'ADN de troisième génération basée sur l'utilisation de nano pores pour séquencer l'ADN, à l'aide d'un nano pore d'exonucléase et d'un séquençage direct des brins. En février 2012, à l'occasion de la réunion annuelle Advances in Genome Biology & Technology à Marco Island, l'ONT a annoncé sa nouvelle plate-forme ADN " GridION " et présenté en même temps une version jetable de la même plate-forme. 'Min ION'. La plate-forme MinION a suscité beaucoup d'intérêt car lorsqu'elle sera commercialisée, elle deviendra le premier séquenceur jetable miniaturisé aussi petit qu'une clé USB et le coût sera d'environ 900 \$. L'ONT prévoit commercialiser la «première génération» de sa GridION et de sa MinION auprès d'un nombre restreint de clients.

### **4.2.1. Séquençage d'exonucléase de nanopores**

Dans la première méthode, une molécule d'ADN est capturée par une exonucléase processive qui est attachée de manière covalente à l'entrée du nanopore. L'exonucléase clive l'ADN base par base et le laisse tomber dans le nano pore. Les bases sont identifiées lorsqu'elles atteignent un confinement dans un nano pore d'a-hémolysine à une vitesse assurant une résolution de base

unique. La vitesse de translocation des bases d'ADN voyageant à travers un nano pore de 5 nm a été un problème majeur dans le passé car la résolution de base unique n'était pas réalisable. Essentiellement, un nano pore d'une longueur de 5 nm aura entre 10 et 15 nucléotides traversant le nano pore à un moment donné (76). En outre, la vitesse de translocation est de l'ordre de 1 nucléotide par ms, ce qui est trop rapide pour résoudre une base individuelle. Pour atteindre la détection de base unique, la vitesse doit être de l'ordre de 1 nucléotide / ms. Une façon de ralentir le débit est d'utiliser des adaptateurs non covalents et en particulier la cyclo dextrine (CD) a été l'adaptateur le plus efficace en combinaison avec le pore de protéine  $\alpha$ -hémolysine conçu et utilisé par ONT, un autre est l'adaptation de la concentration en sel (87). La découverte historique publiée par Clarke et al. montre que CD a été attaché à des cystéines spécifiques dans le nano pore pour réaliser l'identification de base (79). Les bases sont clivées par l'exonucléase et entrent immédiatement en contact avec l'adaptateur CD, se liant brièvement avant de se déplacer sur et à travers le nano pore en raison d'une différence de concentration en sel. La liaison temporaire crée un événement de perturbation qui est lu par une puce électronique associée au nano pore permettant de déterminer l'ordre des bases en raison d'un changement de courant. Les différences de courant sont telles que le système est capable de différencier les quatre bases d'ADN A, G, C et T et en plus le système est capable de distinguer les bases méthyl cytosine qui sont importantes dans l'analyse de la méthylation de l'ADN (79). Bien que ce soit une preuve de principe que les nano pores dans les bonnes conditions (modifications du pore, ajout d'un adaptateur, modification de la concentration en sel et contrôle de la température) puissent distinguer les quatre bases, d'autres développements seront nécessaires l'arène commerciale. De plus, le débit exigera des réseaux de

ces nano pores pour s'assurer que le séquençage des nano pores répond aux attentes actuelles d'une séquence complète du génome humain en une seule journée.

#### **4.2.2. Séquençage de brin de nano pores**

La deuxième méthode en cours de développement par ONT est le séquençage des brins de nano pores. La différence majeure avec le séquençage des exonucléases nano pores est qu'au lieu d'individualiser les bases en utilisant une enzyme exonucléase pour la détection par le nano pore, la molécule d'ADN est rendue monocaténaire et transloquée à travers le nano pore. Le problème qui doit être résolu pour que cette approche fonctionne pour le séquençage est la résolution d'une base individuelle dans le nano pore. La taille du nucléotide par rapport à la taille du nano pore résulte en plusieurs bases se trouvant dans la zone de constriction du nano pore et le courant est bloqué par un groupe de bases. Lorsque le brin d'ADN transloque à travers les nano pores, il doit être traduit en séquence. ONT ont publiquement présenté qu'ils ont réussi à utiliser cette approche pour séquencer 1 phage, en déconvoluant les changements de courant à travers le nanopore en utilisant un algorithme de Viterbi. La méthode a été privilégiée par ONT car elle est potentiellement plus rapide et plus précise car chaque base est physiquement attachée à la précédente ce qui permet de lire correctement l'ordre des bases. Le principal facteur limitant deviendra la processivité de l'enzyme utilisée, mais si suffisamment de créneaux sont disponibles, le débit dépassera certainement celui de toute autre plateforme de séquençage actuellement disponible à un coût espéré de 1000 \$ pour un génome complet. Les ONT doivent encore publier les résultats du séquençage de l'ADN, mais ont confirmé le séquençage d'un génome viral de 5 kb ayant un taux

d'erreur d'appel de base de 4%. La société vise à avoir un taux d'erreur maximum de 2% lors de la commercialisation en 2012. Une caractéristique impressionnante de cette technologie est que les lectures individuelles sont de l'ordre de 10s de Kbases. Les données de ce type révolutionnent le séquençage de l'ADN.

#### **4.3. Séquençage in situ (séquençage des cellules individuelles) - Séquençage de 4ème génération**

Le séquençage de 1ère génération a prévalu pendant près de 30 ans, la 2ème génération se poursuit depuis 7 ans, la 3ème génération est proche et la 4ème génération est déjà à l'horizon. Les cellules voisines dans le contexte de la tumeur ne sont pas identiques et il deviendra de plus en plus important de pouvoir mesurer cela. Le groupe de Mats Nilsson a fourni des exemples d'analyses de transcrits de nombreuses cellules différentes en parallèle dans leur contexte cellulaire (88). Ceci est une preuve de concept du séquençage de 4ème génération.

#### **4.4. Manipulation de chromosomes complets, techniques de cartographie optique**

En raison des procédures de préparation des échantillons utilisés dans le séquençage de deuxième génération, les informations contextuelles sont perdues. L'information haplotypique est particulièrement utile dans l'assemblage de nouveaux génomes, mais elle est également importante pour les problèmes biologiques. La présence de deux variants mutants sur des chromosomes identiques ou opposés est biologiquement pertinente. Les méthodes de prédiction des haplotypes utilisent le contexte familial ou l'inférence statistique.



Dans les molécules d'ADN réorganisées aléatoires, telles qu'elles sont rencontrées dans de nombreuses cellules tumorales, ces stratégies se décomposent. Cependant, la détermination moléculaire est possible en utilisant des méthodes telles que la cartographie optique (89). Au sein de READNA, nous avons consacré beaucoup d'efforts aux techniques de manipulation des molécules d'ADN individuelles et avons développé des systèmes micro- et nano-fluidiques pour les étirer et les rendre disponibles pour une étude plus approfondie (90).

En conclusion, le séquençage de l'ADN a atteint un point où il peut être appliqué de manière holistique. Les transitions entre les générations de méthodes de séquençage de l'ADN ont perturbé les changements dans les ordres de grandeur de la production et les changements complets dans la nature du type de séquence fourni par une méthode et la diminution spectaculaire des coûts. Alors que le séquençage d'ADN de première génération a prévalu pendant 30 ans, les dernières années ont vu une vague de changements avec la 2ème génération et déjà l'apparition de méthodes d'analyse des acides nucléiques de 3ème et 4ème génération. Chaque nouvelle génération n'a pas remplacé la précédente. Il existe un chevauchement important de différentes méthodes qui aident à compléter l'ensemble d'outils requis pour l'analyse génomique complète. La recherche ou la question clinique à portée de main détermine le dispositif optimal ou son utilisation. La deuxième génération qui est l'état de l'art actuel et la méthodologie la plus répandue, par la nature de sa séquence (beaucoup de lectures courtes) met une pression considérable sur les procédures de calcul et d'analyse des données. Au fur et à mesure que nous avançons dans le séquençage de 3ème génération, avec des lectures plus longues, les demandes de

calcul vont probablement diminuer. Cependant, en raison de l'augmentation de la production globale, le traitement bio-informatique en aval des résultats deviendra encore plus important. Les méthodes d'analyse des acides nucléiques de quatrième génération remettront en question les procédures de traitement de l'image et la résolution qu'elles peuvent atteindre.

TABLE 1

Overview of sequencing methods and characteristics

|                                           | Generation | Input          | Template enrichment           | Nature of sequencing template | Sequencing chemistry                  | Detection                              | Particularity                 | Reference                                                             |
|-------------------------------------------|------------|----------------|-------------------------------|-------------------------------|---------------------------------------|----------------------------------------|-------------------------------|-----------------------------------------------------------------------|
| Sanger sequencing                         | 1st        | DNA            | PCR                           | Pool                          | Primer extension, dideoxy termination | Radioactivity, fluorescence            | Used for Human Genome Project | Sanger [1]                                                            |
| Maxam-Gilbert                             | 1st        | DNA            | Sufficient template required  | Pool                          | Chemical base-specific cleavage       | Radioactivity                          |                               | Maxam-Gilbert [6]                                                     |
| 454-Roche                                 | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Primer extension one base at a time   | Real-time chemiluminescence            | Pyrosequencing                | Margulies [11]                                                        |
| Polonator                                 | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Oligonucleotide ligation              | Fluorescence imaging                   |                               | Shendure [17]                                                         |
| Solexa-Illumina                           | 2nd        | DNA            | Bridge PCR                    | Amplified Clone               | Single-base primer extension          | Fluorescence imaging                   |                               | Bentley [15]                                                          |
| Helicos                                   | 2nd        | DNA            | None                          | Individual DNA molecule       | Single-base primer extension          | Fluorescence imaging                   | Single molecule               | Harris [19]                                                           |
| Applied Biotechnologies-Life Technologies | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Oligonucleotide ligation              | Fluorescence imaging                   | Two-base encoding             | McKernan [39]                                                         |
| Complete Genomics                         | 2nd        | DNA            | Rolling circle replication    | Amplified Clone               | Oligonucleotide ligation              | Fluorescence imaging                   | Proprietary system            | Drmanac [20]                                                          |
| IonTorrent                                | 2nd        | DNA            | Emulsion PCR                  | Amplified Clone               | Primer extension one base at a time   | pH measurement in semiconductor device | Non-optical                   | Rothberg [18]                                                         |
| Pacific Biosciences                       | 2.5th      | DNA            | None                          | Individual DNA molecule       | Primer extension                      | Real-time fluorescence imaging         | Zero-mode wave guides         | Eid [9]                                                               |
| Oxford Nanopore (Exonuclease)             | 3rd        | DNA            | None                          | Individual DNA molecule       | Exonuclease digestion                 | Nanopore                               | Exonuclease nanopore          | <a href="http://www.nanoporetech.com">http://www.nanoporetech.com</a> |
| Oxford Nanopore (Strand)                  | 3rd        | DNA            | None                          | Individual DNA molecule       | Direct measurement                    | Nanopore                               |                               | <a href="http://www.nanoporetech.com">http://www.nanoporetech.com</a> |
| <i>In situ</i>                            | 4th        | Tissue section | <i>In situ</i> rolling circle | Rolling circle amplicon       | Primer extension                      | Fluorescence imaging                   |                               | Larsson [36]                                                          |

Tableau1 : Vue d'ensemble des méthodes de séquençage et des caractéristiques



## *IV. La médecine personnalisée*

## **1. Qu'est-ce que la médecine personnalisée ?**

La médecine personnalisée est un domaine vaste, à développement rapide des soins de santé fondé sur des informations uniques de chaque personne cliniques, génétiques, génomiques et environnementaux. Les soins de santé qui embrasse la médecine personnalisée est une approche intégrée, coordonnée et fondées sur des preuves pour l'individualisation des soins aux patients dans tout le processus de la état normal à la maladie .La médecine personnalisée utilise notre compréhension moléculaire de la maladie pour améliorer les stratégies de soins de santé préventives alors que les gens sont toujours bien et commencent les traitements médicamenteux dans les premiers stades de la maladie. L'objectif global de la médecine personnalisé est d'optimiser les soins médicaux et les résultats pour chaque individu, ce qui entraîne une personnalisation sans précédent des soins aux patients.

### **1.1. La Médecine Personnalisée et l'Histoire de santé familiale**

L'Histoire de santé familiale (HSF) est un outil simple mais précieux pour la fourniture de renseignements personnels sur la santé et sur les risques à certains maladies. Reflétant la combinaison complexe de facteurs partagés génétiques, environnementaux, et de style de vie, un HSF robuste peut approximativement donner des renseignements sur les risques génétiques/génomiques et l'intègrent dans les soins aux patients. L'évaluation du HSF aiderait à identifier les personnes à risque plus élevé de maladie, permettant des étapes préemptives et préventives, y compris les changements de style de vie, des projections de la santé, essai et le traitement précoce selon le cas (voir tableau 2). Par conséquent, HSF détient un énorme potentiel pour améliorer les soins de santé préventifs parmi les populations larges et

diversifiées de façon personnellement Pertinente. Toutefois, l'évaluation et intégration de l'HSF n'ont pas été embrassé par la communauté sanitaire et demeure une ressource largement inexploitée (91). Le défi de l'intégration de l'HSF dans la santé du public implique trois composantes essentielles: (a) des méthodes de collecte standard et accessibles; (b) l'accès du fournisseur de soins de santé; et (c) des conseils cliniques pour l'interprétation et l'utilisation.

Actuellement, la collecte d'informations de l'HSF peuvent être incomplète, peuvent être difficile à interpréter, et / ou peuvent varier considérablement dans le contenu entre les fournisseurs de soins de santé du patient. En outre, les fournisseurs peuvent avoir une connaissance et une formation insuffisante afin d'interpréter fidèlement l'HSF pour déterminer le risque des syndromes génétiques héréditaires. (92)

**Tableau 2 : Risques associés au fait d'avoir un parent au premier degré atteint d'une maladie (93)**

| <b>Condition</b>     | <b>Rapport de cotes de prévalence</b> | <b>Références</b> |
|----------------------|---------------------------------------|-------------------|
| Diabète de type 2    | 2.3-2.8                               | 84, 85            |
| Maladie coronarienne | 4.5                                   | 186               |
| Cancer du côlon      | 2 – 2,6                               | 90                |
| Le cancer du sein    | 1.6 – 2                               | 90                |
| Cancer du poumon     | 1.7 – 2,5                             | 90                |

## **1.2. Évaluation du risque de maladie chronique**

Un élément fondamental de la médecine personnalisée est une évaluation standard des risques pour la santé (ERS) visant à évaluer la probabilité qu'un individu développe les maladies chroniques les plus courantes (ou les événements pathologiques). Les ERS fondées sur des données probantes couplées à des modèles prédictifs faciliteront l'évaluation du risque de maladie d'un patient et la priorisation des stratégies cliniques pour y remédier. L'un des ERS les plus largement reconnus est le modèle de maladie cardiaque coronarienne de Framingham, développé à partir de l'étude de Framingham sur le cœur débuté en 1948 (94). L'évaluation du risque de cancer du sein selon le modèle Gail et ses versions modifiées sont également des outils largement acceptés (95).

Malgré la valeur prouvée de nombreux outils ERS, ils n'ont généralement pas été inclus dans l'évaluation formelle des patients en raison du manque de normes des données cliniques requises ou les algorithmes utilisés, ainsi que le manque d'intégration dans les systèmes de technologie de l'information sur la santé (96). Cependant, les ERS devraient constituer la base de la prévision et de la stratification des risques qui sont fondamentales pour les soins de santé personnalisés. Et à mesure que de nouvelles informations biologiques et des marqueurs moléculaires prédictifs sont découverts et validés, il est essentiel que ces nouveaux marqueurs soient évalués itérativement pour leur capacité à augmenter la capacité prédictive des modèles cliniques existants. S'ils le font, ils devraient être incorporés dans la ERS et devenir le point de départ pour des soins de santé personnalisés.

### **1.3. Système de soutien à la décision clinique**

Pour optimiser l'utilisation de la HSF et des ERS, les systèmes de soutien à la décision clinique (SSD) fournissent aux médecins et aux patients des connaissances filtrées conçues pour orienter et améliorer les soins de santé sur le lieu des soins (97). Les systèmes SSD informatisés sont de plus en plus utilisés et des informations spécifiques au patient — poids, médicaments actuels, antécédents médicaux et même l'HSF — peuvent être intégrées dans des formulaires conviviaux pour aider à gérer le diagnostic et le traitement.

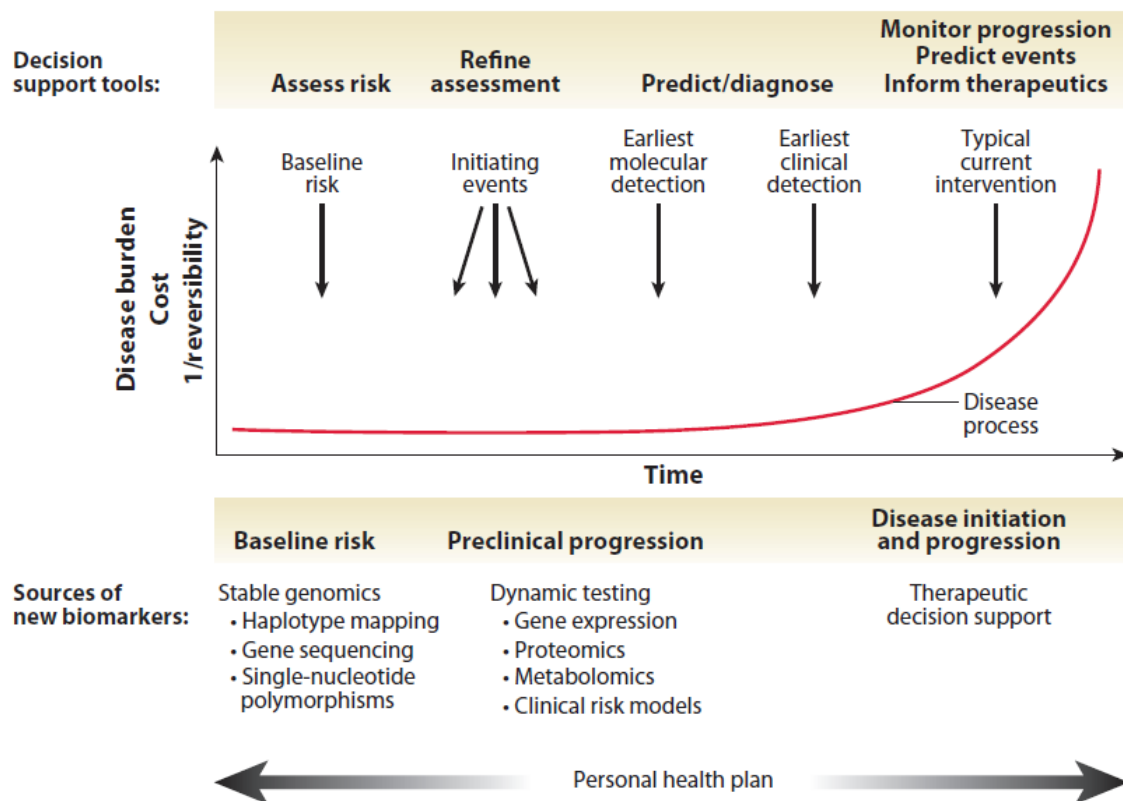
Il a été démontré que les systèmes ERS améliorent les pratiques de prescription, améliorent les soins préventifs et améliorent le respect des normes de soins fondées sur des données probantes (98, 99, 100). À mesure que la ERS s'intégrera aux dossiers de santé électroniques, l'utilisation des données personnelles sur la santé deviendra transparente, donnant aux cliniciens accès à l'information et aux connaissances des patients pour aider à créer des plans de santé adaptés aux besoins individuels.

