

Faculté des Sciences et Techniques
Settat

THÈSE DE DOCTORAT

 Pour l'obtention de grade de Docteur en *Sciences et Techniques*

Formation Doctorale: Biologie, Santé et Environnement

Spécialité: Santé et Environnement

Sous le thème

 Analyse bioinformatique du microbiome intestinal et des
micro-organismes contenus dans les eaux de puits identifiés
par séquençage moléculaire
Région rurale Sidi Ayache de la province de Kénitra, Maroc

Présentée par :

Soumaya EL HALAS

Soutenue le: 15 Juillet 2021

A la Faculté des Sciences et Techniques de Settat devant le jury composé de :

Pr. Abderaouf HILALI	PES	Université Hassan I	Président
Pr. Abderazak KHADMAOUI	PES	Université Ibn Tofail	Rapporteur
Pr. Hind AASSILA	PH	Université Hassan I	Rapporteur
Pr. Aziz BAIDANI	PH	Université Hassan I	Examineur
Pr. Lahoucine HILALI	PES	Université Hassan I	Examineur
Pr. Katrine WHITESON	PH	University California Irvine	Examineur
Pr. Mohamed OUHSSINE	PES	Université Ibn Tofail	Co-Directeur de thèse
Pr. Bahya CHARAF	PES	Université Hassan I	Directeur de thèse

Année Universitaire: 2020/2021

Résumé

L'accès à l'eau potable est une composante nécessaire à la santé et au bien-être humain. C'est un défi mondial dans les pays développés et ceux en voie de développement. Le Maroc n'en faisant pas une exception connaît une contamination des nappes phréatiques dans la quasi-totalité des communautés rurales. En effet, les eaux de puits de la zone rurale de Sidi Ayache, dans la province de Kénitra, au Maroc, restent la principale source d'eau potable et d'assainissement des ménages. Cependant, la présence des fosses septiques à proximité des puits ont suscité des inquiétudes quant à la sécurité des eaux souterraines du bassin versant.

Afin de déduire l'impact des fosses septiques sur l'eau, le sol, le végétal, l'animal et l'homme ; une enquête par questionnaire a été établie pour évaluer la qualité des eaux de puits dans la zone rurale d'étude, à travers le témoignage des habitants, pour avoir un retour sur les différents usages des eaux de puits. 413 ménages ont été concernés par cette enquête. Les résultats préliminaires ont révélé que la plupart des gens doutent de la qualité des eaux de puits. Ils affirment que toutes les maladies observées dans la région pourraient provenir de la mauvaise qualité des eaux souterraines de mauvais goût et donc non potables. Il a également été observé que dans certains ménages, il n'y a aucun respect de la distance entre les fosses septiques et les puits. Il y a même un ménage qui a transformé le puits en fosse septique. Ce comportement et d'autres peuvent avoir des effets néfastes sur les composants de l'environnement.

Notre travail de recherche consistait dans un premier temps à analyser la qualité microbiologique des eaux de puits dans la région rurale d'étude, par séquençage moléculaire des microorganismes. Les résultats ont confirmé la contamination par plusieurs virus et bactéries pathogènes (*Legionella*, *Acinetobacter*, *Pseudomonas aeruginosa*, *Mycobacterium*, *Salmonella*, *Vibrio rotiferianus*, *Stenotrophomonas*, *Sphingomonas*, *Flavobacterium*, *Caulobacter*, *Nitrosomonas*, *Nitrospira*, *Streptomyces*...).

Ensuite, pour déterminer l'impact de la consommation des eaux de ces puits sur le microbiome intestinal ; nous avons analysé le microbiome intestinal des sujets marocains qui boivent l'eau contaminée et comment leurs microbiomes se comparent à ceux d'individus vivant en milieu urbain et consommant les eaux potables. Pour ce faire, nous avons étudié les microbiomes intestinaux de trois cohortes au Maroc. Les personnes vivant au Maroc rural, les Marocains vivant en zone urbaine et les sujets migrants d'autres pays africains qui vivent maintenant dans le Maroc urbain. Les trente sujets ruraux consommaient l'eau directement des eaux souterraines.

Nous avons analysé la composition du microbiome par séquençage d'amplicon ARNr 16S/ITS et de « Shotgun ». Les résultats ont indiqué que la diversité n'était pas affectée par les sources d'eau et que les microbiomes intestinaux des individus ruraux étaient plus semblables les uns aux autres que les microbiomes des individus urbains.

MOTS CLES : qualité de l'eau, séquençage de l'ARNr 16S, séquençage des ITS, séquençage « Shotgun », microbiome intestinal, communauté bactérienne, communauté fongique, sujets ruraux et urbains, populations marocaines et africaines.

ملخص

يعد الحصول على مياه الشرب المأمونة مكوناً ضرورياً لصحة الإنسان ورفاهيته. إنه تحد عالمي في البلدان المتقدمة والنامية. المغرب، دون استثناء، يعاني من تلوث المياه الجوفية في جميع المجتمعات الريفية تقريباً. وبالفعل، تظل مياه الآبار في ريف سيدي عياش بإقليم القنيطرة بالمغرب المصدر الرئيسي لمياه الشرب والصرف الصحي المنزلي. ومع ذلك، فإن لوجود خزانات الصرف الصحي بالقرب من الآبار آثار ومخاوف بشأن سلامة المياه الجوفية في مستجمعات المياه.

من أجل استنتاج تأثير خزانات الصرف الصحي على المياه والتربة والنباتات والحيوانات والبشر؛ تم عمل استبيان لتقييم جودة مياه الآبار في منطقة الدراسة الريفية، من خلال شهادة السكان، للحصول على تغذية راجعة حول الاستخدامات المختلفة لمياه الآبار. 413 أسرة معنية بهذا المسح. كشفت النتائج الأولية أن معظم الناس لديهم شكوك حول جودة مياه الآبار. يزعمون أن جميع الأمراض التي لوحظت في المنطقة يمكن أن تأتي من نوعية المياه الجوفية الرديئة التي لا طعم لها وبالتالي غير الصالحة للشرب. وقد لوحظ أيضاً أنه في بعض المنازل لا يوجد احترام للمسافة بين خزانات الصرف الصحي والآبار. حتى أن هناك أسرة حولت البئر إلى خزان للصرف الصحي. يمكن أن يكون لهذا السلوك وغيره آثار ضارة على مكونات البيئة.

يتكون عملنا البحثي أولاً وقبل كل شيء من تحليل الجودة الميكروبيولوجية لمياه الآبار في منطقة الدراسة الريفية، من خلال التسلسل الجزيئي للكائنات الحية الدقيقة. وأكدت النتائج الإصابة ببيكتيريا وفيروسات.

بعد ذلك، لتحديد تأثير مياه الشرب من هذه الآبار على ميكروبيوم الأمعاء؛ قمنا بتحليل الميكروبيوم المعوي للأشخاص المغاربة الذين يشربون المياه الملوثة وكيف يقارن الميكروبيوم لديهم بتلك الخاصة بالآفراد الذين يعيشون في المناطق الحضرية ويستهلكون مياه الشرب. للقيام بذلك، قمنا بدراسة ميكروبيومات الأمعاء لثلاث مجموعات في المغرب. الأشخاص الذين يعيشون في المناطق الريفية بالمغرب، والمغاربة الذين يعيشون في المناطق الحضرية، والمهاجرون من البلدان الأفريقية الأخرى الذين يعيشون الآن في المناطق الحضرية بالمغرب. استهلك الثلاثون شخصاً ريفياً المياه مباشرة من المياه الجوفية.

قمنا بتحليل تكوين الميكروبيوم بواسطة تسلسل أمبليكون rRNA S/ITS16 وتسلسل « Shotgun » وأشارت النتائج إلى أن التنوع لم يتأثر بمصادر المياه وأن الميكروبات المعوية للآفراد الريفيين كانت أكثر تشابهاً مع بعضها البعض من الميكروبيومات للآفراد الحضريين. الكلمات الرئيسية: جودة المياه، تسلسل S16، تسلسل ITS، تسلسل « Shotgun »، ميكروبيوم الأمعاء، المجتمع البكتيري، المجتمع الفطري، المواضيع الريفية والحضرية، السكان المغاربة والآفارقة.