### **1.4. La médecine personnalisée et le génome humain**

La médecine personnalisée est maintenant étayée par l'incorporation d'informations provenant des génomes, de leurs séquences et de leurs produits dérivés (tels que l'ARN, les protéines et les métabolites) dans la prise de décision médicale: les tests de santé basés sur le génome peuvent être utilisés pour prédire le risque, dépister les porteurs, établir des diagnostics cliniques et des pronostics pour les individus, et diriger la gestion clinique (101). Ainsi, l'information sur le génome humain permet maintenant aux fournisseurs de créer des plans de soins optimisés à chaque stade d'une maladie, en mettant l'accent sur les soins de santé réactifs plutôt que préventifs (102, 103, 104).

Tout au long du continuum de la santé à la maladie (figure 1), il existe maintenant plusieurs moments importants au cours desquels les applications génomiques personnalisent les soins de santé (105, 106). La susceptibilité à la maladie et le risque peuvent maintenant être quantifiés et anticipés pendant la santé et même à la naissance en utilisant des évaluations basées sur l'ADN qui ne changent pas au cours de la vie d'une personne. Les profils de transcription, les profils d'expression des protéines et les niveaux de métabolites combinés avec les modalités d'imagerie dynamique peuvent fournir des moyens plus précis de dépister les personnes à haut risque de développer une maladie dès les premières manifestations moléculaires alors que la maladie est subclinique. Cette même information peut fournir un diagnostic définitif et une classification moléculaire qui prédit le pronostic. De même, la pharmacogénomique - l'étude de la variation génétique et de l'efficacité et de la toxicité des médicaments - devient un exemple important d'information génomique appliquée à la médecine clinique (108). La sélection des médicaments peut être guidée à la fois par la constitution génétique sous-jacente du patient ainsi que l'architecture moléculaire de la maladie chez l'individu. Un bon exemple de cette approche peut être vu avec le dosage de warfarine, où les fournisseurs peuvent maintenant utiliser les systèmes de SSD en ligne (<http://www.warfarindosing.org>) pour déterminer la dose appropriée basée sur l'information génotypique et clinique d'un patient (109).





**Figure 5 : Le rôle de l'information génomique dans le continuum de la santé à la maladie.(Référence 93)**

## 2. La boîte à outils du génome humain

Le tableau 3 présente les principales plateformes technologiques du génome formant une boîte à outils pour l'exploration de l'impact du génome et de ses produits exprimés sur la santé et les états pathologiques. L'application clinique de ces technologies constitue la base de soins médicaux personnalisés. (Référence 103)

## **2.1. Variation à l'échelle du génome**

La variation dans le génome humain provient de plusieurs types d'altérations de séquence, y compris les mutations ponctuelles impliquant des changements de paires de bases mono nucléotidiques [polymorphismes mono nucléotidiques (SNP)] et les réarrangements structuraux de la séquence d'ADN (109, 110).

La plupart de nos connaissances sur la variation du génome humain proviennent de données de séquences d'ADN. La technologie de séquençage de nouvelle génération prend en charge le séquençage à haut débit et réduit les coûts au point où un génome peut être séquencé pour 10 000 \$, et on estime que ces coûts seront de 1 000 \$ en un an ou deux (111). À ce niveau de prix, le séquençage du génome d'un patient se situera dans la plage des tests médicaux standards (112, 113). Le séquençage de nouvelle génération a permis le reséquençage de plusieurs génomes humains, mettant en évidence l'évolution des plateformes de séquençage et leur capacité à étendre la couverture de lecture avec une plus grande résolution (114, 115, 116, 117). Le séquençage de nouvelle génération est maintenant utilisé pour comprendre plusieurs maladies - par exemple, le cancer, les maladies génétiques rares et les infections microbiennes - dans le but d'élucider les variantes géniques fonctionnelles sous-jacentes qui sont peut-être responsables de l'état pathologique (118, 119, 120, 121, 122, 123). L'utilisation du séquençage du génome entier avance à la clinique, où les données de séquences d'ADN ont été utilisées pour établir des diagnostics définitifs et même guider le traitement (124, 125, 126).

Le projet international HapMap (<http://hapmap.ncbi.nlm.nih.gov>), achevé en 2005, a développé une carte haplotype du génome humain pour détecter les

régions de déséquilibre de liaison élevé et de faible diversité haplotypique, identifiant les quelques SNP qui sont les variantes causales de la maladie (127, 128, 129). Ces données sont à la base des études d'association pangénomique (GWAS), une stratégie basée sur le génome qui utilise des comparaisons de SNP entre les populations pathogènes et témoins et fournit une méthode non biaisée de relier la variation du génome humain à la maladie clinique. En mai 2011, plus de 800 GWAS avaient été publiés sur 150 maladies et caractères humains, signalant plus de 2 400 SNP avec des associations et des rapports de cotes statistiquement significatifs (130, 131). Par exemple, dans la maladie de Crohn, des études d'association ont découvert 32 variants contribuant à la maladie, avec des variants dans les gènes NOD2 et IL23R associés à une augmentation du risque de maladie (132). Dans le diabète de type 2, au moins 20 variants ont été décrits, avec des variants plus récents émergeant de plus grands ensembles de données GWAS dans le contexte de l'homéostasie du glucose à jeun (133, 134). De même, les GWAS dans la maladie coronarienne (CAD) ont identifié plus de 100 variantes pouvant contribuer au risque de coronaropathie (135, 136, 137, 138, 139). Nombre de ces études ont indépendamment vérifié l'importance du locus 9p21.3. Les gènes contenus dans cette région codent pour l'ANRIL, un ARN régulateur non codant supposé réguler le silençage épigénétique par action sur le locus p15 / CDKN2B-p16 / CDKN2A-p14 / ARF, et les protéines kinases jouant un rôle inflammatoire induit par le TGF- $\beta$  dans l'athérosclérose (140, 141, 142, 143). D'autres ont suggéré que les SNP dans des amplificateurs distincts trouvés dans la région affectent la liaison STAT1 d'une manière dépendant de l'interféron  $\gamma$  dans les cellules endothéliales vasculaires humaines, reliant la susceptibilité CAD à la signalisation inflammatoire dans des cellules vasculaires spécifiques (144). Enfin, une utilisation précoce réussie du développement de

thérapeutiques basées sur les données GWAS dans la dégénérescence maculaire liée à l'âge où des thérapeutiques ciblées ont été développées pour les voies inflammatoires médiées par le complément identifiées par les GWAS (145,146).

Étonnamment, pour la grande majorité des GWAS qui ont été menées, les rapports de cotes des allèles associés au phénotype de la maladie ont été  $<2$  et peuvent ne pas contribuer significativement au risque de maladie. Il y a cependant quelques exceptions: Dans un GWAS de Ge et al. (147), un polymorphisme sur le chromosome 19, 3 kb en amont de l'IL-28B, codant pour l'interféron lambda-3, était associé à un changement de deux fois de la réponse au traitement. Cette découverte a été cliniquement confirmée dans une étude de Thompson et al. (148) ont génotypé des patients recevant un traitement contre le virus de l'hépatite C et ont trouvé que le même polymorphisme était un bon prédicteur d'une réponse virologique soutenue (rapport de cotes 5,2, intervalle de confiance à 95%, 4,1-6,7). Dans l'étude SEARCH Collaborative Group, un polymorphisme localisé sur SLCO1B1, un gène régulant l'absorption hépatique des statines, était associé à un risque accru de myopathie suite à un traitement par statine (odds ratio 4,5, intervalle de confiance 95%, 2,6-7,7) (149). Bien que la pertinence clinique des GWAS ne soit pas aussi répandue que nous l'avions imaginé, ces études ont révélé de nombreux nouveaux gènes contribuant à la pathogenèse qui ont fourni de nouvelles perspectives sur la biologie sous-jacente, les mécanismes de la maladie et les approches thérapeutiques potentiellement nouvelles.

**Tableau 3 : La boîte à outils du génome humain (Référence 93)**

Table 2 The human genome toolbox

| “-omics” approach                          | Generated information                                  | Applications                                                                     | Notable examples                                                                                  |
|--------------------------------------------|--------------------------------------------------------|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|
| Human genome sequence (genomics)           | Whole-genome sequence, SNPs, and CNVs (~10–15 million) | Disease mechanisms<br>Disease diagnosis<br>Pharmacogenomics                      | Age-related macular degeneration (120), HCV virologic response (1), AML (32), warfarin dosing (6) |
| Gene expression profiles (transcriptomics) | Microarrays and RNA sequencing (~25,000 transcripts)   | Disease mechanisms<br>Disease diagnosis<br>Disease prognosis<br>Pharmacogenomics | AML (71), ALL (94), ACS (20), breast cancer (161)                                                 |
| Proteome (proteomics)                      | Protein profiles of specific protein products          | Disease diagnosis                                                                | ACS (143)                                                                                         |
| Metabolome (metabolomics)                  | Metabolic profiles (~1,000–10,000 metabolites)         | Disease mechanisms<br>Pharmacogenomics                                           | ACS (182), drug toxicity (44), cancer profiling (76), CAD (193)                                   |

Abbreviations: ACS, acute coronary syndromes; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CAD, coronary artery disease; CNV, copy number variation; HCV, hepatitis C virus; SNP, single-nucleotide polymorphism. Table adapted from Reference 66.

## 2.2. Transcriptomique

La transcriptomique est l'étude génomique des niveaux d'expression de l'ARN dans une cellule, un tissu ou un fluide biologique, et comprend un spectre de molécules allant de l'ARN messenger (ARNm) aux ARN non codants. Les analyses statistiques et la classification des gènes coexprimés à l'aide de méthodes paramétriques ou non paramétriques peuvent être appliquées pour classer la maladie ou prédire les états pathologiques futurs, ou pour comprendre les mécanismes biologiques sous-jacents de la maladie (150).

Dans le cancer, les données de biopuces ont été utilisées pour le diagnostic, le pronostic et la réponse au traitement (151,152). Ces données ont révélé des sous-ensembles discernables de classes moléculaires dans de nombreux cancers, y compris les cancers du sein, du poumon, du sang et du mélanome (153, 154, 155, 156). Chez les patients atteints de leucémies aiguës, ce diagnostic moléculaire est particulièrement utile pour la prise de décision clinique, où le traitement peut différer en fonction du sous-type de maladie en réseau (157).

L'application de la technologie des micros réseaux à la classification des maladies a été appliquée dans d'autres maladies complexes, y compris les maladies cardiovasculaires, les maladies rhumatismales et leurs voies inflammatoires, et les maladies neurologiques telles que la sclérose en plaques et les troubles psychiatriques tels que la schizophrénie, dépression (158, 159, 160, 161, 162, 163).

Une caractéristique du transcriptome qui a émergé est l'expression des profils d'ARN non codants, en particulier les petits ARN interférents (siRNA) et les microARN (miARN). D'abord identifiés en 1993 chez *Caenorhabditis elegans*, les miARN ont généralement une longueur de 22 nucléotides et inhibent l'expression génique à un niveau post-transcriptionnel en se liant à des régions non traduites complémentaires des ARNm cibles (164,165). Ils ont joué un rôle dans plusieurs maladies, de la régulation des cancers proliférants au développement des cardiomyopathies, de la réponse immunitaire et de la schizophrénie, et même à la promotion de la réplication du virus de l'hépatite C (166, 167, 168, 169, 170, 171). Les modèles de miARN différentiellement exprimés peuvent être utilisés cliniquement dans le diagnostic et le pronostic de la maladie (172, 173,174). Compte tenu de leur rôle dans la complexité de la régulation des gènes, les miARN et même les miARN et les siRNA sont testés (175). Un essai de patients humains avec des tumeurs solides réfractaires a montré des résultats prometteurs après la livraison de siRNA en utilisant un système de nanoparticules ciblées (176). Cependant, l'analyse des miARN peut être difficile en raison de leur manque de spécificité, conduisant à des effets hors cible: En raison de leur petite taille, ils peuvent potentiellement influencer l'expression de plusieurs ARNm différents (177, 178).

Enfin, les technologies de nouvelle génération utilisées pour faire progresser le séquençage du génome entier peuvent être exploitées pour analyser les transcriptomes: le séquençage de l'ARN (RNA-Seq), initialement utilisé pour examiner les courts ARN régulateurs, a été appliqué à l'analyse du transcriptome humain entier pour continuer notre compréhension de la régulation de la transcription (179, 180, 181, 182, 183). Par exemple, les transcriptomes liés au cancer ont été étudiés par RNA Seq, fournissant une nouvelle approche pour comprendre les événements de fusion et de translocation de gènes (184, 185, 186, 187).

### **2.3. Protéomique**

L'étude de la protéomique fait référence à l'étude à grande échelle des protéines, et le protéome se réfère souvent à l'ensemble complet des protéines et de leurs divers dérivés. Dans le contexte de la santé et de la maladie, la protéomique cherche à définir l'ensemble complet des protéines associées à un état physiologique ou pathologique particulier et continue d'être un domaine prioritaire pour la découverte de biomarqueurs et la médecine moléculaire.

Cette zone a été dominée par des approches isotopiques stables, mais les méthodes quantitatives sans marqueur récemment développées, qui reposent sur l'intensité mesurée d'un ion peptidique par rapport à son intensité dans d'autres échantillons, ont l'avantage d'un débit plus élevé et moins d'étapes de manipulation (188,189). Bien que cette technologie soit relativement immature dans ses applications à la santé humaine et à la maladie par rapport à l'ADN, l'ARN et le profil métabolique, ces méthodes, combinées au développement de la spectroscopie de masse, feront progresser la protéomique dans la classification des maladies, le diagnostic, le pronostic et la pharmacogénomique au cours des prochaines années.

## **2.4. Métabolomique**

La métabolomique utilise la spectroscopie de masse et la spectroscopie de résonance magnétique nucléaire pour mesurer les changements dans les petites molécules non protéiques liées à un état biologique ou physiologique. On estime que le métabolome humain contient environ 5 000 métabolites discrets de petites molécules, et l'identification des changements métaboliques associés à la maladie suggère immédiatement des cibles médicamenteuses enzymatiques. Cette approche a été appliquée à un certain nombre d'états pathologiques chroniques, tels que le diabète, l'obésité, les maladies cardiovasculaires, le cancer et les troubles mentaux (190, 191, 192, 193). Dans le domaine des maladies cardiovasculaires, deux études distinctes ont défini des profils de métabolites pour l'ischémie et la coronaropathie, respectivement (194, 195). Le profilage métabolomique a également été utilisé directement pour évaluer la toxicité des médicaments (196, 197, 198)

## **2.5. Épigénomique**

L'épigénomique fait référence à la programmation génétique qui résulte principalement de la méthylation de l'ADN (199). La base moléculaire de cette régulation implique la méthylation de la cytosine et est maintenue par une famille d'ADN méthyltransférases; des quatre bases comprenant l'ADN, seule la cytosine peut être méthylée et non méthylée, et seulement dans le contexte des dinucléotides CpG, qui chevauchent souvent la région promotrice des gènes. Les gènes transcrits activement sont généralement dans un état non méthylé, tandis que les gènes qui sont réduits au silence sont fortement méthylés (200, 201). L'analyse des tumeurs montre une hypométhylation chez les oncogènes et une hyperméthylation régionale chez les gènes suppresseurs de tumeurs (202, 203,



204). Des changements de méthylation ont également été mis en évidence dans d'autres maladies telles que le diabète de type 2, les pathologies cardiovasculaires et les maladies auto-immunes telles que le lupus érythémateux disséminé (205, 206, 207, 208). Les Instituts nationaux de la santé ont lancé le projet Human Epigenome pour identifier, cataloguer et interpréter les profils de méthylation de l'ADN de tous les gènes humains dans tous les principaux tissus, et ce projet a depuis été étendu au niveau international (209, 210, 211).

## **2.6. Biologie des systèmes et médecine des systèmes**

La biologie des systèmes et son domaine appliqué de la médecine systémique reposent sur le principe que les mesures de l'information biologique doivent être globales. Les différents niveaux d'information génomique doivent être étudiés de manière intégrée et les systèmes biologiques fonctionnent dynamiquement, non comme des événements discrets (212). La santé humaine et la maladie sont définies par des voies moléculaires interdépendantes qui peuvent être définies par le génome, l'épigénome, le transcriptome, le protéome et / ou le métabolome. Bien que chacun puisse être analysé individuellement, le cœur de la biologie des systèmes est la création d'un modèle qui intègre toutes les données d'expression et ses changements dynamiques. L'analyse des données de systèmes à partir d'un seul échantillon de sang ou de tissu fournit des informations sur la classification moléculaire et la prédiction pour un état pathologique particulier ainsi que sur le contexte biologique du développement de nouvelles cibles thérapeutiques (213, 214, 215, 216).

### **3. De l'information de génome à la médecine personnalisée**

La traduction d'informations génomiques en médecine clinique nécessite de répondre aux questions suivantes: Quels sont les facteurs moléculaires responsables de la pathogenèse d'une maladie particulière? Pouvons-nous déterminer quels patients développeront une maladie? Comment pouvons-nous développer un régime thérapeutique qui peut être adapté à la biologie unique d'un individu? Le tableau 4 donne des exemples de marqueurs génomiques utilisés cliniquement aujourd'hui dans le continuum de l'état normal à l'état maladie. Beaucoup de ces outils ont été traités dans les sections précédentes, et cette section mettra en évidence l'utilisation de ces outils dans la pratique médicale aujourd'hui.

#### **3.1. Tests préconceptionnels et dépistage des hétérozygotes**

En France, le dépistage préconceptionnel des couples à risque de maladies héréditaires monogéniques s'effectue dans le cadre d'un conseil génétique s'il existe un cas index dans la famille. Mais dans quelques populations présentant une forte prévalence de certaines affections héréditaires le dépistage des hétérozygotes est devenu depuis longtemps systématique :  $\beta$  thalassémies dans les pays méditerranéens et maladie de Tays-Sachs dans la communauté juive Ashkénase des États-Unis. À Chypre, le dépistage de la  $\beta$  thalassémie a fait baisser le nombre de cas de 51 en 1974 à simplement 5 entre 1991 et 2001 ; aucun cas n'a été observé entre 2002 et 2007. Aux États-Unis 1379 couples à risque de Tays-Sachs ont été identifiés (aboutissant soit à la non réalisation de l'union projetée, soit à un diagnostic prénatal, soit à un diagnostic préimplantatoire) et l'incidence de la maladie a diminué de 90 % ; dans cette communauté juive le programme de dépistage a été étendu à 16 autres maladies

récessives. En Israël, la détection systématique est recommandée pour 8 maladies dont la fréquence des hétérozygotes est supérieure à 1/60 et gratuite pour 4 d'entre elles (mucoviscidose, Tays-Sachs, dysautonomie familiale, thalassémie) ; elle est également encouragée pour des maladies dont la fréquence des hétérozygotes est supérieure à 1/110 (Nieman Pick, glycogénoses, etc.). Aux États-Unis le dépistage des hétérozygotes est recommandé pour des maladies fréquentes comme la mucoviscidose ou l'amyotrophie spinale ; la firme « Genzyme genetics » propose ainsi un test pour 98 des mutations les plus fréquentes du gène CFTR (un couple à risque sur 1 400).