 Soumaya
EL HALAS

 Analyse bioinformatique du microbiome intestinal et des micro-organismes contenus
dans les eaux de puits identifiés par séquençage moléculaire
Région rurale Sidi Ayache de la province de Kénitra, Maroc

 BSE
2020/2021,

Thèse de doctorat en Biologie, Santé et Environnement

**Analyse bioinformatique du microbiome intestinal et des micro-organismes
contenus dans les eaux de puits, identifiés par séquençage moléculaire
Région rurale Sidi Ayache de la province de Kénitra, Maroc**

Soumaya EL HALAS

Sous la direction du Pr. Bahia CHARAF, Pr. Mohamed OUHSSINE et Pr. Katrine WHITESON

Présentée et soutenue publiquement le 15 Juillet 2021

Composition du Jury :

- Pr. Bahia CHARAF, Université Hassan I (Encadrante)
- Pr. Mohamed OUHSSINE, Université Ibn Tofail (Co-encadrant)
- Pr. Abderaouf HILALI, Université Hassan I (Président)
- Pr. Hind AASSILA, Université Hassan I (Rapporteur)
- Pr. Abderazak KHADMAOUI, Université Ibn Tofail (Rapporteur)
- Pr. Lahoucine HILALI, Université Hassan I (Examineur)
- Pr. Aziz BAIDANI, Université Hassan I (Examineur)
- Pr. Katrine WHITESON, University of California Irvine (American Advisor)
- Ms. Rebecca B. GEFFNER, MACECE Executive Director (Invitée)
- Ms. Hafsa EL BASTAMI, MACECE Deputy Executive Director (Invitée)

Déclaration d'originalité

L'Université Hassan Premier n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse ; ces opinions doivent être considérées comme propres à leur auteur.

Remerciement

Je souhaite exprimer en premier lieu ma profonde gratitude à ma très chère encadrante, professeur **Bahya CHARAF**, pour m'avoir permis de travailler sur ce projet passionnant. Ses connaissances approfondies et son bon cœur m'ont aidé à surmonter plusieurs obstacles qui auraient été beaucoup plus difficiles à affronter.

J'adresse de chaleureux remerciements à mon co-encadrant, professeur **Mohamed OUHSSINE** pour ses conseils avisés et son écoute qui ont été l'un des facteurs les plus influents et les plus motivants de mon travail de thèse de doctorat.

Je remercie le professeur **Katrine WHITESON** de m'avoir accueilli dans son laboratoire pour effectuer les analyses de séquençage dans les meilleures conditions.

Un grand merci au professeur **Abderaouf HILALI** pour l'honneur qu'il m'a fait en présidant le jury de cette thèse, je lui exprime toute ma profonde reconnaissance.

Je voudrais remercier les rapporteurs de cette thèse, professeur **Abderazak KHADMAOUI** et professeur **Hind AASSILA** pour le temps consacré à la lecture de ce travail ainsi que pour les commentaires m'ayant permis de l'améliorer.

J'associe à ces remerciements professeur **Lahoucine HILALI** et professeur **Aziz BAIDANI** pour avoir accepté d'examiner mon travail. J'en suis reconnaissante.

Mon immense gratitude va également à **Ms. Rebecca B. GEFFNER** et **Ms. Hafsa EL BASTAMI** pour m'avoir honorée de leur présence.

Je tiens également à remercier le **CNRST** et **Fulbright MACECE** de m'avoir financé pendant ma période de recherche.

Ce fut également un véritable privilège et un honneur de travailler avec les membres du laboratoire de biologie moléculaire **Whiteson lab**: Matthew, Julio, Andrew, Jason, Tara, Joann, Stephen, Stephan, Bretton, Clark and Cyril. Leur aide a été déterminante dans ma réussite mentale et académique à l'Université de California Irvine. Je les remercie pour les nombreux moments formidables que nous avons partagés à l'intérieur et à l'extérieur du laboratoire.

En outre, je remercie **Abdelali BALAHBIB** de m'avoir aidé à collecter les échantillons.

Enfin, je tiens à remercier mon merveilleux mari **Badreddine** pour sa patience et son soutien.

Résumé

RESUME :

L'accès à l'eau potable est une composante nécessaire à la santé et au bien-être humain. C'est un défi mondial dans les pays développés et ceux en voie de développement. Le Maroc n'en faisant pas une exception connaît une contamination des nappes phréatiques dans la quasi-totalité des communautés rurales. En effet, les eaux de puits de la zone rurale de Sidi Ayache, dans la province de Kénitra, au Maroc, restent la principale source d'eau potable et d'assainissement des ménages. Cependant, la présence des fosses septiques à proximité des puits ont suscité des inquiétudes quant à la sécurité des eaux souterraines du bassin versant.

Afin de déduire l'impact des fosses septiques sur l'eau, le sol, le végétal, l'animal et l'homme ; une enquête par questionnaire a été établie pour évaluer la qualité des eaux de puits dans la zone rurale d'étude, à travers le témoignage des habitants, pour avoir un retour sur les différents usages des eaux de puits. 413 ménages ont été concernés par cette enquête. Les résultats préliminaires ont révélé que la plupart des gens doutent de la qualité des eaux de puits. Ils affirment que toutes les maladies observées dans la région pourraient provenir de la mauvaise qualité des eaux souterraines de mauvais goût et donc non potables. Il a également été observé que dans certains ménages, il n'y a aucun respect de la distance entre les fosses septiques et les puits. Il y a même un ménage qui a transformé le puits en fosse septique. Ce comportement et d'autres peuvent avoir des effets néfastes sur les composants de l'environnement.

Notre travail de recherche consistait dans un premier temps à analyser la qualité microbiologique des eaux de puits dans la région rurale d'étude, par séquençage moléculaire des microorganismes. Les résultats ont confirmé la contamination par plusieurs virus et bactéries pathogènes (*Legionella*, *Acinetobacter*, *Pseudomonas_aeruginosa*, *Mycobacterium*, *Salmonella*, *Vibrio_rotiferianus*, *Stenotrophomonas*, *Sphingomonas*, *Flavobacterium*, *Caulobacter*, *Nitrosomonas*, *Nitrospira*, *Streptomyces*...).

Ensuite, pour déterminer l'impact de la consommation des eaux de ces puits sur le microbiome intestinal ; nous avons analysé le microbiome intestinal des sujets marocains qui boivent l'eau contaminée et comment leurs microbiomes se comparent à ceux d'individus vivant en milieu urbain et consommant les eaux potables. Pour ce faire, nous avons étudié les microbiomes intestinaux de trois cohortes au Maroc. Les personnes vivant au Maroc rural, les Marocains vivant en zone urbaine et les sujets migrants d'autres pays africains qui vivent maintenant dans le Maroc urbain. Les trente sujets ruraux consommaient l'eau directement des eaux souterraines.

Nous avons analysé la composition du microbiome par séquençage d'amplicon ARNr 16S/ITS et de « Shotgun ». Les résultats ont indiqué que la diversité n'était pas affectée par les sources d'eau et que les microbiomes intestinaux des individus ruraux étaient plus semblables les uns aux autres que les microbiomes des individus urbains.

MOTS CLES : qualité de l'eau, séquençage de l'ARNr 16S, séquençage des ITS, séquençage « Shotgun », microbiome intestinal, communauté bactérienne, communauté fongique, sujets ruraux et urbains, populations marocaines et africaines.

ABSTRACT:

Access to safe drinking water is a necessary component of human health and well-being. It is a global challenge in developed and developing countries. Morocco, not making an exception, is experiencing groundwater contamination in almost all rural communities. Indeed, well water in the rural area of Sidi Ayache, in the province of Kenitra, in Morocco, remains the main source of drinking water and household sanitation. However, the presence of septic tanks near the wells has raised concerns about the safety of groundwater in the watershed.

In order to deduce the impact of septic tanks on water, soil, plants, animals and humans; a questionnaire survey was established to assess the quality of well water in the rural study area, through the testimony of residents, to obtain feedback on the different uses of well water. 413 households were concerned by this survey. Preliminary results revealed that most people have doubts about the quality of well water. They claim that all the diseases observed in the region could come from the poor quality of water and

therefore not drinkable. It has also been observed that in some households there is no respect for the distance between septic tanks and wells. There is even a household who transformed the well into a septic tank. This behavior and others can have adverse effects on components of the environment.