En France, le dépistage néonatal systématique concerne cinq maladies : mucoviscidose, phénylcétonurie, hypothyroïdies, hyperplasies congénitales des surrénales et drépanocytose ; il s'agit d'un dépistage biologique des homozygotes. L'apparition des techniques de séquençage à haut débit est susceptible d'entraîner le passage à un dépistage neonatal génotypique, identifiant également de ce fait les hétérozygotes et éventuellement étendu à plusieurs centaines de maladies dont les gènes et les mutations pathogènes sont connus et ceci pour moins de 400 dollars (217). Les NGS rendent ainsi possible le diagnostic préconceptionnel généralisé, ce qui soulève de multiples questions tant éthiques que socio-économiques. Sur un tel marché potentiel se positionnent dès à présent des firmes commercialisant des DTC (Direct Test to Consumer). Ainsi « Amby Screen » propose le dépistage de 75 maladies pédiatriques à transmission récessive ou liée à l'X. « 23 and me » offre un test plus étendu incluant, outre le génotypage de mutations responsables de maladies monogéniques, la recherche de polymorphismes associés à des maladies dites plurifactorielles et un profil pharmacogénétique, et cela pour 99 dollars.

### 3.2. Diagnostic prénatal et préimplantatoire

Les diagnostics prénataux ne peuvent être pratiqués en France que dans un cadre réglementaire strict par des centres pluridisciplinaires et des laboratoires agréés ; dans le cas du diagnostic préimplantatoire seuls quatre centres sont autorisés par l'Agence de Biomédecine : Paris, Strasbourg, Montpellier et Nantes. Les indications diagnostiques sont restreintes puisque la loi n'autorise l'interruption de grossesse que pour une affection d'une particulière gravité et reconnue comme incurable. Le rapport 2010 de l'Agence fait état de 55 000 diagnostics prénataux en cytogénétique (essentiellement recherche de trisomie) dont près de 4 000 positifs et de 2 750 en génétique moléculaire dont 544 positifs. Le diagnostic prénatal est effectué sur biopsie de trophoblaste à 12 semaines d'aménorrhée (entraînant 1 à 2 % d'avortements), sur cellules amniotiques à 16 semaines (0,2 à 0,5 % d'avortements), sur sang foetal après 20 semaines (3 % d'avortements) ou sur tissu foetal au troisième trimestre de la grossesse. Un diagnostic préimplantatoire nécessite une stimulation ovarienne hormonale, la collecte d'ovules suivie d'insémination artificielle, une biopsie embryonnaire au troisième jour (J3), un diagnostic en 24-48 heures (caryotype avec hybridation *in situ* ou recherche d'une mutation) et un transfert *in utero* de 1 à 3 embryons sains à J4/J5 ; chacune de ces étapes comporte un pourcentage d'échecs important et en définitive le nombre de grossesses ainsi obtenu est faible (218).

Les pertes foetales liées aux prélèvements invasifs pratiqués pour le diagnostic prénatal sont estimées entre 300 et 600 par an, d'où l'intérêt majeur d'un diagnostic non invasif par analyse des acides nucléiques foetaux circulant dans le sang maternel. La présence d'ADN foetal ou cffDNA (cell free fetal

DNA) dans le sang maternel a été démontrée en 1997 et sa concentration estimée à plus de 10 % de l'ADN total circulant ; il disparaît rapidement du sang maternel après l'accouchement, contrairement aux cellules mononuclées fœtales qui pourraient persister plusieurs années. Détectable dès 5 à 6 semaines d'aménorrhée, il a une origine trophoblastique ou placentaire, et provient sans doute de cellules en apoptose ; il est *Bull. Acad. Natle Méd.*, 2014, 198, no 1, 101-117, séance du 7 janvier 2014 106 constitué majoritairement de petits fragments dont la taille moyenne est de 143 bases, ce qui correspond à la longueur d'ADN enroulé sur un nucléosome. La première application de l'analyse de l'ADN fœtal du sang maternel a été la détermination du sexe du fœtus et potentiellement des maladies liées à l'X : il s'agit d'un examen inscrit à la nomenclature des actes de biologie (B500). En 2011 la HAS a également donné un avis favorable à la prise en charge de la détermination du rhésus fœtal chez les femmes rhésus négatif : actuellement environ 6000 examens par an, mais il s'agit d'un « marché » potentiel de 160 à 180 000 tests annuels. En utilisant des « primers » spécifiques il était également possible de faire différents diagnostics ciblés ( $\beta$  thalassémie, hémophilie A...). L'apparition des techniques de séquençage à haut débit, en séquençant à la fois l'ADN circulant maternel et l'ADN fœtal et en déterminant leurs proportions relatives (par détermination des SNPs pour reconstituer les haplotypes), a ouvert de nouvelles perspectives : diagnostic des maladies génétiques et surtout des aneuploïdies (219, 220). La recherche d'une trisomie repose sur la mise en évidence d'un petit excès de matériel provenant de ce chromosome par rapport à l'ADN maternel ; il implique un séquençage aléatoire des petits fragments d'ADN avec une grande profondeur suivi d'une analyse bioinformatique complexe. Les premiers résultats récemment publiés sont très bons pour la trisomie 21 (sensibilité 100 %

et 0,1 % de faux positifs) et la trisomie 18 (sensibilité 100 % et 0,3 % de faux positifs), sensiblement moins performants en cas de trisomie 13 (sensibilité de 91,4 % et 0,9 % de faux positifs). Bien entendu les possibilités offertes par les NGS dans ce domaine vont se multiplier ; elles devraient pouvoir également s'appliquer dans l'avenir au diagnostic préimplantatoire, mais aussi à la détection des petites délétions et des variations du nombre de copies ou CNV (« Copy Number Variation »), restreignant ainsi le champ de la cytogénétique classique (221).

### **3.3. Prédiction du risque**

L'information génomique change le paysage médical en fournissant des outils qui transforment la pratique en prévision et en prévention du risque. Par exemple, les femmes porteuses de mutations dans les gènes BRCA1 et BRCA2 à forte pénétrance présentent un risque accru de susceptibilité au cancer du sein (222). Le dépistage clinique de ces mutations génétiques a conduit à des mesures préventives standard telles que des interventions chirurgicales ou radio thérapeutiques pour aider les patients à minimiser leur risque (223). Dans le cancer colorectal, les méthodes cliniques pour détecter l'instabilité des microsatellites dans les gènes de réparation des mésappariements MLH1 et MSH2 identifient les patients à risque et permettent aux médecins de les surveiller de près grâce au dépistage fréquent du cancer du côlon (224). Des facteurs de risque génétiques cliniquement utilisables ont également émergé pour le syndrome du QT long (LQTS) et le syndrome du QT court (225). LQTS est hérité selon un schéma autosomique dominant, et la variation du phénotype de la maladie est associée à des mutations dans 12 gènes différents de susceptibilité LQTS, ce qui justifie de tester génétiquement les patients pour ces

mutations (226). Par exemple, les  $\beta$ -bloquants sont un traitement efficace pour les patients avec LQTS1 (KCNQ1) mais pas pour ceux avec LQTS2 (KCNH2) ou LQTS3 (SCN5A), démontrant comment le génotypage de ces patients peut aider à guider le traitement (227). L'utilisation de tests de variation génétique utilisant le génotypage et le séquençage combinés avec des stratégies interventionnelles peut retarder la progression de la maladie ou prévenir complètement la maladie.

### **3.4. Diagnostic de la maladie et caractérisation moléculaire**

Une application précoce de la technologie du génome à la médecine clinique est dans la classification des états pathologiques. Les analyses génomiques et moléculaires ont révélé des sous-types distincts de la maladie, qui ont été traditionnellement définis par de larges phénotypes cliniques ou descriptifs. Parce que la réponse thérapeutique peut différer entre chaque sous-type, cela a des implications claires pour guider le traitement clinique.

L'oncologie offre de multiples exemples de la façon dont la médecine génomique a changé la compréhension de la maladie. Plusieurs génomes de cancer ont été séquencés, y compris la leucémie myéloïde aiguë, le cancer du poumon non à petites cellules (NSCLC) et le cancer du poumon à petites cellules (228, 229, 230). L'Atlas du génome du cancer, un effort multi-institutionnel de 1,5 milliard de dollars sur dix ans, vise à caractériser et à réaliser une analyse génomique multidimensionnelle du cancer (231). L'un des premiers cancers étudiés était le glioblastome multiforme (GBM); 206 échantillons ont été analysés pour le nombre de copies d'ADN, l'expression des gènes et les états de méthylation de l'ADN. De nouvelles mutations fréquentes du gène de la sous-unité régulatrice de la kinase PI3 ont été trouvées ainsi qu'une

association entre l'état de méthylation de MGMT (biomarqueur pour la réponse au témozolomide) et les mutations de réparation de mésappariement dans le GBM post-traitement (232). Une autre conclusion importante à tirer de ce projet est la caractérisation de sous-types distincts de GBM (classique, mésoenchymateuse, pro neurale) basés sur des mutations génétiques causales dans EGFR, NF1 et PDGFR / IDH1, respectivement (233).

La capacité à trouver plusieurs profils d'expression génique discrets dans un type de tumeur reflète l'hétérogénéité de la maladie. L'un des premiers exemples d'utilisation de profils de micro réseaux d'expression génique pour classer les tumeurs a été utilisé pour distinguer la leucémie myéloïde aiguë de la leucémie lymphoblastique aiguë (234). Des méthodes similaires ont également été appliquées avec succès pour caractériser, au niveau moléculaire, une large gamme de cancers du sein, de cancers du poumon et de lymphomes (235, 236, 237).

L'utilisation de la génomique pour stratifier la maladie n'est pas propre à l'oncologie. Les syndromes coronariens aigus (SCA) décrit une classe de maladies impliquant l'occlusion des artères coronaires, entraînant une lésion ischémique du cœur. Plusieurs études ont démontré l'utilité de la métabolomique et de la protéomique pour distinguer les états SCA de l'infarctus aigu du myocarde et de l'angor instable (238, 239). Une étude récente du protéome plaquettaire chez des patients avec et sans sus-décalage du segment ST a mis en évidence des différences d'activation plaquettaire et, en particulier, des taux élevés de la protéine SPARC (240). L'expression du gène du sang périphérique a également été associée à la sévérité de la coronaropathie, et un test clinique diagnostique a été développé sur la base d'observations similaires (CorusTM



CAD) et validé dans un essai prospectif multicentrique avec des résultats impressionnants, améliorant la précision de l'évaluation des maladies coronariennes de 16% -20% (241, 242). Ainsi, un certain nombre de plateformes de tests génomiques complexes sont utilisées dans la clinique ou sont préparées pour une utilisation clinique afin d'aider au diagnostic et d'éclairer la prise de décision thérapeutique.

**Tableau 4 : Diagnostics moléculaires disponibles cliniquement sur le continuum de la santé à la maladie dans le cancer et les maladies cardiovasculaires (référence 105)**

**Table 3 Clinically available molecular diagnostics along the continuum from health to disease in cancer and cardiovascular disease**

| Time point in clinical decision making | Cancer                                                                                                |                                                                                                 | Cardiovascular disease                                                            |                                                        |
|----------------------------------------|-------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------|
|                                        | Test                                                                                                  | Indication                                                                                      | Test                                                                              | Indication                                             |
| Risk/susceptibility                    | <i>BRCA1, BRCA2</i><br><i>HNPCC, MLH1, MSH2</i><br><i>TP53, PTEN</i>                                  | Breast<br>Colon<br>Sarcomas                                                                     | <i>KIF6, 9p21</i><br>Familion® 5-gene profile                                     | CAD<br>LQTS                                            |
| Screening                              | HPV genotypes                                                                                         | Cervical                                                                                        | Corus™ CAD                                                                        | CAD                                                    |
| Diagnosis                              | Lymphochip                                                                                            | Lymphoma                                                                                        | Corus CAD                                                                         | CAD                                                    |
| Prognosis                              | Oncotype DX® (21-gene assay)<br>MammaPrint® (70-gene assay)<br>Her2/neu, ER, PR                       | Breast                                                                                          | TnI, BNP, CRP                                                                     | ACS                                                    |
| Pharmacogenomics                       | Her2/neu<br><i>UGT1A1</i><br><i>KRAS</i><br><i>EGFR</i><br>Amplichip®; DMET™<br><i>CYP2D6/CYP2C19</i> | Herceptin<br>Irinotecan<br>Cetuximab<br>Erlotinib, gefitinib<br>Various others<br>(see Table 2) | <i>KIF6, SLCO1B1</i><br>Amplichip; DMET<br><i>CYP2D6/CYP2C19</i><br><i>VKORC1</i> | Statins<br>Warfarin<br>Various others<br>(see Table 2) |
| Monitoring                             | CTCs                                                                                                  | Tumor recurrence or progression                                                                 | AlloMap® gene profile                                                             | Transplant rejection                                   |

Abbreviations: ACS, acute coronary syndromes; BNP, brain natriuretic peptide; CAD, coronary artery disease; CRP, C-reactive protein; CTC, circulating tumor cell; ER, estrogen receptor; HPV, human papillomavirus; LQTS, long QT syndrome; PR, progesterone receptor. Table adapted from Reference 64.

### 3.5. Pronostic de la maladie

Les modèles de signatures tumeur-expression génique, combinés à des données cliniquement pertinentes telles que les issues de survie, fournissent l'un des exemples les plus clairs de l'utilisation clinique de l'information génomique

en tant qu'outils pronostiques (243, 244, 245, 246). MammaPrint R\_ et Oncotype DXR\_ sont deux de ces tests pronostiques qui sont maintenant cliniquement disponibles pour une utilisation dans le cancer. L'oncotype DX est une signature de 21 gènes dérivée d'une réaction en chaîne de la polymérase en temps réel d'échantillons inclus dans la paraffine et utilisée pour prédire la récurrence à distance sur une période de 10 ans (247, 248, 249). Il est actuellement en cours de validation dans un essai clinique de phase III, mais a une plus grande signification statistique par rapport aux facteurs de prédiction histologiques traditionnels (250, 251). MammaPrint, approuvé par la Food and Drug Administration américaine en 2007, a été développé au Netherlands Cancer Institute, qui a créé une signature à 70 gènes pour discriminer les patients à haut et faible risque de récurrence métastatique dans les cinq ans (252). Des essais cliniques basés sur cette signature ont depuis été lancés pour valider la signature, et les auteurs de ces études ont rapporté une prédiction précise (253, 254). Une méta-analyse a récemment montré que les informations pronostiques d'Oncotype DX jouent un rôle important dans la prise de décision clinique et ont conduit à une réduction de la chimiothérapie tout en augmentant les thérapies hormonales adjuvantes (255).

Les approches génomiques ont également contribué au pronostic après transplantation rénale. Une étude a été menée chez des receveurs de greffe de rein qui n'ont pas rejeté leurs greffes rénales malgré l'arrêt du traitement immunosuppresseur pour analyser les profils d'expression génique des globules blancs circulants. Cela a conduit à la découverte d'une expression génique spécifique dans les cellules B qui distinguait les patients tolérants des patients non tolérants (256). Les résultats de ceci impliquent que des biomarqueurs

peuvent être développés qui peuvent prédire quels patients devraient continuer à recevoir l'immunosuppression des années après la chirurgie de transplantation et quels patients peuvent avoir ceci discontinué.

### **3.6. Pharmacogénomique**

La pharmacogénomique est l'utilisation de l'information génomique pour déterminer ou prédire la réponse à la thérapie. En oncologie, le génome somatique de la tumeur a été utilisé pour identifier les biomarqueurs qui prédisent et guident la thérapie pour une réponse optimale. Par exemple, dans la leucémie lymphoblastique aiguë, environ 20% des patients deviendront réfractaires au traitement. Deux études utilisant l'expression génique différentielle ont montré que les gènes impliqués dans la chimiorésistance étaient également associés à un pronostic plus sombre, jetant les bases de futures études sur les mécanismes de résistance (257, 258).

La thérapie ciblée est souvent synonyme de pharmacogénomique. Le gène récepteur du facteur de croissance, le récepteur du facteur de croissance épithélial humain (HER2), s'est révélé amplifié dans 25% à 30% des cancers du sein, et sa surexpression corrélée avec un mauvais pronostic (259,260). Le Trastuzumab, un anticorps monoclonal qui cible HER2, est efficace pour réduire la charge tumorale, et le dépistage des patients pour l'expression de HER2 / neu est devenu une routine et une exigence pour l'utilisation du médicament (261, 262, 263). Dans le CBNPC, les effets antitumoraux ont été démontrés en utilisant des inhibiteurs du récepteur du facteur de croissance épidermique (EGFR), un récepteur de surface cellulaire fréquemment muté pour provoquer la prolifération du cancer (264, 265, 266). Bien que l'essai Iressa de survie au cancer du poumon - un essai de phase III randomisé et contrôlé par placebo - ait

montré que le gefitinib, un inhibiteur de l'EGFR, ne procurait aucun avantage de survie à la population générale, un sous-groupe de non-fumeurs asiatiques mettre en évidence le potentiel de la génomique pour diriger et focaliser la thérapie (267). Parce que seulement 5% à 15% des patients blancs portent une mutation somatique EGFR, comparativement à 25% -35% des patients asiatiques atteints de CPNPC, ces études soulignent le besoin d'un dépistage génétique supplémentaire avant d'administrer un traitement (268).

Un exemple important de pharmacogénétique concerne la prise en charge de la warfarine (269). En raison de l'index thérapeutique étroit du médicament, diverses complications sont associées à son traitement, même après ajustement de la dose en fonction de l'âge, du sexe, du poids, de l'état pathologique, du régime alimentaire et des médicaments concomitants. L'étude des propriétés pharmacocinétiques de la warfarine a indiqué l'implication de deux gènes dans la détermination de la dose de warfarine. L'un de ces gènes code le CYP2C9, responsable de la majeure partie de la clairance métabolique (~ 80%) de l'énantiomère S, plus puissant sur le plan pharmacologique, de la warfarine. Les deux CYP2C9 \* 2 et \* 3 provoquent une réduction de la clairance de la S-warfarine, avec une variation de 10 fois du génotype liée à l'activité la plus élevée (CYP2C9 \* 1 / \* 1) à la plus faible (CYP2C9 \* 3 / \* 3). De nombreuses études ont associé ces génotypes à une sensibilité initiale à la dose, à une stabilisation retardée de la dose d'entretien, à des retards de sortie de l'hôpital et à des complications hémorragiques accrues (270). Le second gène identifié comme prédicteur du dosage est la protéine complexe de la vitamine K époxyde réductase 1 (VKORC1), ciblée par la warfarine. La prise en compte du génotype VKORC1 ou de l'haplotype associé au génotype CYP2C9 et des facteurs tels

que l'âge et la taille corporelle sont estimés représenter 35% à 60% de la variabilité des besoins posologiques en warfarine (271).

Malgré le fait que plusieurs groupes indépendants ont reproduit ces données, des études cliniques prospectives sont nécessaires pour établir si la dose initiale peut être adaptée aux patients par le génotypage CYP2C9 et VKORC1, couplée avec des variables cliniques connues (272). Bien que ces études soient actuellement en cours, le comité consultatif de pharmacologie clinique de la FDA a reconnu l'importance et le potentiel du génotypage du CYP2C9 et du VKORC1 durant la phase précoce du traitement par warfarine, et l'étiquette du médicament a été modifiée en conséquence en août 2007.

### **3.7. Surveillance de la réponse des maladies à la thérapie**

Le profil d'expression génique des cellules mononuclées du sang périphérique (PBMC) est couramment utilisé dans certains centres pour surveiller l'état des greffes après transplantation d'organes solides, l'essentiel étant de développer de meilleurs tests pour évaluer le rejet du greffon (273, 274,275). La recherche en transplantation cardiaque, peut-être plus que tout autre organe, a ouvert la voie à la mise au point de nouveaux tests sanguins permettant de faire la distinction entre la quiescence et le rejet aigu. L'étude d'observation génique de rejet d'allo greffon cardiaque (CARGO), une étude multicentrique phare en trois étapes dans le domaine de la transplantation cardiaque, a montré que le profil d'expression génique des PBMC pouvait être utilisé pour distinguer avec précision les patients post-transplantation versus rejet aigu (276). Ceci a été réalisé en identifiant d'abord un ensemble de 11 gènes qui distinguaient le mieux entre le rejet aigu et la quiescence, puis en utilisant une cohorte indépendante pour valider le schéma de classification dans

un mode prospectif et aveugle. En utilisant un système de notation de 0 à 40, avec de faibles scores représentant un faible risque de rejet de greffe, les chercheurs ont utilisé un seuil de 30 sur les patients plus d'un an après la transplantation et la validation du test a donné une valeur prédictive négative de 99,6% et une valeur prédictive positive de 6,8% pour le rejet aigu. Ce test est disponible dans le commerce sous le nom AlloMap R<sub>+</sub> (XDx; Brisbane, CA) et coûte environ 3 000 \$ par test. Bien que coûteuse, l'analyse des coûts basée sur les données de la cohorte CARGO a montré que ce test de profil d'expression génique est moins cher que la biopsie et permettrait d'économiser jusqu'à 12 millions de dollars en coûts de santé par an (277). Pour que cela soit vrai, cependant, le besoin de biopsie devrait être évité, et le profil d'expression génique devrait suffire chez les patients considérés comme présentant un faible risque de rejet. Un essai prospectif, multicentrique, randomisé mais non aveugle appelé Atténuation Invasive de Surveillance par Expression Génique (IMAGE) a montré qu'une stratégie de surveillance non invasive utilisant le profil d'expression génique n'est pas inférieure à la biopsie endomyocardique pour diagnostiquer un dysfonctionnement d'allogreffe cardiaque dans le rejet. Les auteurs ont conclu que l'utilisation de cette stratégie de surveillance du rejet réduit significativement le nombre de biopsies sans risque accru d'événements indésirables, faisant de l'étude un exemple clé où le profil d'expression génique des PBMC peut être utilisé pour guider la prise de décision thérapeutique (278).

#### **4. Génomes interactives et microbiomes**

Bien que la génomique joue un rôle important dans le diagnostic, le pronostic, le traitement et la prévention de la maladie, il reste beaucoup à savoir sur l'interaction humaine avec l'environnement. Des études récentes de la

réponse de l'hôte aux pathogènes microbiologiques ont démontré des changements d'expression génique très spécifiques en réponse à différents virus; par exemple, une signature d'expression génique basée sur l'hôte a été récemment identifiée qui pourrait être utilisée pour la détection précoce d'une infection virale (279,280). Les mêmes données ont également été utilisées pour montrer que les profils d'expression génique distinguent les infections bactériennes et virales (Figure 6). Les profils d'expression des gènes dans les cellules hôtes ont également été modifiés en fonction d'autres facteurs de stress environnementaux tels que le tabagisme et les états pathologiques comme l'asthme et la bronchopneumopathie chronique obstructive, et pourraient permettre de comprendre la pathogenèse et la susceptibilité au risque (281, 282, 283). Ainsi, les interactions gène-environnement mesurées en utilisant la réponse du génome de l'hôte commencent à fournir de nouveaux biomarqueurs cliniquement utiles et potentiellement diagnostiques.

#### **4.1. Génomes pathogènes**

Les génomes des pathogènes humains sont également séquencés dans l'espoir de fournir des informations pertinentes sur la pathogenèse de l'infection ou des cibles thérapeutiques potentielles. En 1995, le premier génome bactérien achevé a été séquencé par *Haemophilus influenzae* (284). Depuis lors, plus de 1 000 génomes bactériens et plus de 3 000 génomes viraux ont été complètement séquencés (285). Cette information génomique a fourni des données précieuses pour les outils génomiques comparatifs, en comparant les souches pathogènes avec les souches non pathogènes pour identifier les protéines qui sont utilisées pour la pathogenèse, conduisant potentiellement à de nouveaux médicaments avec une toxicité limitée (286). L'analyse protéomique comparative entre les

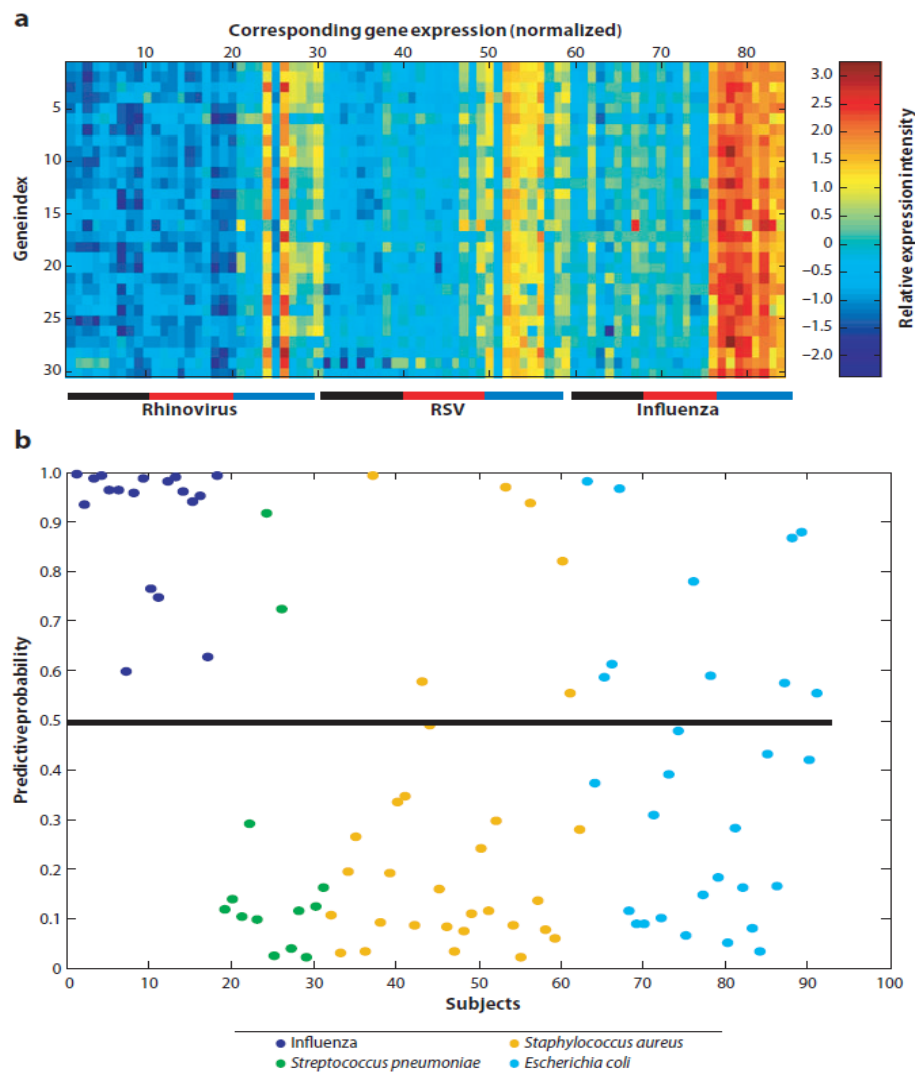
phénotypes virulents et avirulents de *Rickettsia Prowazekii* a révélé des mutations adaptatives dans huit gènes responsables de sa virulence (287). Le séquençage de génomes bactériens pour des cibles vaccinales telles que les protéines de surface bactériennes a également connu un grand succès clinique: Des séquençages vaccnologiques inversés des protéines cibles ont conduit au développement d'un vaccin prometteur pour le séro groupe B de *Neisseria meningitidis* récemment introduit dans les essais cliniques (288). Une autre approche pour comprendre comment les pathogènes développent une résistance a été de suivre les changements génomiques des isolats de *Staphylococcus aureus* au cours du temps chez un patient après l'exposition au traitement antibiotique jusqu'à la résistance, une stratégie qui a identifié 35 mutations dans 31 loci (289).

#### **4.2. Le microbiome humain**

Les interactions synergiques entre les humains et notre flore corporelle normale doivent être comprises, car il y a 10 fois plus de bactéries que le nombre de cellules humaines dans le corps: Le côlon humain contient à lui seul plus de 400 espèces bactériennes comprenant environ 10<sup>13</sup>-10<sup>14</sup> microorganismes (290,291). À travers l'étude de la variation génétique, le microbiome décrit l'écologie fragile qui existe entre la communauté des microbes symbiotiques, commensaux et pathologiques et notre espace corporel partagé. Potentiellement, cette connaissance pourrait être utilisée pour fournir des tests de diagnostic rapides, prévoir l'efficacité des antibiotiques et surveiller les environnements pour les contaminants (292, 293). La microbiomique a été utilisée pour identifier des infections virales inconnues ou pour diagnostiquer la résistance aux antibiotiques dans des infections telles que *Staphylococcus aureus*



résistant à la méthicilline (294, 295). De plus, plusieurs maladies ont été associées à des déséquilibres à grande échelle dans le microbiome intestinal, y compris la maladie inflammatoire de l'intestin, la diarrhée résistante aux antibiotiques et l'obésité (296, 297). Le Projet sur le microbiome humain est conçu pour comprendre ces interactions sur le plan physiologique et pathologique, en créant non seulement une base de données contenant des pathogènes, mais aussi un ensemble de contrôle de la flore normale (298). Au fur et à mesure que notre conscience de l'équilibre délicat entre l'hôte et les micro-organismes continue de croître, les applications du microbiome vont continuer de croître.



**Figure 2**

Gene expression signatures diagnose influenza and other respiratory viral infections in humans. (a) A microarray-based gene expression signature can accurately differentiate between symptomatic and asymptomatic individuals following exposure to rhinovirus, respiratory syncytial virus (RSV), and influenza. In the key, black represents the preinoculation baseline (healthy), red represents symptomatic postinoculation, and blue represents asymptomatic postinoculation. (b) A microarray-based gene expression signature can also distinguish individuals with bacterial infection from those with viral infection. The region above the black line is where we predict a high probability of viral infection; the region below is where the probability is low. Figure adapted from Reference 241.

**Figure 6 : Les signatures d'expression génique permettent de diagnostiquer la grippe et d'autres infections virales respiratoires chez les humains. (Référence 280)**

## **5. Intégration de la génomique dans la pratique clinique**

Plus l'information génomique est utilisée pour amplifier nos connaissances sur le mécanisme et le traitement de la maladie, mieux elle expose où les lacunes résident dans la méthodologie, l'infrastructure et la politique pour sa mise en œuvre en médecine clinique. L'abondance de la recherche utilisant des outils génomiques et la prospérité des données dont nous disposons nous obligent à faire des analyses techniques de ces résultats pour suivre leur production. Plusieurs suggestions ont été faites pour créer un cadre pour cette analyse, qui devrait impliquer ce qui suit pour chaque test: capacités techniques et normes opérationnelles, capacités de diagnostic et impact sur la prise de décision clinique, méthodes d'intégration comportementale et cognitive des médecins, rapport coût-efficacité, et les résultats pour la santé (299).

Une méta-analyse récente a révélé des zones de déficience dans le parcours de la clinique au chevet du patient, en commençant par la main-d'œuvre en soins primaires qui est à l'avant-garde de l'interaction avec le patient (300). Les médecins ne se sentent souvent pas préparés à administrer la médecine génomique, et les patients eux-mêmes ne comprennent pas clairement la valeur et l'intégration des tests génomiques dans leurs soins médicaux. Ce manque d'éducation s'accompagne d'un besoin accru d'évaluations des résultats pour les interventions génomiques, conduisant à une intégration limitée de la médecine génomique dans la pratique réelle. Avec ces obstacles identifiés, comment devrions-nous procéder?

## **5.1. Education**

L'un des principaux obstacles à la traduction de l'information génomique en pratique clinique est le manque de compréhension de la génomique par le fournisseur et l'accès limité aux outils appropriés. Plusieurs enquêtes suggèrent une déficience croissante des connaissances en génétique chez les médecins en moyenne (301, 302, 303). En plus de se familiariser avec les progrès rapides de la recherche génomique, les fournisseurs doivent également comprendre comment interpréter les résultats. Un baromètre direct de ceci est trouvé avec l'histoire de famille: Malgré le fait que 87% des médecins reconnaissent son importance, les médecins de premier recours se sont efforcés d'interpréter les schémas familiaux de la maladie et de mettre en œuvre l'information sur les antécédents familiaux dans leurs processus de travail (304, 305). Il est impératif d'améliorer la formation médicale entourant la médecine génétique et génomique pour le domaine et son intégration dans la pratique clinique pour avancer (306).

## **5.2. Système de soutien à la décision clinique**

SSD doit être utilisé pour promouvoir et intégrer les outils génomiques dans le flux de travail clinique. Dans les essais contrôlés randomisés, il a été démontré que les interventions de la SSD amélioraient significativement les soins prodigués aux patients et pouvaient avoir des implications significatives pour l'administration de recommandations basées sur la génomique aux médecins occupés (307, 308, 309). Si elle est élargie pour inclure les prévisions de risque pour la santé basées sur l'information génomique, des recommandations individualisées et en temps réel sur la pharmacogénomique ou les profils de risque pourraient être fournies sur la base d'une recherche à jour.

Le développement d'une infrastructure centrale et nationale pour la SSD et d'autres ressources génomiques, avec accès à l'information génomique et non génomique des patients, est essentiel pour toute mise en œuvre future (309). L'accès des médecins aux technologies de la SSD sera important pour son utilisation cohérente. De plus, ces systèmes ont eux-mêmes besoin de tests fondés sur des données probantes pour s'assurer qu'ils améliorent la prestation des soins de santé.

### **5.3. Vie privée**

À mesure que l'information génétique croît dans le contexte de la prise de décision clinique, les patients s'inquiètent progressivement de la discrimination génétique (310, 311). Au niveau fédéral, en 2005, le Congrès américain a adopté la Loi sur la transférabilité et la responsabilité en matière d'assurance maladie, qui prévoyait une protection contre la discrimination génétique pour les régimes de santé collectifs. La loi de 2008 sur la non-discrimination génétique a été adoptée pour fournir une protection supplémentaire (312). Plus de 30 états ont fourni leur propre législation traitant de cette question, créant une mosaïque de protection contre l'utilisation de l'information génétique par les assureurs pour discriminer les patients (313). Cependant, il faudra également élargir l'accès à l'éducation pour les patients, y compris des informations sur les composants génétiques spécifiques de leurs maladies et sur la façon dont ils affectent les droits au traitement et à la vie privée. Élargir le rôle des conseillers en génétique dans l'éducation des patients comblerait ce besoin et allégerait grandement le fardeau du système de soins de santé.

#### **5.4. Politique réglementaire**

La réglementation des produits pharmaceutiques à base de génomique et le niveau de rigueur auquel ils sont testés constituent un autre domaine de politique nécessaire (314). Ici, la FDA a été proactive, comme l'illustrent les plus de 200 étiquettes de produits qui pointent vers l'influence de la variation génétique sur la réponse aux médicaments (315). Ses efforts récents pour influencer les tests génétiques directement auprès des consommateurs, qui permettent aux consommateurs d'accéder à la variation génétique au sein de leur génome se rapportant à la maladie, sont peut-être encore plus remarquables (316). Bien que certains aient fait valoir que la validation adéquate fait défaut, il est clair qu'il existe un marché pour de telles entreprises (317). Cependant, la valeur clinique de leurs résultats n'a pas été définie, et la question à laquelle il faut répondre est, Quelle est la valeur de ces tests pour fournir aux cliniciens des données utilisables? Pour répondre à ces préoccupations, en juin 2010, la FDA a annoncé qu'elle commencerait à réglementer les tests directs aux consommateurs en tant qu'instrumentation médicale (318). Reste à savoir si ces mesures vont freiner le progrès scientifique ou protéger les consommateurs de la confusion par des informations cliniquement non pertinentes.

#### **5.5. Normes**

La réglementation croissante des outils génomiques dans la recherche et la clinique souligne le besoin de normes nationales de pratique pour toutes les infrastructures génomiques, des approches statistiques à la conception d'études cliniques et même à une infrastructure de bio banque qui fournit aux chercheurs des données provenant de multiples études. Les référentiels centralisés et normalisés pour le stockage, l'enregistrement et l'annotation d'échantillons

biologiques sont parmi les ressources les plus favorables à la médecine génomique (319). Des lignes directrices sur les pratiques exemplaires ont été publiées depuis 2005 pour orienter les efforts de bio banques, et des efforts ont déjà été déployés au niveau national pour normaliser la collecte d'échantillons biologiques: Le plan directeur du Réseau national de prélèvements biologiques a été proposé en 2005 et piloté par le Programme pilote de réseau national de bio spécificités du Programme spécialisé de recherche sur la prostate de l'Institut National du Cancer pour stocker des échantillons biologiques dans de grandes bio banques en réseau (320, 321). Les questions d'accessibilité devraient également être abordées, en accordant aux chercheurs extérieurs un accès gratuit à l'information tout en protégeant la propriété originale de l'échantillon. L'éthique de l'utilisation d'un large consentement du patient pour stocker les échantillons et les utiliser est toujours discutée, mais tous les efforts doivent être faits pour protéger la vie privée des patients et séparer les données sur l'identité des patients des données de chaque échantillon (322).

De plus, des normes communautaires sont nécessaires dans les approches statistiques et mathématiques des données génomiques. En raison du manque de grandes populations, des études statistiquement puissantes pour valider les résultats de la recherche sur les maladies chroniques sont difficiles à concevoir. Tous les efforts devraient être faits pour tester des modèles basés sur la génomique dans des essais cliniques, conçus avec des contrôles appropriés, avant que ces modèles soient introduits dans la clinique. La reproductibilité des résultats doit être démontrée avant leur mise en œuvre clinique (323). La randomisation et l'aveuglement sont également essentiels pour éliminer le biais des observateurs. Étant donné que le risque de faux positifs est élevé, des normes de contrôle et de normalisation doivent être en place pour chaque étude respective afin de pouvoir comparer les résultats.

## **5.6. Recherche sur l'efficacité comparative**

Étant donné le paysage actuel de la pratique fondée sur des données probantes, à quelle norme les approches génomiques de la médecine devraient-elles être appliquées? Pour résoudre ces problèmes, la recherche sur l'efficacité comparative (REC) est devenue une stratégie pour évaluer systématiquement les données à mesure que les résultats de la recherche entrent dans la pratique clinique. En 2009, 1,1 milliard de dollars de l'American Recovery and Reinvestment Act ont été affectés à la REC.

La REC est défini par l'Institute of Medicine comme «la génération et la synthèse de preuves qui comparent les avantages et les inconvénients des méthodes alternatives pour prévenir, diagnostiquer, traiter et surveiller un état clinique ou pour améliorer la prestation des soins, à la fois au niveau individuel et au niveau de la population »(324). En substance, la REC comprend une comparaison directe entre deux approches de soins et leur efficacité dans des contextes réels (325). Bien que certains ont fait valoir que la REC pourrait être utilisé à mauvais escient pour entraver l'avancement médical, la mise en œuvre de la médecine génomique exigera des preuves que la REC peut fournir, ce qui, nous l'espérons, encouragera l'innovation et réduira les coûts des soins de santé (326, 327). En comparant l'efficacité des outils génomiques dans le diagnostic, le pronostic, le traitement et la prédiction du risque par rapport aux normes d'or dans les soins actuels, des preuves peuvent être générées pour démontrer ou nier l'efficacité de plusieurs approches.

## **5.7. Propriété intellectuelle**

Le développement rapide des stratégies basées sur la génomique nécessite un investissement important et une coopération entre les intérêts publics et



privés. Malheureusement, au fur et à mesure que cette recherche se poursuit, l'insécurité entoure la propriété intellectuelle. Une question qui a retenu l'attention des juristes est la capacité des organisations, des entreprises et même des gouvernements à détenir des brevets sur des gènes nouvellement découverts. En 2009, un procès a été intenté au nom de l'American Civil Liberties Union et de la Public Patent Foundation pour invalider les brevets détenus sur les gènes BRCA1 et BRCA2, leur principal argument étant que ces gènes sont des produits de la nature (328). Les plaignants ont soutenu que les brevets (détenus par Myriad Genetics, l'Université de l'Utah Research Corporation et le gouvernement des États-Unis) entravaient la recherche. Fin mars 2010, un juge fédéral a révoqué des brevets sur les gènes BRCA1 et BRCA2 et potentiellement invalidé des centaines d'autres, confirmant qu'ils étaient des produits de la nature et non, comme le soutenaient les compagnies, de l'ADN isolé humain (329). Un appel a été déposé et ces contestations juridiques posent des questions intéressantes sur la façon dont la recherche en génomique devrait se dérouler.

## **5.8. Remboursement**

Si la médecine personnalisée doit être largement pratiquée, il est essentiel que les assureurs de santé, privés et gouvernementaux, disposés à fournir une couverture pour les tests génétiques, se dirigent vers l'intégration (330). À ce stade, les assureurs-maladie ont refusé de le faire parce qu'ils n'avaient pas reçu d'incitatifs qui les aideraient à obtenir des avantages à long terme en matière de coûts. Cependant, des progrès récents ont été réalisés pour encourager la médecine personnalisée: De nouveaux modèles ont montré une méthode de coordination de la prévention avec des résultats de maladie favorables et des coûts réduits (331). Comme les preuves renforcent l'importance de prévenir la

maladie en plus de la traiter, un nouveau rôle pourrait émerger pour la médecine génomique qui pourrait être soutenu par les compagnies d'assurance. Le gouvernement fédéral fournit une motivation supplémentaire, qui a joué un rôle accru dans l'introduction d'initiatives législatives visant à encourager l'adoption de la médecine personnalisée (332).

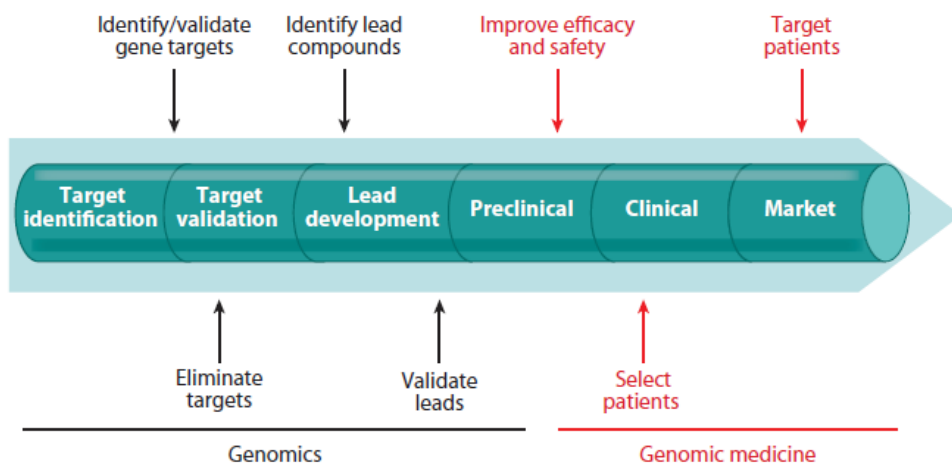
## **6. Quand la "médecine personnalisée" est "médecine" ?**

Il semble que les principaux obstacles se situent dans trois catégories: la collecte d'informations adéquates et appropriées sur les patients, le dosage de ces informations pour qu'elles deviennent cliniquement pertinentes et l'interprétation de ces tests par les professionnels de la santé en charge des soins aux patients.

Malgré les défis qui nous attendent pour l'intégration complète de la médecine personnalisée dans la prise de décision clinique, une revue du paysage de la pratique médicale suggère que la médecine personnalisée est arrivée. Bien qu'il soit facile de se demander pourquoi les progrès scientifiques de la recherche génomique ne se sont pas exactement traduits par les résultats révolutionnaires attendus en médecine, il serait stupide de négliger l'impact profond qu'elle a déjà fait. De plus, l'effet subjectif perçu peut être attribué à des technologies en retard et à des approches analytiques combinées à une lutte pour mettre en œuvre des résultats éprouvés en clinique. Bien que des améliorations puissent être apportées à la base de connaissances et à l'infrastructure, il est également important de garder une perspective historique: Les progrès de la médecine fondés sur une pensée scientifique novatrice prennent beaucoup de temps et d'investissements. Un rappel important de ce fait est l'histoire de 30 ans du paclitaxel, qui a débuté en 1962 avec la collecte d'écorce d'arbre et qui n'est entrée dans les essais cliniques de phase I qu'en 1984 (333). En moyenne, il faut

17 ans pour que 14% de la recherche scientifique soit incorporée dans la pratique clinique quotidienne (334, 335). L'utilisation de la génomique pour comprendre la biologie de la maladie et sélectionner des cibles pour le développement de médicaments et l'utilisation de la médecine génomique pour concevoir des essais cliniques et sélectionner des patients pour des thérapies ciblées devraient augmenter le taux de réussite et réduire le temps de développement de médicaments (Figure7).

Comme Kuhn l'affirme clairement, les changements dans la pensée scientifique ne peuvent se produire que par ce qui doit sembler être des exemples disjoints, chacun ayant son propre impact. Mais s'ils sont pris ensemble, il devient clair que la médecine personnalisée n'était pas une promesse faite à la hâte, ni un objectif inaccessible; l'ensemble des preuves concernant l'efficacité de l'utilisation de la médecine génomique pour prescrire un traitement individualisé continue de croître. Bien que l'investissement dans la médecine génomique ne soit pas trivial, il fournit la base et les méthodes systématiques pour les progrès futurs. Plus de travail est nécessaire, plus de réflexion est nécessaire, et plus de lignes directrices doivent être établies, mais nous sommes constamment en train de traverser dans un nouveau paradigme où l'optimisation de la pratique médicale peut s'adresser spécifiquement à chaque patient nécessitant un traitement.



**Figure 3**

The pathway for identifying and validating drug targets for pharmacological therapies, showing a schematic representation of the process from biomarker discovery to a marketable drug. Genomic studies can potentially reduce the time it takes for drug discovery through enhanced understanding of the underlying biology and the discovery of relevant biomarkers; genomic medicine can enable drug development and improve efficacy and safety by identifying and targeting clinical studies to patient subsets most likely to respond to the drug.

**Figure 7 : La voie pour l'identification et la validation des cibles médicamenteuses pour les thérapies pharmacologiques, montrant une représentation schématique du processus de la découverte de biomarqueurs à un médicament commercialisable.**



## *V. Conclusion*

Alors que les progrès technologiques ont accéléré l'identification des biomarqueurs de la susceptibilité, du traitement et de la progression des maladies, le défi consiste à transporter ces découvertes vers les soins cliniques. Si la promesse de la génomique doit être réalisée, nous devons tester / adopter des méthodes scientifiques pour documenter comment ces technologies pourraient améliorer les soins de santé et prévenir les maladies. Les voies par lesquelles les marqueurs génomiques doivent voyager pour aider les cliniciens comprennent quatre phases: 1) Découverte, élucidation des voies et réplication chez les populations indépendantes, 2) Validation de l'efficacité et utilité clinique chez les populations à risque, 3) Exigences réglementaires pour l'approbation de la dispensation clinique, et 4) Démonstration de l'impact, de l'efficacité et des aspects sanitaires l'intégrité individuelle et la protection contre la discrimination génomique, inégalement respectée dans de nombreuses juridictions.

L'établissement de relations entre de nombreux marqueurs génomiques et des phénotypes cliniques est très difficile. À grande échelle, les études cliniques prospectives sont difficiles mais très nécessaires pour démontrer l'utilité du génotypage comme base pour la médecine individualisée. Il convient toutefois de noter que les tests génétiques peuvent être utiles sans une connaissance complète de la fonction biologique des variants génétiques en question. Une stratégie innovante qui relie de nombreux résultats cliniques aux marqueurs génomiques pléiotropiques inclut le PheWAS récemment proposé, une étude d'association à l'échelle du phénomène utilisée pour découvrir de nouvelles associations de maladies géniques. Ce qui est essentiel, c'est un programme multidisciplinaire de recherche en traduction qui englobe des recherches

cliniques et axées sur la population et le développement d'une collaboration entre les intervenants pour faciliter la traduction fondée sur des données probantes et le remboursement des tests. Alors que la médecine personnalisée entre dans la pratique médicale traditionnelle, les médecins et autres prestataires de soins devront administrer ou donner des conseils sur l'application d'un nombre croissant de tests génomiques, de thérapies pharmacogénomiques et de décisions thérapeutiques basées sur des preuves prédictives et l'estimation des risques. Les insuffisances des programmes de génétique dans les écoles de médecine pourraient entraver l'adoption de nouvelles connaissances génétiques. Cependant, plusieurs programmes ont récemment vu le jour qui préparent la prochaine génération de médecins, d'infirmières, de pharmaciens et d'autres travailleurs de la santé à la vague de médecine génomique à venir.

La perception du public est également très importante, car l'adoption clinique de nombreuses interventions génomiques proposées dépend de la compréhension du public. La médecine participative est un nouveau scénario où les patients deviennent des participants actifs dans leurs propres soins médicaux. L'environnement, le mode de vie, l'alimentation, les antécédents familiaux et les observations des patients sur leurs propres symptômes doivent être combinés aux données des biomarqueurs pour fournir une vision globale et intégrée des facteurs influençant le risque, la progression et les résultats du traitement. Les individus peuvent maintenant obtenir leur information génomique personnelle par le biais de tests génétiques directement auprès des consommateurs, mais quel impact, le cas échéant, cela aurait-il sur leur mode de vie et leur santé par rapport à un effet délétère? Jusqu'à présent, quelques études ont été réalisées, avec des preuves encourageantes mais limitées suggérant que l'information

génétiq ue pourrait entra  ner des changements dans les comportements li  s    l'hygi  ne de vie, comme l'abandon du tabac. Des   tudes plus approfondies de cette question sont n  cessaires mais, surtout, la r  alisation des soins de sant   personnalis  s, tout en s'appuyant sur la connaissance g  nomique de l'individu, requiert des informations environnementales tout aussi essentielles et la compr  hension de la capacit   de chaque individu    promouvoir sa sant  . Tous ces   l  ments sont importants pour promouvoir les quatre «P» de l'avenir de la m  decine: la m  decine pr  dictive, pr  ventive, personnalis  e et participative.





## *VI. Résumé*

## Résumé

**Titre:** Le séquençage de l'ADN et médecine personnalisée

**Auteur:** Mohamed En-nouali

**Most-clés:** Génomique- NGS–Médecine individualisée-Médecine personnalisée

L'évolution technique du séquençage des ADN s'est faite avec une rapidité exponentielle au cours des dernières années. La diffusion dans les laboratoires hospitaliers des nouveaux appareils de séquençage à haut débit (ou NGS) permet de passer de l'étude ponctuelle d'un gène à une analyse globale portant sur des dizaines ou centaines de gènes ou sur l'ensemble de l'exome ; elle modifie de ce fait les concepts médicaux et les conditions d'exercice, notamment en génétique et en cancérologie. Cette possibilité d'une recherche simultanée de mutations dans la séquence d'un grand nombre de gènes trouve aujourd'hui ses applications dans le diagnostic, éventuellement néonatal, dans le dépistage systématique des hétérozygotes et le diagnostic préconceptionnel généralisé. En oncogénétique constitutionnelle, les NGS réalisent également l'analyse simultanée des gènes impliqués dans les formes « héréditaires » de cancers. Par ailleurs dans l'ensemble des affections malignes, les NGS identifient l'ensemble des anomalies génomiques affectant le tissu malade, en s'efforçant de faire la distinction entre mutations « drivers » importantes pour la progression tumorale et mutations « passagers » ou accessoires; en s'efforçant également d'identifier des mutations « druggable » susceptibles de faire l'objet de thérapies ciblées. Enfin le séquençage systématique de l'ensemble des gènes impliqués dans le métabolisme et la réponse aux médicaments est le support d'une pharmacogénétique individuelle. En définitive, le séquençage des génomes tumoraux et constitutionnels, l'identification des mutations somatiques, la détermination des variants pharmacogénétiques, sont les éléments déterminants d'une médecine dite personnalisée. Les premiers résultats de ces indications thérapeutiques ciblées montrent un gain en termes de durée de rémission et de survie; reste néanmoins posée la question du coût de ces traitements. Ces énormes capacités de séquençage des génomes soulèvent également un grand nombre de questions réglementaires et éthiques.

## Summary

**Title:** DNA sequencing and personalized medicine

**Author:** Mohamed En-nouali

**Keywords:** Genomics–Next Generation Sequencing (NGS)-Individualized medicine–Personalized medicine.

DNA sequencing technologies have advanced at an exponential rate in recent years. More and more hospital laboratories are acquiring new high-throughput sequencing devices, (“ next-generation sequencers ”, NGS), allowing them to analyze tens or hundreds of genes, or even the entire exome. This is having a major impact on medical concepts and practices, especially with respect to genetics and oncology. This ability to search for mutations simultaneously in a large number of genes is finding applications in the diagnosis of Mendelian diseases (including at birth), routine screening for heterozygotes, and preconception diagnosis. Data on multifactorial diseases are still preliminary, but it should soon be possible to identify “strong ” factors of genetic predisposition that have so far been beyond the scope of genome-wide association studies (GWAS). In the field of constitutional oncogenetics, NGS can also be used for simultaneous analysis of genes involved in “hereditary ” cancers. More generally, NGS can identify all genomic abnormalities in a given malignant tissue (hemopathy or solid tumor), and has the potential to distinguish between important mutations (those that drive tumor progression) from “bystander ” or accessory mutations, and also to identify “ druggable ” mutations amenable to targeted therapies. Systematic sequencing of all the genes involved in drug metabolism and responsiveness will lead to individualized pharmacogenetics (warfarin therapy). Finally, sequencing of the tumoral and constitutional genomes, identification of somatic mutations, and detection of pharmacogenetic variants will open up the era of personalized medicine. The first results of these targeted therapeutic indications show a gain in the duration of remission and survival, although the cost-effectiveness of these approaches remains to be determined. Finally, this huge capacity for genome sequencing raises a number of regulatory and ethical issues.

# ملخص

العنوان: تسلسل الحمض النووي والطب الشخصي

من طرف: محمد النوالي

الكلمات الأساسية : علم الجينوم- الجيل الجديد التسلسل- الطب الفردي- الطب الشخصي

وقد تم التطور التقني لتسلسل الحمض النووي بسرعة هائلة في السنوات الأخيرة. المزيد والمزيد من مختبرات المستشفيات تكتسب جديد الأجهزة الإنتاجية عالية التسلسل ("الجيل القادم من التسلسل"، ج.ق.ت). مما يسمح لهم بتحليل عشرات أو مئات من الجينات، أو حتى إكسوم بأكمله. هذا له تأثير كبير على المفاهيم والممارسات الطبية، وخاصة فيما يتعلق علم الوراثة والأورام. هذه القدرة على البحث عن الطفرات في وقت واحد في عدد كبير من الجينات تم العثور لها على تطبيقات في تشخيص الأمراض المندلية، الفحص الروتيني للمتخالف، والتشخيص ما قبل الحمل. البيانات حول الأمراض متعددة العوامل لا تزال أولية، ولكن من الممكن قريباً تحديد العوامل "القوية" من الاستعداد الوراثي التي كانت حتى الآن خارج نطاق دراسات الارتباط على نطاق الجينوم (غواس). في مجال علم الوراثة السرطانية، ويمكن أيضاً أن تستخدم ج.ق.ت للتحليل في وقت واحد من الجينات التي تشارك في السرطان "الوراثي" (21 الجينات سرطان الثدي، 6 القولون وجينات السرطان، وما إلى ذلك). وبصورة أعم، يمكن أن يحدد ج.ق.ت جميع الشذوذ الجيني (الحذف، نقل المواقع، الطفرات) في نسيج خبيث معين (اعتلال المعدة أو الورم الصلب)، ولديه القدرة على التمييز بين الطفرات الهامة (تلك التي تدفع تطور الورم) من "المارة" أو الطفرات التبعية، وأيضاً لتحديد الطفرات "قابلة للتلف" قابلة للعلاجات المستهدفة (تراستوزوماب، وجيفيتينيب...). التسلسل المنهجي لجميع الجينات المشاركة في استقلاب الدواء والاستجابة يؤدي إلى علم الوراثة الدوائي الفردي. وأخيراً، فإن تسلسل الجينوم الورمي، وتحديد الطفرات الجسدية، والكشف عن فارماكوجينيتيك فتح عصر الطب الشخصي. النتائج الأولى لهذه المؤشرات العلاجية المستهدفة تظهر مكسباً في مدة الشفاء من المرض والبقاء على قيد الحياة، على الرغم من أن فعالية وتكلفة هذا النهج لا يزالان يتعين تحديده. وأخيراً، فإن هذه القدرة الضخمة على تسلسل الجينوم تثير عدداً من القضايا التنظيمية والأخلاقية.



## *VII. Bibliographie*

- [1] "deoxyribonucleic acid". Merriam-Webster Dictionary.
- [2] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2014). Molecular Biology of the Cell (6th ed.). Garland. p. Chapter 4: DNA, Chromosomes and Genomes. ISBN 9780815344322.
- [3] Purcell A. "DNA". Basic Biology.
- [4] Nuwer R (18 July 2015). "Counting All the DNA on Earth". The New York Times. New York: The New York Times Company. ISSN 0362-4331. Retrieved 2015-07-18.
- [5] "The Biosphere: Diversity of Life". Aspen Global Change Institute. Basalt, CO. Retrieved 2015-07-19.
- [6] Russell P (2001). iGenetics. New York: Benjamin Cummings. ISBN 0-8053-4553-1.
- [7] Mashaghi A, Katan A (2013). "A physicist's view of DNA". De Physicus. 24e (3): 59–61. Bibcode:2013arXiv1311.2545M. arXiv:1311.2545v1 [3](#).
- [8] Saenger W (1984). Principles of Nucleic Acid Structure. New York: Springer-Verlag. ISBN 0-387-90762-9.
- [9] ^ :<sup>a</sup> <sup>b</sup> Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Peter W (2002). Molecular Biology of the Cell (Fourth ed.). New York and London: Garland Science. ISBN 0-8153-3218-1. OCLC 145080076.

- [10] Irobalieva RN, Fogg JM, Catanese DJ, Catanese DJ, Sutthibutpong T, Chen M, Barker AK, Ludtke SJ, Harris SA, Schmid MF, Chiu W, Zechiedrich L (October 2015). "Structural diversity of supercoiled DNA". *Nature Communications*. 6: 8440. Bibcode: 2015NatCo...6E8440I. PMC 4608029 . PMID 26455586. doi: 10.1038/ncomms9440.
- [11] <sup>a b c</sup> Watson JD, Crick FH (April 1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid" (PDF). *Nature*. 171 (4356): 737–8. Bibcode:1953Natur.171..737W. PMID 13054692. doi:10.1038/171737a0.
- [12] Mandelkern M, Elias JG, Eden D, Crothers DM (October 1981). "The dimensions of DNA in solution". *Journal of Molecular Biology*. 152 (1): 153–61. PMID 7338906. doi:10.1016/0022-2836(81)90099-1.
- [13] Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, et al. (May 2006). "The DNA sequence and biological annotation of human chromosome 1". *Nature*. 441(7091): 315–21. Bibcode:2006Natur.441..315G. PMID 16710414. doi:10.1038/nature04727.
- [14] Watson JD, Crick FH (April 1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid" (PDF). *Nature*. 171 (4356): 737–8. Bibcode:1953Natur.171..737W. PMID 13054692. doi:10.1038/171737a0.
- [15] <sup>a b c</sup> Berg J., Tymoczko J. and Stryer L. (2002) *Biochemistry*. W. H. Freeman and Company ISBN 0-7167-4955-6

- [16] Abbreviations and Symbols for Nucleic Acids, Polynucleotides and their Constituents IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Retrieved 3 January 2006.
- [17] <sup>a b</sup> Ghosh A, Bansal M (April 2003). "A glossary of DNA structures from A to Z". *Acta Crystallographica. Section D, Biological Crystallography*. 59 (Pt 4): 620–6. PMID 12657780. doi:10.1107/S0907444903003251.
- [18] Created from PDB 1D65
- [19] Yakovchuk P, Protozanova E, Frank-Kamenetskii MD (2006). "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix". *Nucleic Acids Research*. 34 (2): 564–74. PMC 1360284 . PMID 16449200. doi:10.1093/nar/gkj454.
- [20] Burton E. Tropp – "Molecular Biology"- Jones and Barlett Learning, ISBN 978-0-7637-8663-2
- [21] "Watson-Crick Structure of DNA – 1953". Steven Carr. Memorial University of Newfoundland. Retrieved 13 July 2016.
- [22] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. (February 2001). "The sequence of the human genome". *Science*. 291 (5507): 1304–51. Bibcode:2001Sci...291.1304V. PMID 11181995. doi:10.1126/science.1058040.
- [23] Thanbichler M, Wang SC, Shapiro L (October 2005). "The bacterial nucleoid: a highly organized and dynamic structure". *Journal of Cellular Biochemistry*. 96 (3): 506–21. PMID 15988757. doi:10.1002/jcb.20519.



- [24] Wolfsberg TG, McEntyre J, Schuler GD (February 2001). "Guide to the draft human genome". *Nature*. 409 (6822): 824–6. Bibcode:2001Natur.409..824W. PMID 11236998. doi:10.1038/35057000.
- [25] Gregory TR (January 2005). "The C-value enigma in plants and animals: a review of parallels and an appeal for partnership". *Annals of Botany*. 95 (1): 133–46. PMID 15596463. doi:10.1093/aob/mci009.
- [26] Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, et al. (June 2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project". *Nature*. 447 (7146): 799–816. Bibcode:2007Natur.447..799B. PMC 2212820 . PMID 17571346. doi:10.1038/nature05874.
- [27] Created from PDB 1MSW
- [28] Pidoux AL, Allshire RC (March 2005). "The role of heterochromatin in centromere function". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 360 (1455): 569–79. PMC 1569473 . PMID 15905142. doi:10.1098/rstb.2004.1611.
- [29] Harrison PM, Hegyi H, Balasubramanian S, Luscombe NM, Bertone P, Echols N, Johnson T, Gerstein M (February 2002). "Molecular fossils in the human genome: identification and analysis of the pseudogenes in chromosomes 21 and 22". *Genome Research*. 12 (2): 272–80. PMC 155275 . PMID 11827946. doi:10.1101/gr.207102.

- [30] Harrison PM, Gerstein M (May 2002). "Studying genomes through the aeons: protein families, pseudogenes and proteome evolution". *Journal of Molecular Biology*. 318 (5): 1155–74. doi:10.1016/S0022-2836(02)00109-2. PMID 12083509.
- [31] Albà M (2001). "Replicative DNA polymerases". *Genome Biology*. 2 (1): REVIEWS3002. doi:10.1186/gb-2001-2-1-reviews3002. PMC 150442 . PMID 11178285.
- [32] Miescher, Friedrich (1871) "Ueber die chemische Zusammensetzung der Eiterzellen"(On the chemical composition of pus cells), *Medicinisch-chemische Untersuchungen*, 4 : 441–460. From p. 456: "Ich habe mich daher später mit meinen Versuchen an die ganzen Kerne gehalten, die Trennung der Körper, die ich einstweilen ohne weiteres Präjudiz als lösliches und unlösliches Nuclein bezeichnen will, einem günstigeren Material überlassend."(Therefore, in my experiments I subsequently limited myself to the whole nucleus, leaving to a more favorable material the separation of the substances, that for the present, without further prejudice, I will designate as soluble and insoluble nuclear material ("Nuclein").)
- [33] Dahm R (January 2008). "Discovering DNA: Friedrich Miescher and the early years of nucleic acid research". *Human Genetics*. 122 (6): 565–81. doi:10.1007/s00439-007-0433-0. PMID 17901982.
- [34] See:
- Albrecht Kossel (1879) "Ueber Nuclein der Hefe" (On nuclein in yeast) *Zeitschrift für physiologische Chemie*, 3 : 284–291.

- Albrect Kossel (1880) "Ueber Nuclein der Hefe II" (On nuclein in yeast, Part 2) *Zeitschrift für physiologische Chemie*, 4 : 290–295.
- Albrect Kossel (1881) "Ueber die Verbreitung des Hypoxanthins im Thier- und Pflanzenreich" (On the distribution of hypoxanthins in the animal and plant kingdoms) *Zeitschrift für physiologische Chemie*, 5 : 267–271.
- Albrect Kossel, *Untersuchungen über die Nucleine und ihre Spaltungsprodukte*[Investigations into nuclein and its cleavage products] (Strassburg, Germany: K.J. Trübne, 1881), 19 pages.
- Albrect Kossel (1882) "Ueber Xanthin und Hypoxanthin" (On xanthin and hypoxanthin), *Zeitschrift für physiologische Chemie*, 6 : 422–431.
- Albrect Kossel (1883) "Zur Chemie des Zellkerns" (On the chemistry of the cell nucleus), *Zeitschrift für physiologische Chemie*, 7 : 7–22.
- Albrect Kossel (1886) "Weitere Beiträge zur Chemie des Zellkerns" (Further contributions to the chemistry of the cell nucleus), *Zeitschrift für Physiologische Chemie*, 10 : 248–264. Available online at: Max Planck Institute for the History of Science, Berlin, Germany. On p. 264, Kossel remarked presciently: "Der Erforschung der quantitativen Verhältnisse der vier stickstoffreichen Basen, der Abhängigkeit ihrer Menge von den physiologischen Zuständen der Zelle, verspricht wichtige Aufschlüsse über die elementaren physiologisch-chemischen Vorgänge." (The study of the quantitative relations of the four nitrogenous bases — [and] of the dependence of their quantity on the physiological states of the cell — promises important insights into the fundamental physiological-chemical processes.)

- [35] Jones ME (September 1953). "Albrecht Kossel, a biographical sketch". The Yale Journal of Biology and Medicine. National Center for Biotechnology Information. 26 (1): 80–97. PMC 2599350 [3](#). PMID 13103145.
- [36] Levene P (1 December 1919). "The structure of yeast nucleic acid". J Biol Chem. 40 (2): 415–24.
- [37] See:
1. W. T. Astbury and Florence O. Bell (1938) "Some recent developments in the X-ray study of proteins and related structures," Cold Spring Harbor Symposia on Quantitative Biology, 6 : 109–121. Available on-line at: University of Leeds.
  2. Astbury, W. T., (1947) "X-ray studies of nucleic acids," Symposia of the Society for Experimental Biology, 1 : 66–76. Available on-line at: Oregon State University.
- [38] Koltsov proposed that a cell's genetic information was encoded in a long chain of amino acids. See:
1. Н. К. Кольцов, "Физико-химические основы морфологии" (The physical-chemical basis of morphology) – speech given at the 3rd All-Union Meeting of Zoologist, Anatomists, and Histologists at Leningrad, U.S.S.R., December 12, 1927.
  2. Reprinted in: Успехи экспериментальной биологии (Advances in Experimental Biology), series B, 7 (1) : ?-? (1928).
  3. Reprinted in German as: Nikolaj K. Koltzoff (1928) "Physikalisch-chemische Grundlagen der Morphologie" (The physical-chemical basis of morphology), Biologisches Zentralblatt, 48 (6) : 345–369.

4. In 1934, Koltsov contended that the proteins that contain a cell's genetic information replicate. See: N. K. Koltzoff (October 5, 1934) "The structure of the chromosomes in the salivary glands of *Drosophila*," *Science*, 80 (2075) : 312–313. From page 313: "I think that the size of the chromosomes in the salivary glands [of *Drosophila*] is determined through the multiplication of genonemes. By this term I designate the axial thread of the chromosome, in which the geneticists locate the linear combination of genes; ... In the normal chromosome there is usually only one genoneme; before cell-division this genoneme has become divided into two strands."
- [39] Soyfer VN (September 2001). "The consequences of political dictatorship for Russian science". *Nature Reviews. Genetics*. 2 (9): 723–9. doi:10.1038/35088598. PMID 11533721.
- [40] Griffith F (January 1928). "The Significance of Pneumococcal Types". *The Journal of Hygiene*. 27 (2): 113–59. doi:10.1017/S0022172400031879. PMC 2167760 . PMID 20474956.
- [41] Lorenz MG, Wackernagel W (September 1994). "Bacterial gene transfer by natural genetic transformation in the environment". *Microbiological Reviews*. 58 (3): 563–602. PMC 372978 . PMID 7968924.
- [42] Avery OT, Macleod CM, McCarty M (February 1944). "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type III". *The Journal of Experimental Medicine*. 79 (2): 137–58. doi:10.1084/jem.79.2.137. PMC 2135445 . PMID 19871359.

- [43] Hershey AD, Chase M (May 1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage". *The Journal of General Physiology*. 36 (1): 39–56. doi:10.1085/jgp.36.1.39. PMC 2147348 . PMID 12981234.
- [44] The B-DNA X-ray pattern on the right of this linked image was obtained by Rosalind Franklin and Raymond Gosling in May 1952 at high hydration levels of DNA and it has been labeled as "Photo 51"
- [45] Regis, Ed (2009) *What Is Life?: investigating the nature of life in the age of synthetic biology*. Oxford: Oxford University Press ISBN 0-19-538341-9; p. 52
- [46] Nature Archives Double Helix of DNA: 50 Years
- [47] "Original X-ray diffraction image". [Osulibrary.oregonstate.edu](http://osulibrary.oregonstate.edu). Retrieved 6 February 2011.
- [48] The Nobel Prize in Physiology or Medicine 1962 [Nobelprize .org](http://Nobelprize.org) Accessed 22 December 06
- [49] Maddox B (January 2003). "The double helix and the 'wronged heroine'" (PDF). *Nature*. 421 (6921): 407–8. Bibcode:2003Natur.421..407M. doi:10.1038/nature01399. PMID 12540909.
- [50] Crick, F.H.C. On degenerate templates and the adaptor hypothesis (PDF). Archived1 October 2008 at the Wayback Machine. [genome.wellcome.ac.uk](http://genome.wellcome.ac.uk) (Lecture, 1955). Retrieved 22 December 2006.

- [51] Meselson M, Stahl FW (July 1958). "THE REPLICATION OF DNA IN ESCHERICHIA COLI". *Proceedings of the National Academy of Sciences of the United States of America*. 44 (7): 671–82. Bibcode:1958PNAS...44..671M. doi:10.1073/pnas.44.7.671. PMC 528642 . PMID 16590258.
- [52] The Nobel Prize in Physiology or Medicine 1968 Nobelprize.org Accessed 22 December 06
- [53] Pray, L. (2008) Discovery of DNA structure and function: Watson and Crick. *Nature Education* 1(1):100
- [54] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 1977;74(12):5463–7.
- [55] Kircher M, Kelso J. High-throughput DNA sequencing – concepts and limitations. *Bioessays* 2010;32(6):524–36.
- [56] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822): 860–921.
- [57] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science* 2001;291(5507):1304–51.
- [58] Finishing the euchromatic sequence of the human genome. *Nature* 2004; 431(7011):931–45.

- [59] Maxam AM, Gilbert W. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America* 1977;74(2):560–4.
- [60] Metzker ML. Sequencing technologies – the next generation. *Nature Reviews Genetics* 2010;11(1):31–46.
- [61] Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. *Human Molecular Genetics* 2010;19(R2):R227–40.
- [62] Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. *Science* 2009;323(5910):133–8.
- [63] Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, et al. The challenges of sequencing by synthesis. *Nature Biotechnology* 2009;27(11):
- [64] Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bembien LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437(7057):376–80.
- [65] Ronaghi M, Uhlen M, Nyren P. A sequencing method based on real-time pyrophosphate. *Science* 1998;281(5375):363–5.
- [66] Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P. Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry* 1996;242(1):84–9.
- [67] Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008;452(7189):872–6.



- [68] Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 2008;456(7218):53–9.
- [69] Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Research* 2008;18(7):1051–63.
- [70] Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, et al. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 2005;309(5741):1728–32.
- [71] Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 2011;475(7356):348–52.
- [72] Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, et al. Singlemolecule DNA sequencing of a viral genome. *Science* 2008;320(5872):106–9.
- [73] Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 2010;327(5961):78–81.
- [74] Kasianowicz JJ, Brandin E, Branton D, Deamer DW. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences of the United States of America* 1996;93(24): 13770–73.

- [75] Bayley H. Sequencing single molecules of DNA. *Current Opinion in Chemical Biology* 2006;10(6):628–37.
- [76] Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, et al. The potential and challenges of nanopore sequencing. *Nature Biotechnology* 2008;26(10):1146–53.
- [77] Derrington IM, Butler TZ, Collins MD, Manrao E, Pavlenok M, Niederweis M, et al. Nanopore DNA sequencing with MspA. *Proceedings of the National Academy of Sciences of the United States of America* 2010;107(37):16060–65.
- [78] Schneider GF, Dekker C. DNA sequencing with nanopores. *Nature Biotechnology* 2012;30(4):326–8.
- [79] Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology* 2009;4(4):265–70.
- [80] Wu HC, Astier Y, Maglia G, Mikhailova E, Bayley H. Protein nanopores with covalently attached molecular adapters. *Journal of the American Chemical Society* 2007;129(51):16142–48.
- [81] Astier Y, Braha O, Bayley H. Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *Journal of the American Chemical Society* 2006;128(5):1705–10.

- [82] Stoddart D, Heron AJ, Klingelhoefer J, Mikhailova E, Maglia G, Bayley H. Nucleobase recognition in ssDNA at the central constriction of the  $\alpha$ -hemolysin pore. *Nano Letters* 2010;10(9):3633–7.
- [83] Manrao EA, Derrington IM, Laszlo AH, Langford KW, Hopper MK, Gillgren N, et al. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nature Biotechnology* 2012;30(4):349–53.
- [84] Cherf GM, Lieberman KR, Rashid H, Lam CE, Karplus K, Akeson M. Automated forward and reverse ratcheting of DNA in a nanopore at 5-A precision. *Nature Biotechnology* 2012;30(4):344–8.
- [85] Garaj S, Hubbard W, Reina A, Kong J, Branton D, Golovchenko JA. Graphene as a subnanometre trans-electrode membrane. *Nature* 2010;467(7312):190–3.
- [86] Dekker C. Solid-state nanopores. *Nature Nanotechnology* 2007;2(4):209–15.
- [87] Banerjee A, Mikhailova E, Cheley S, Gu LQ, Montoya M, Nagaoka Y, et al. Molecular bases of cyclodextrin adapter interactions with engineered protein nanopores. *Proceedings of the National Academy of Sciences of the United States of America* 2010;107(18):8165–70.
- [88] Larsson C, Grundberg I, Söderberg O, Nilsson M. In situ detection and genotyping of individual mRNA molecules. *Nature Methods* 2010;7(5):395–7.

- [89] Aston C, Mishra B, Schwartz DC. Optical mapping and its potential for largescale sequencing projects. *Trends in Biotechnology* 1999;17(7):297–302.
- [90] Westerlund F, et al. Fluorescence enhancement of single DNA molecules con-fined in Si/SiO<sub>2</sub> nanochannels. *Lab Chip* 2010;10(16):2049–51.
- [91] Guttmacher AE, Collins FS, Carmona RH. 2004. The family history: more important than ever.
- [92] Rich E, BurkeW, Heaton C, Haga S, Pinsky L, et al. 2004. Reconsidering the family history in primary care. *J. Gen. Intern. Med.* 19:273–80
- [93] Ginsburg GS, Willard HF. 2009. Genomic and personalized medicine: foundations and applications. *Transl. Res.* 154:277–87
- [94] KannelW, Dawber T, Kagan A, Revotskie N, Stokes J III. 1961. Factors of risk in the development of coronary heart disease: six year follow-up experience: the Framingham Study. *Ann. Intern.Med.* 55:33–50
- [95] Gail M, Greene M. 2000. Gail model and breast cancer. *Lancet* 355:1017
- [96] Liu J, Wyatt J, Altman D. 2006. Decision tools in health care: focus on the problem, not the solution. *BMC Med. Inform. Decis. Mak.* 6:4
- [97] Osheroff JA, Teich JM, Middleton B, Steen EB, Wright A, Detmer DE. 2007. A roadmap for national action on clinical decision support. *J. Am. Med. Inform. Assoc.* 14:141–45

- [98] Bennett JW, Glasziou PP. 2003. Computerised reminders and feedback in medication management: a systematic review of randomised controlled trials. *Med. J. Aust.* 178:217–22
- [99] Shea S, DuMouchel W, Bahamonde L. 1996. A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting. *J. Am. Med. Inform. Assoc.* 3:399–409
- [100] Walton R, Dovey S, Harvey E, Freemantle N. 1999. Computer support for determining drug dose: systematic review and meta-analysis. *BMJ* 318:984–90
- [101] Collins FS. 1999. Shattuck lecture: medical and societal consequences of the Human Genome Project. *N. Engl. J. Med.* 341:28–37
- [102] Abrahams E, Silver M. 2009. The case for personalized medicine. *J. Diabetes Sci. Technol.* 3:680–84
- [103] Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, et al. 2009. Repeatability of published microarray gene expression analyses. *Nat. Genet.* 41:149–55
- [104] Secr. Advis. Comm. Genet. Health Soc. 2008. U.S. system of oversight of genetic testing: a response to the charge of the secretary of health and human services. Rep., Dep. Health Hum. Serv., Washington, DC
- [105] Ginsburg GS. 2011. Application of molecular technologies to clinical medicine. In *Cecil Medicine*, ed. L Goldman, AI Schafer, pp. 199–203. Philadelphia: Elsevier. 24th ed.

- [106] Willard H, Angrist M, Ginsburg G. 2005. Genomic medicine: genetic variation and its impact on the future of health care. *Philos. Trans. R. Soc. B* 360:1543–50
- [107] Weinshilboum R, Wang L. 2004. Pharmacogenomics: bench to bedside. *Nat. Rev. Drug Discov.* 3:739–48
- [108] Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, et al. 2009. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N. Engl. J. Med.* 360:753–64
- [109] Feero WG, Guttmacher AE, Collins FS. 2010. Genomic medicine: an updated primer. *N. Engl. J. Med.* 362:2001–11
- [110] Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, et al. 2006. Global variation in copy number in the human genome. *Nature* 444:444–54
- [111] Bonetta L. 2010. Whole-genome sequencing breaks the cost barrier. *Cell* 141:917–19
- [112] Bentley DR. 2006. Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* 16:545–52
- [113] Service RF. 2006. The race for the \$1000 genome. *Science* 311:1544–46
- [114] Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, et al. 2007. The diploid genome sequence of an individual human. *PLoS Biol.* 5:e254
- [115] Pushkarev D, Neff NF, Quake SR. 2009. Single-molecule sequencing of an individual human genome *Nat. Biotechnol.* 27:847–50

- [116] Wang J, Wang W, Li R, Li Y, Tian G, et al. 2008. The diploid genome sequence of an Asian individual. *Nature* 456:60–65
- [117] Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, et al. 2008. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452:872–76
- [118] Bolze A, Byun M, McDonald D, Morgan NV, Abhyankar A, et al. 2010. Whole-exome-sequencing-based discovery of human FADD deficiency. *Am. J. Hum. Genet.* 87:873–81
- [119] Dahl F, Stenberg J, Fredriksson S, Welch K, Zhang M, et al. 2007. Multigene amplification and massively
- [120] Hornsey M, Loman N, Wareham DW, Ellington MJ, Pallen MJ, et al. 2011. Whole-genome comparison of two *Acinetobacter baumannii* isolates from a single patient, where resistance developed during tigecycline therapy. *J. Antimicrob. Chemother.* 66:1499–503
- [121] Kennemann L, Didelot X, Aebischer T, Kuhn S, Drescher B, et al. 2011. *Helicobacter pylori* genome evolution during human infection. *Proc. Natl. Acad. Sci. USA* 108:5033–38
- [122] Lupski JR, Reid JG, Gonzaga-Jauregui C, Rio Deiros D, Chen DC, et al. 2010. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N. Engl. J. Med.* 362:1181–91
- [123] Tsurusaki Y, Osaka H, Hamanoue H, Shimbo H, Tsuji M, et al. 2011. Rapid detection of a mutation causing X-linked leucoencephalopathy by exome sequencing. *J. Med. Genet.* In press

- [124] Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, et al. 2010. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328:636–39
- [125] Welch JS, Westervelt P, Ding L, Larson DE, Klco JM, et al. 2011. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA* 305:1577–84
- [126] Worthey EA, Mayer AN, Syverson GD, Helbling D, Bonacci BB, et al. 2011. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet. Med.* 13:255–62
- [127] Deloukas P, Bentley D. 2004. The HapMap project and its application to genetic studies of drug response. *Pharmacogenomics J.* 4:88–90
- [128] Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, et al. 2002. The structure of haplotype blocks in the human genome. *Science* 296:2225–29
- [129] Weiss KM, Clark AG. 2002. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* 18:19–24
- [130] Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TO. 2011. A catalog of published genome-wide association studies. <http://www.genome.gov/gwastudies>
- [131] Johnson AD, O'Donnell CJ. 2009. An open access database of genome-wide association results. *BMC Med. Genet.* 10:6



- [132] Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, et al. 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* 40:955–62
- [133] McCarthy MI. 2010. Genomics, type 2 diabetes, and obesity. *N. Engl. J. Med.* 363:2339–50
- [134] Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445:881–85
- [135] Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, et al. 2007. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 316:1491–93
- [136] McCarthy JJ, Parker A, Salem R, Moliterno DJ, Wang Q, et al. 2004. Large scale association analysis for identification of genes underlying premature coronary heart disease: cumulative perspective from analysis of 111 candidate genes. *J. Med. Genet.* 41:334–41
- [137] . McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, et al. 2007. A common allele on chromosome 9 associated with coronary heart disease. *Science* 316:1488–91
- [138] Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, et al. 2007. Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* 357:443–53
- [139] Marshall E. 2009. Lawsuit challenges legal basis for patenting human genes. *Science* 324:1000–1

- [140] Hannon GJ, Beach D. 1994. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature* 371:257–61
- [141] Pasmant E, Laurendeau I, Heron D, Vidaud M, Vidaud D, Bieche I. 2007. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.* 67:3963–69
- [142] . Visel A, Zhu Y, May D, Afzal V, Gong E, et al. 2010. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* 464:409–12
- [143] . Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, et al. 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* 38:662–74
- [144] Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, et al. 2011. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature* 470:264–68
- [145] Huang Y, Qiao F, Atkinson C, Holers VM, Tomlinson S. 2008. A novel targeted inhibitor of the alternative pathway of complement and its therapeutic application in ischemia/reperfusion injury. *J. Immunol.* 181:8068–76
- [146] Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, et al. 2005. Complement factor H polymorphism in age-related macular degeneration. *Science* 308:385–89

- [147] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, et al. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 461:399–401
- [148] Thompson AJ, Muir AJ, Sulkowski MS, Ge D, Fellay J, et al. 2010. Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. *Gastroenterology* 139:120–29
- [149] Link E, Parish S, Armitage J, Bowman L, Heath S, et al. 2008. SLCO1B1 variants and statin-induced myopathy: a genomewide study. *N. Engl. J. Med.* 359:789–99
- [150] Simon R. 2003. Diagnostic and prognostic prediction using gene expression profiles in high-dimensional microarray data. *Br. J. Cancer* 89:1599–604
- [151] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, et al. 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* 27:1160–67
- [152] Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, et al. 2002. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 346:1937–47
- [153] Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, et al. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503–11

- [154] Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, et al. 2001. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl. Acad. Sci. USA* 98:13790–95
- [155] 17. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, et al. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406:536–40 18. 34
- [156] Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–36
- [157] Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531–37
- [158] Bray NJ. 2008. Gene expression in the etiology of schizophrenia. *Schizophr. Bull.* 34:412–18
- [159] Comabella M, Martin R. 2007. Genomics in multiple sclerosis: current state and future directions. *J. Neuroimmunol.* 187:1–8
- [160] Goes FS, Sanders LL, Potash JB. 2008. The genetics of psychotic bipolar disorder. *Curr. Psychiatry Rep.* 10:178–89
- [161] Kittleson MM, Hare JM. 2005. Molecular signature analysis: the potential of gene-expression analysis in cardiomyopathy. *Future Cardiol.* 1:793–808

- [162] Mehta D, Menke A, Binder EB. 2010. Gene expression studies in major depression. *Curr. Psychiatry Rep.* 12:135–44
- [163] Van Baarsen LG, Bos CL, van der Pouw Kraan TC, Verweij CL. 2009. Transcription profiling of rheumatic diseases. *Arthritis Res. Ther.* 11:207
- [164] Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–97
- [165] Lee RC, Feinbaum RL, Ambros V. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75:843–54
- [166] Feng J, Sun G, Yan J, Noltner K, Li W, et al. 2009. Evidence for X-chromosomal schizophrenia associated with microRNA alterations. *PLoS One* 4:e6121
- [167] Iorio MV, Croce CM. 2009. MicroRNAs in cancer: small molecules with a huge impact. *J. Clin. Oncol.* 27:5848–56
- [168] Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. 2005. Modulation of hepatitis C virus RNA abundance by a liver-specific microRNA. *Science* 309:1577–81
- [169] O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. 2005. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 435:839–43
- [170] Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, et al. 2007. Requirement of *bic/microRNA-155* for normal immune function. *Science* 316:608–11

- [171] Tatsuguchi M, Seok HY, Callis TE, Thomson JM, Chen JF, et al. 2007. Expression of microRNAs is dynamically regulated during cardiomyocyte hypertrophy. *J. Mol. Cell. Cardiol.* 42:1137–41
- [172] Budhu A, Ji J, Wang XW. 2010. The clinical potential of microRNAs. *J. Hematol. Oncol.* 3:37
- [173] Calin GA, Croce CM. 2006. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* 6:857–66
- [174] Kartha RV, Subramanian S. 2010. MicroRNAs in cardiovascular diseases: biology and potential clinical applications. *J. Cardiovasc. Transl. Res.* 3:256–70
- [175] Xiao J, Yang B, Lin H, Lu Y, Luo X, Wang Z. 2007. Novel approaches for gene-specific interference via manipulating actions of microRNAs: examination on the pacemaker channel genes HCN2 and HCN4. *J. Cell. Physiol.* 212:285–92
- [176] Davis ME, Zuckerman JE, Choi CH, Seligson D, Tolcher A, et al. 2010. Evidence of RNAi in humans from systemically administered siRNA via targeted nanoparticles. *Nature* 464:1067–70
- [177] Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. 2003. Prediction of mammalian microRNA targets. *Cell* 115:787–98
- [178] Thomas M, Lieberman J, Lal A. 2010. Desperately seeking microRNA targets. *Nat. Struct. Mol. Biol.* 17:1169–74
- [179] Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322:1845–48

- [180] He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. 2008. The antisense transcriptomes of human cells. *Science* 322:1855–57
- [181] Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40:1413–15
- [182] Sugarbaker DJ, Richards WG, Gordon GJ, Dong L, De Rienzo A, et al. 2008. Transcriptome sequencing of malignant pleural mesothelioma tumors. *Proc. Natl. Acad. Sci. USA* 105:3521–26
- [183] Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, et al. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321:956–60
- [184] Guffanti A, Iacono M, Pelucchi P, Kim N, Solda G, et al. 2009. A transcriptional sketch of a primary human breast cancer by 454 deep sequencing. *BMC Genomics* 10:163
- [185] Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, et al. 2009. Transcriptome sequencing to detect gene fusions in cancer. *Nature* 458:97–101
- [186] Palanisamy N, Ateeq B, Kalyana-Sundaram S, Pflueger D, Ramnarayanan K, et al. 2010. Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma. *Nat. Med.* 16:793–98

- [187] Zhao Q, Caballero OL, Levy S, Stevenson BJ, Iseli C, et al. 2009. Transcriptome-guided characterization of genomic rearrangements in a breast cancer cell line. *Proc. Natl. Acad. Sci. USA* 106:1886–91
- [188] Du X, Callister SJ, Manes NP, Adkins JN, Alexandridis RA, et al. 2008. A computational strategy to analyze label-free temporal bottom-up proteomics data. *J. Proteome Res.* 7:2595–604
- [189] Issaq HJ, Veenstra TD. 2008. Would you prefer multiple reaction monitoring or antibodies with your biomarker validation? *Expert Rev. Proteomics* 5:761–63
- [190] Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. 2009. Metabolomics applied to diabetes research: moving from information to knowledge. *Diabetes* 58:2429–43
- [191] Griffin JL, Shockcor JP. 2004. Metabolic profiles of cancer cells. *Nat. Rev. Cancer* 4:551–61
- [192] Kaddurah-Daouk R, McEvoy J, Baillie RA, Lee D, Yao JK, et al. 2007. Metabolomic mapping of atypical antipsychotic effects in schizophrenia. *Mol. Psychiatry* 12:934–45
- [193] Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, et al. 2009. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* 9:311–26
- [194] Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, et al. 2005. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation* 112:3868–75



- [195] Shah SH, Hauser ER, Bain JR, Muehlbauer MJ, Haynes C, et al. 2009. High heritability of metabolomics profiles in families burdened with premature cardiovascular disease. *Mol. Syst. Biol.* 5:258
- [196] Ebbels TM, Keun HC, Beckonert OP, Bollard ME, Lindon JC, et al. 2007. Prediction and classification of drug toxicity using probabilistic modeling of temporal metabolic data: the consortium on metabonomic
- [197] Keun HC. 2006. Metabonomic modeling of drug toxicity. *Pharmacol. Ther.* 109:92–106
- [198] Nicholson JK, Connelly J, Lindon JC, Holmes E. 2002. Metabonomics: a platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* 1:153–61
- [199] . Sharma S, Kelly TK, Jones PA. 2010. Epigenetics in cancer. *Carcinogenesis* 31:27–36
- [200] Bird AP. 1986. CpG-rich islands and the function of DNA methylation. *Nature* 321:209–13
- [201] Prendergast GC, Ziff EB. 1991. Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science* 251:186–89
- [202] Archey WB, McEachern KA, Robson M, Offit K, Vaziri SA, et al. 2002. Increased CpG methylation of the estrogen receptor gene in BRCA1-linked estrogen receptor-negative breast cancers. *Oncogene* 21:7034–41
- [203] Feinberg AP, Vogelstein B. 1983. Hypomethylation of ras oncogenes in primary human cancers. *Biochem. Biophys. Res. Commun.* 111:47–54

- [204] Rosty C, Ueki T, Argani P, Jansen M, Yeo CJ, et al. 2002. Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation. *Am. J. Pathol.* 160:45–50
- [205] Corwin EJ. 2004. The concept of epigenetics and its role in the development of cardiovascular disease: commentary on “New and emerging theories of cardiovascular disease.” *Biol. Res. Nurs.* 6:11–16, 21–23
- [206] Devaskar SU, Thamocharan M. 2007. Metabolic programming in the pathogenesis of insulin resistance. *Rev. Endocr. Metab. Disord.* 8:105–13
- [207] Ji Q, Hao X, Meng Y, Zhang M, Desano J, et al. 2008. Restoration of tumor suppressor miR-34 inhibit human p53-mutant gastric cancer tumorspheres. *BMC Cancer* 8:266
- [208] . Lu Q, Kaplan M, Ray D, Zacharek S, Gutsch D, Richardson B. 2002. Demethylation of ITGAL(CD11a) regulatory sequences in systemic lupus erythematosus. *Arthritis Rheum.* 46:1282–91
- [209] Am. Assoc. Cancer Res. Hum. Epigenome Task Force, Eur. Union Netw. Excell. Sci. Advis. Board. 2008. Moving AHEAD with an international human epigenome project. *Nature* 454:711–15
- [210] Bernstein BE, Meissner A, Lander ES. 2007. The mammalian epigenome. *Cell* 128:669–81

- [211] Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, et al. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* 38:1378–85
- [212] Willard HF, Ginsburg GS. 2009. *Genomic and Personalized Medicine*. Amsterdam: Elsevier. 1478 pp.
- [213] Gonzalez-Angulo AM. 2010. Reply to D.H. Roukos and D.A.X. Schinagl et al. *J. Clin. Oncol.* 28:e282–83
- [214] Gonzalez-Angulo AM, Hennessy BT, Mills GB. 2010. Future of personalized medicine in oncology: a systems biology approach. *J. Clin. Oncol.* 28:2777–83
- [215] Roukos DH. 2010. Beyond HER2 and trastuzumab: heterogeneity, systems biology, and cancer origin research may guide the future for personalized treatment of very early but aggressive breast cancer. *J. Clin. Oncol.* 28:e279–80
- [216] Wang K, Lee I, Carlson G, Hood L, Galas D. 2010. Systems biology and the discovery of diagnostic biomarkers. *Dis. Markers* 28:199–207
- [217] Bell C. J., Dinwiddie D C., Miller N. A. et al. — Carrier testing for severe childhood recessive diseases by next generation sequencing. *Sci. Transl. Med.*, 2011, 3, 65ra4.
- [218] Steffann J., Frydman N., Burlet Ph. et al. Le diagnostic pré-implantatoire couplé au typage HLA: l'expérience française. *Bull. Acad. Natle Med.*, 2011, 195(4-5), 1015-22.
- [219] Fan H. C., Gu W., Wang J. et al. Non-invasive prenatal measurement of the fetal genome. *Nature*. 2012, 487, 320-24.

- [220] Snyder M.W., Simmons L.E., Kitzman J.O. et al.—Noninvasive fetal genome sequencing: a primer. *Prenatal diagnosis*. 2013, 33, 547-54.
- [221] Dugoff L.—Application of genomic technology in prenatal diagnosis. *N. Engl. J. Med.*, 2012, 367, 2249-50.
- [222] Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M. 2010. Genetic susceptibility to breast cancer. *Mol. Oncol.* 4:174–91
- [223] Trainer AH, Lewis CR, Tucker K, Meiser B, Friedlander M, Ward RL. 2010. The role of BRCA mutation testing in determining breast cancer therapy. *Nat. Rev. Clin. Oncol.* 7:708–17
- [224] Vilar E, Gruber SB. 2010. Microsatellite instability in colorectal cancer: the stable evidence. *Nat. Rev. Clin. Oncol.* 7:153–62
- [225] Hedley PL, Jorgensen P, Schlamowitz S, Wangari R, Moolman-Smook J, et al. 2009. The genetic basis of long QT and short QT syndromes: a mutation update. *Hum. Mutat.* 30:1486–511
- [226] Tester DJ, Ackerman MJ. 2011. Genetic testing for potentially lethal, highly treatable inherited cardiomyopathies/ channelopathies in clinical practice. *Circulation* 123:1021–37
- [227] Shimizu W. 2008. Clinical impact of genetic studies in lethal inherited cardiac arrhythmias. *Circ. J.* 72:1926–36
- [228] Lee W, Jiang Z, Liu J, Haverty PM, Guan Y, et al. 2010. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature* 465:473–77

- [229] Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, et al. 2008. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 456:66–72
- [230] Pleasance ED, Stephens PJ, O’Meara S, McBride DJ, Meynert A, et al. 2010. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 463:184–90
- [231] Garber K. 2005. Human Cancer Genome Project moving forward despite some doubts in community. *J. Natl. Cancer Inst.* 97:1322–24
- [232] Cancer Genome Atlas Res. Netw. 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455:1061–68
- [233] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, et al. 2010. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR,
- [234] Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531–37
- [235] Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, et al. 2001. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl. Acad. Sci. USA* 98:13790–95

- [236] Huang JZ, Sanger WG, Greiner TC, Staudt LM, Weisenburger DD, et al. 2002. The t(14;18) defines a unique subset of diffuse large B-cell lymphoma with a germinal center B-cell gene expression profile. *Blood* 99:2285–90
- [237] Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, et al. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* 98:10869–74
- [238] Mateos-Caceres PJ, Garcia-Mendez A, Lopez Farre A, Macaya C, Nunez A, et al. 2004. Proteomic analysis of plasma from patients during an acute coronary syndrome. *J. Am. Coll. Cardiol.* 44:1578–83
- [239] Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, et al. 2005. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation* 112:3868–75
- [240] . Fernandez Parguina A, Grigorian-Shamajian L, Agra RM, Teijeira-Fernandez E, Rosa I, et al. 2010. Proteins involved in platelet signaling are differentially regulated in acute coronary syndrome: a proteomic study. *PLoS One* 5:e13404
- [241] Rosenberg S, Elashoff MR, Beineke P, Daniels SE, Wingrove JA, et al. 2010. Multicenter validation of the diagnostic accuracy of a blood-based gene expression test for assessing obstructive coronary artery disease in nondiabetic patients. *Ann. Intern. Med.* 153:425–34
- [242] Wingrove JA, Daniels SE, Sehnert AJ, Tingley W, Elashoff MR, et al. 2008. Correlation of peripheral blood gene expression with the extent of coronary artery stenosis. *Circ. Cardiovasc. Genet.* 1:31–38

- [243] Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, et al. 2008. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N. Engl. J. Med.* 359:1995–2004
- [244] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, et al. 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* 27:1160–67
- [245] Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, et al. 2002. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 346:1937–47
- [246] Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–36
- [247] Cobleigh MA, Tabesh B, Bitterman P, Baker J, Cronin M, et al. 2005. Tumor gene expression and prognosis in breast cancer patients with 10 or more positive lymph nodes. *Clin. Cancer Res.* 11:8623–31
- [248] Cronin M, Pho M, Dutta D, Stephens JC, Shak S, et al. 2004. Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am. J. Pathol.* 164:35–42
- [249] Sparano JA, Paik S. 2008. Development of the 21-gene assay and its application in clinical practice and clinical trials. *J. Clin. Oncol.* 26:721–28

- [250] Kelly CM, Krishnamurthy S, Bianchini G, Litton JK, Gonzalez-Angulo AM, et al. 2010. Utility of oncotype DX risk estimates in clinically intermediate risk hormone receptor-positive, HER2-normal, grade II, lymph node-negative breast cancers. *Cancer* 116:5161–67
- [251] Paik S, Shak S, Tang G, Kim C, Baker J, et al. 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* 351:2817–26
- [252] Van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, et al. 2002. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* 347:1999–2009
- [253] . Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, et al. 2006. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J. Natl. Cancer Inst.* 98:1183–92
- [254] Glas AM, Floore A, Delahaye LJ, Witteveen AT, Pover RC, et al. 2006. Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics* 7:278
- [255] Hornberger J, Chien R. 2010. Meta-analysis of the decision impact of the 21-gene breast cancer Recurrence Score in clinical practice. Presented at Annu. San Antonio Breast Cancer Symp., 33rd, San Antonio, TX
- [256] Newell KA, Asare A, Kirk AD, Gisler TD, Bourcier K, et al. 2010. Identification of a B cell signature associated with renal transplant tolerance in humans. *J. Clin. Investig.* 120:1836–47



- [257] Holleman A, Cheok MH, den Boer ML, Yang W, Veerman AJ, et al. 2004. Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. *N. Engl. J. Med.* 351:533–42
- [258] Lugthart S, Cheok MH, den Boer ML, Yang W, Holleman A, et al. 2005. Identification of genes associated with chemotherapy crossresistance and treatment response in childhood acute lymphoblastic leukemia. *Cancer Cell* 7:375–86
- [259] King CR, Kraus MH, Aaronson SA. 1985. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. *Science* 229:974–76
- [260] Press MF, Pike MC, Chazin VR, Hung G, Udoe JA, et al. 1993. Her-2/neu expression in node-negative breast cancer: direct tissue quantitation by computerized image analysis and association of overexpression with increased risk of recurrent disease. *Cancer Res.* 53:4960–70
- [261] Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, et al. 1999. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J. Clin. Oncol*
- [262] Maughan KL, Lutterbie MA, Ham PS. 2010. Treatment of breast cancer. *Am. Fam. Physician* 81:1339–46

- [263] Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, et al. 2001. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med.* 344:783–92
- [264] Herbst RS, Maddox AM, Rothenberg ML, Small EJ, Rubin EH, et al. 2002. Selective oral epidermal growth factor receptor tyrosine kinase inhibitor ZD1839 is generally well-tolerated and has activity in non-small-cell lung cancer and other solid tumors: results of a phase I trial. *J. Clin. Oncol.* 20:3815–25
- [265] Masui H, Kawamoto T, Sato JD, Wolf B, Sato G, Mendelsohn J. 1984. Growth inhibition of human tumor cells in athymic mice by anti-epidermal growth factor receptor monoclonal antibodies. *Cancer Res.* 44:1002–7
- [266] Ranson M, Hammond LA, Ferry D, Kris M, Tullo A, et al. 2002. ZD1839, a selective oral epidermal growth factor receptor-tyrosine kinase inhibitor, is well tolerated and active in patients with solid, malignant tumors: results of a phase I trial. *J. Clin. Oncol.* 20:2240–50
- [267] Thatcher N, Chang A, Parikh P, Rodrigues Pereira J, Ciuleanu T, et al. 2005. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* 366:1527–37

- [268] Sequist LV, Bell DW, Lynch TJ, Haber DA. 2007. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J. Clin. Oncol.* 25:587–95
- [269] Voora D, McLeod H, Eby C, Gage B. 2005. The pharmacogenetics of coumarin therapy. *Pharmacogenomics* 6:503–13
- [270] Higashi MK, Veenstra DL, Kondo LM, Wittkowsky AK, Srinouanprachanh SL, et al. 2002. Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. *JAMA* 287:1690–98
- [271] Sconce EA, Khan TI, Wynne HA, Avery P, Monkhouse L, et al. 2005. The impact of CYP2C9 and VKORC1 genetic polymorphism and patient characteristics upon warfarin dose requirements: proposal for a new dosing regimen. *Blood* 106:2329–33
- [272] Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, et al. 2009. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N. Engl. J. Med.* 360:753–64
- [273] Driscoll C, Cashion A, Hathaway D, Thompson C, Conley Y, et al. 2006. Blood gene expression profiling in liver transplant recipients with hepatitis C virus and posttransplantation diabetes mellitus. *Transplant. Proc.* 38:3646–48
- [274] Martínez-Llordella M, Puig-Pey I, Orlando G, Ramoni M, Tisone G, et al. 2007. Multiparameter immune profiling of operational tolerance in liver transplantation. *Am. J. Transplant.* 7:309–19

- [275] Starling R, Pham M, Valantine H, Miller L, Eisen H, et al. 2006. Molecular testing in the management of cardiac transplant recipients: initial clinical experience. *J. Heart Lung Transplant.* 25:1389–95
- [276] Deng M, Eisen H, Mehra M, Billingham M, Marboe C, et al. 2006. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. *Am. J. Transplant.* 6:150–60
- [277] Evans R, Williams G, Baron H, Deng M, Eisen H, et al. 2005. The economic implications of noninvasive molecular testing for cardiac allograft rejection. *Am. J. Transplant.* 5:1553–58
- [278] Pham MX, Teuteberg JJ, Kfoury AG, Starling RC, Deng MC, et al. 2010. Gene-expression profiling for rejection surveillance after cardiac transplantation. *N. Engl. J. Med.* 362:1890–900
- [279] Chen B, Chen M, Paisley J, Zaas A, Woods C, et al. 2010. Bayesian inference of the number of factors in gene-expression analysis: application to human virus challenge studies. *BMC Bioinforma.* 11:552
- [280] Zaas AK, Chen M, Varkey J, Veldman T, Hero AO III, et al. 2009. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host Microbe* 6:207–17
- [281] Polonikov AV, Ivanov VP, Solodilova MA. 2009. Genetic variation of genes for xenobiotic-metabolizing enzymes and risk of bronchial asthma: the importance of gene-gene and gene-environment interactions for disease susceptibility. *J. Hum. Genet.* 54:440–49
- [282] Seibold MA, Schwartz DA. 2011. The lung: the natural boundary between nature and nurture. *Annu. Rev. Physiol.* 73:457–78

- [283] Simpson A, John SL, Jury F, Niven R, Woodcock A, et al. 2006. Endotoxin exposure, CD14, and allergic disease: an interaction between genes and the environment. *Am. J. Respir. Crit. CareMed.* 174:386–92
- [284] Fleischmann RD, AdamsMD, White O, Clayton RA, Kirkness EF, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512
- [285] Seib KL, Dougan G, Rappuoli R. 2009. The key role of genomics in modern vaccine and drug design for emerging infectious diseases. *PLoS Genet.* 5:e1000612
- [286] Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, et al. 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 190:6881–93
- [287] Bechah Y, El Karkouri K, Mediannikov O, Leroy Q, Pelletier N, et al. 2010. Genomic, proteomic, and transcriptomic analysis of virulent and avirulent *Rickettsia prowazekii* reveals its adaptive mutation capabilities. *Genome Res.* 20:655–63
- [288] Giuliani MM, Adu-Bobie J, Comanducci M, Arico B, Savino S, et al. 2006. A universal vaccine for serogroup B meningococcus. *Proc. Natl. Acad. Sci. USA* 103:10834–39

- [289] Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, et al. 2007. Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc. Natl. Acad. Sci. USA* 104:9451–56
- [290] Davies J. 2001. In a map for human life, count the microbes, too. *Science* 291:2316
- [291] Savage DC. 1977. Microbial ecology of the gastrointestinal tract. *Annu. Rev. Microbiol.* 31:107–33
- [292] Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, et al. 2006. Metagenomic analysis of the human distal gut microbiome. *Science* 312:1355–59
- [293] Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. 2007. The human microbiome project. *Nature* 449:804–10
- [294] Francois P, Scherl A, Hochstrasser D, Schrenzel J. 2007. Proteomic approach to investigate MRSA. *Methods Mol. Biol.* 391:179–99
- [295] Long WH, Xiao HS, Gu XM, Zhang QH, Yang HJ, et al. 2004. A universal microarray for detection of SARS coronavirus. *J. Virol. Methods* 121:57–63
- [296] Ley RE, Turnbaugh PJ, Klein S, Gordon JI. 2006. Microbial ecology: human gut microbes associated with obesity. *Nature* 444:1022–23
- [297] Peterson DA, Frank DN, Pace NR, Gordon JI. 2008. Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. *Cell Host Microbe* 3:417–27

- [298] Peterson J, Garges S, Giovanni M, McInnes P, Wang L, et al. 2009. The NIH human microbiome project. *Genome Res.* 19:2317–23
- [299] Douglas PS, Ginsburg GS. 2008. Clinical genomic testing: getting it right. *J. Cardiovasc. Transl. Res.* 1:17–20
- [300] Scheuner MT, Sieverding P, Shekelle PG. 2008. Delivery of genomic medicine for common chronic adult diseases: a systematic review. *JAMA* 299:1320–34
- [301] Lapham EV, Kozma C, Weiss JO, Benkendorf JL, Wilson MA. 2000. The gap between practice and genetics education of health professionals: HuGEM survey results. *Genet. Med.* 2:226–31
- [302] Metcalfe S, Hurworth R, Newstead J, Robins R. 2002. Needs assessment study of genetics education for general practitioners in Australia. *Genet. Med.* 4:71–77
- [303] Wilkins-Haug L, Hill LD, Power ML, Holzman GB, Schulkin J. 2000. Gynecologists' training, knowledge, and experiences in genetics: a survey. *Obstet. Gynecol.* 95:421–24
- [304] Goddard KA, Moore C, Ottman D, Szegda KL, Bradley L, Khoury MJ. 2007. Awareness and use of direct-to-consumer nutrigenomic tests, United States, 2006. *Genet. Med.* 9:510–17
- [305] Gramling R, Nash J, Siren K, Eaton C, Culpepper L. 2004. Family physician self-efficacy with screening for inherited cancer risk. *Ann. Fam. Med.* 2:130–32

- [306] FruehFW, Gurwitz D. 2004. From pharmacogenetics to personalized medicine: a vital need for educating health professionals and the community. *Pharmacogenomics* 5:571–79
- [307] Eisenstein EL, Lobach DF, Kawamoto K, Edwards R, Willis JM, et al. 2009. A randomized clinical trial of clinical decision support in a rural community health network serving lower income individuals: study design and baseline characteristics. *Stud. Health Technol. Inform.* 143:220–26
- [308] Kawamoto K, Houlihan CA, Balas EA, Lobach DF. 2005. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 330:765
- [309] Kawamoto K, Lobach DF, Willard HF, Ginsburg GS. 2009. A national clinical decision support infrastructure to enable the widespread and consistent practice of genomic and personalized medicine. *BMC Med. Inform. Decis. Mak.* 9:17
- [310] Apse KA, Biesecker BB, Giardiello FM, Fuller BP, Bernhardt BA. 2004. Perceptions of genetic discrimination among at-risk relatives of colorectal cancer patients. *Genet. Med.* 6:510–166. Archey WB, McEachern KA, Robson M, Offit K, Vaziri SA, et al. 2002. Increased CpG methylation of the estrogen receptor gene in BRCA1-linked estrogen receptor-negative breast cancers. *Oncogene* 21:7034–41
- [311] Hall MA, McEwen JE, Barton JC, Walker AP, Howe EG, et al. 2005. Concerns in a primary care population about genetic discrimination by insurers. *Genet. Med.* 7:311–16



- [312] Klitzman R. 2010. Views of discrimination among individuals confronting genetic disease. *J. Genet. Couns.* 19:68–83
- [313] Haga SB, Willard HF. 2006. Defining the spectrum of genome policy. *Nat. Rev. Genet.* 7:966–72
- [314] Vastag B. 2006. New clinical trials policy at FDA. *Nat. Biotechnol.* 24:1043
- [315] Ginsburg GS, Willard HF. 2009. Genomic and personalized medicine: foundations and applications. *Transl. Res.* 154:277–87
- [316] Fed. Trade Comm. 2006. At-home genetic tests: A healthy dose of skepticism may be the best prescription. <http://www.ftc.gov/bcp/edu/pubs/consumer/health/hea02.shtm>
- [317] Hunter DJ, Khoury MJ, Drazen JM. 2008. Letting the genome out of the bottle: Will we get our wish? *N. Engl. J. Med.* 358:105–7
- [318] McGuire AL, Evans BJ, Caulfield T, Burke W. 2010. Regulating direct-to-consumer personal genome testing. *Science* 330:181–82
- [319] Ginsburg GS, Burke TW, Febbo P. 2008. Centralized biorepositories for genetic and genomic research. *JAMA* 299:1359–61
- [320] Friede A. 2003. National Biospecimen Network Blueprint. Durham, NC: Constella
- [321] Int. Soc. Biol. Environ. Repos. 2005. International Society for Biological and Environmental Repositories: 2004 Annual Meeting. Philadelphia: Taylor & Francis

- [322] HanssonMG, Dillner J, Bartram CR, Carlson JA, Helgesson G. 2006. Should donors be allowed to give broad consent to future biobank research? *Lancet Oncol.* 7:266–69
- [323] Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, et al. 2009. Repeatability of published microarray gene expression analyses. *Nat. Genet.* 41:149–55
- [324] Inst. Med. Comm. Comp. Eff. Res. Prioritization. 2009. Initial National Priorities for Comparative Effectiveness Research. Washington, DC: National Academies. 227 pp
- [325] . Umscheid CA. 2010. Maximizing the clinical utility of comparative effectiveness research. *Clin. Pharmacol. Ther.* 88:876–79
- [326] Garber AM, Tunis SR. 2009. Does comparative-effectiveness research threaten personalized medicine? *N. Engl. J. Med.* 360:1925–27
- [327] Woodward C. 2009. Obama sparks an ideological donnybrook with his push to compare medical treatments. *CMAJ* 180:807
- [328] Marshall E. 2009. Lawsuit challenges legal basis for patenting human genes. *Science* 324:1000–1
- [329] Marshall E. 2010. Cancer gene patents ruled invalid. *Science* 328:153
- [330] Dinan MA, Simmons LA, Snyderman R. 2010. Commentary: Personalized health planning and the Patient Protection and Affordable Care Act: an opportunity for academic medicine to lead health care reform. *Acad. Med.* 85:1665–68

- [331] Whellan DJ, Gaulden L, GattisWA, Granger B, Russell SD, et al. 2001. The benefit of implementing a heart failure disease management program. *Arch. Intern. Med.* 161:2223–28
- [332] Lee SS, Mudaliar A. 2009. Racing forward: the Genomics and PersonalizedMedicine Act. *Science* 323:342
- [333] Goodman J, Walsh V. 2001. The Story of Taxol: Nature and Politics in the Pursuit of an Anti-Cancer Drug. Cambridge: Camb. Univ. Press. 282 pp.
- [334] Burns DK. 2010. Developing pharmacogenetic evidence throughout clinical development. *Clin. Pharmacol. Ther.* 88:867–70
- [335] Westfall JM, Mold J, Fagnan L. 2007. Practice-based research: “Blue Highways” on the NIH roadmap. *JAMA* 297:403–6

## *Serment d'Hippocrate*

*Au moment d'être admis à devenir membre de la profession médicale, je m'engage solennellement à consacrer ma vie au service de l'humanité.*

- *Je traiterai mes maîtres avec le respect et la reconnaissance qui leur sont dus.*
- *Je pratiquerai ma profession avec conscience et dignité. La santé de mes malades sera mon premier but.*
- *Je ne trahirai pas les secrets qui me seront confiés.*
- *Je maintiendrai par tous les moyens en mon pouvoir l'honneur et les nobles traditions de la profession médicale.*
- *Les médecins seront mes frères.*
- *Aucune considération de religion, de nationalité, de race, aucune considération politique et sociale ne s'interposera entre mon devoir et mon patient.*
- *Je maintiendrai le respect de la vie humaine dès la conception.*
- *Même sous la menace, je n'userai pas de mes connaissances médicales d'une façon contraire aux lois de l'humanité.*
- *Je m'y engage librement et sur mon honneur.*

# قسم أبقراط

بسم الله الرحمان الرحيم

أقسم بالله العظيم

في هذه اللحظة التي يتم فيها قبولي عضوا في المهنة الطبية أتعهد علانية:

- بأن أكرس حياتي لخدمة الإنسانية .
- وأن أحترم أساتذتي وأعترف لهم بالجميل الذي يستحقونه .
- وأن أمارس مهنتي بوانزع من ضميري وشر في جاعلا صحة مريض هدي في الأول .
- وأن لا أفشي الأسرار المعهودة إلي .
- وأن أحافظ بكل ما لدي من وسائل على الشرف والتقاليد النبيلة لمهنة الطب .
- وأن أعتبر سائر الأطباء إخوة لي .
- وأن أقوم بواجبي نحو مرضاي بدون أي اعتبار ديني أو وطني أو عرقي أو سياسي أو اجتماعي .
- وأن أحافظ بكل حزم على احترام الحياة الإنسانية منذ نشأتها .
- وأن لا أستعمل معلوماتي الطبية بطريق يضر بحقوق الإنسان مهما لاقيت من تهديد .
- بكل هذا أتعهد عن كامل اختيار ومقسما بالله .

والله على ما أقول شهيد .

جامعة محمد الخامس – الرباط  
كلية الطب والصيدلة بالرباط

أطروحة رقم: 434

سنة : 2017

## تسلسل الحمض النووي والطب الشخصي

### أطروحة

قدمت ونوقشت علانية يوم : .....

من طرف

**السيد: محمد النوالي**

المزوداد في: 12 نونبر 1992 بمراكش

### لنيل شهادة الدكتوراه في الطب

الكلمات الأساسية: علم الجينوم – تسلسل الجيل الجديد – الطب الفردي – الطب الشخصي.

### تحت إشراف اللجنة المكونة من الأساتذة

رئيس

مشرف

أعضاء

السيد: توفيق ضاكة

أستاذ في علم وظائف الأعضاء

السيد: عمر شقيري

أستاذ في علم الأنسجة والأجنة

السيد: فؤاد زويدية

أستاذ في علم التشريح الدقيق

السيد: يونس الرحالي

أستاذ في الصيدلة الغالبية