Our research work consisted first of all in analyzing the microbiological quality of well water in the rural study area, by molecular sequencing of microorganisms. The results confirmed the contamination by some pathogenic bacteria and viruses.

Then, to determine the impact of drinking water from these wells on the gut microbiome; We analyzed the intestinal microbiome of Moroccan subjects who drink contaminated water and how their microbiomes compare to those of individuals living in urban areas and consuming drinking water. To do this, we studied the gut microbiomes of three cohorts in Morocco. People living in rural Morocco, Moroccans living in urban areas and migrant subjects from other African countries who now live in urban Morocco. The thirty rural subjects consumed water directly from groundwater.

We analyzed the composition of the microbiome by 16S rRNA / ITS amplicon sequencing and "Shotgun" sequencing. The results indicated that diversity was not affected by water sources and that the gut microbiomes of rural individuals were more similar to each other than the microbiomes of urban individuals.

KEYWORDS: water quality, 16S rRNA sequencing, ITS sequencing, shotgun sequencing, gut microbiome, bacterial community, fungal community , rural and urban subjects, Moroccan and African populations.

نبذة مختصرة:

يعد الحصول على مياه الشرب المأمونة مكونًا ضروريًا لصحة الإنسان ورفاهيته. إنه تحد عالمي في البلدان المتقدمة والنامية. المغرب، دون استثناء، يعاني من تلوث المياه الجوفية في جميع المجتمعات الريفية تقريبًا. وبالفعل، تظل مياه الآبار في ريف سيدي عياش بإقليم القنيطرة بالمغرب المصدر الرئيسي لمياه الشرب والصرف الصحي المنزلي. ومع ذلك، فإن لوجود خزانات الصرف الصحي بالقرب من الآبار آثار ومخاوف بشأن سلامة المياه الجوفية في مستجمعات المياه.

من أجل استنتاج تأثير خزانات الصرف الصحي على المياه والتربة والنباتات والحيوانات والبشر؛ تم عمل استبيان لتقييم جودة مياه الآبار في منطقة الدراسة الريفية، من خلال شهادة السكان، للحصول على تغذية راجعة حول الاستخدامات المختلفة لمياه الآبار. 413 أسرة معنية بهذا المسح. كشفت النتائج الأولية أن معظم الناس لديهم شكوك حول جودة مياه الآبار. يزعمون أن جميع الأمراض التي لوحظت في المنطقة يمكن أن تأتي من نوعية المياه الجوفية الرديئة التي لا طعم لها وبالتالي غير الصالحة للشرب. وقد لوحظ أيضاً أنه في بعض المنازل لا يوجد احترام للمسافة بين خزانات الصرف الصحي والآبار. حتى أن هناك أسرة حولت البئر إلى خزان للصرف الصحي. يمكن أن يكون لهذا السلوك وغيره آثار ضارة على مكونات البيئة.

يتكون عملنا البحثي أولاً وقبل كل شيء من تحليل الجودة الميكروبيولوجية لمياه الآبار في منطقة الدراسة الريفية، من خلال التسلسل الجزيئي للكائنات الحية الدقيقة. وأكدت النتائج الإصابة ببكتيريات وفيروسات.

بعد ذلك، لتحديد تأثير مياه الشرب من هذه الآبار على ميكروبيوم الأمعاء؛ قمنا بتحليل الميكروبيوم المعوي للأشخاص المغاربة الذين يشربون المياه الملوثة وكيف يقارن الميكروبيوم لديهم بتلك الخاصة بالأفراد الذين يعيشون في المناطق الحضرية ويستهلكون مياه الشرب. للقيام بذلك، قمنا بدراسة ميكروبيومات الأمعاء لثلاث مجموعات في المغرب. الأشخاص الذين يعيشون في المناطق الريفية بالمغرب، والمغاربة الذين يعيشون في المناطق الحضرية، والمهاجرون من البلدان الأفريقية الأخرى الذين يعيشون الآن في المناطق الحضرية بالمغرب. استهلك الثلاثون شخصاً ريفياً المياه مباشرة من المياه الجوفية.

قمنا بتحليل تكوين الميكروبيوم بواسطة تسلسل أمبليكون 16S/ITS rRNA وتسلسل « Shotgun » وأشارت النتائج إلى أن التنوع لم يتأثر بمصادر المياه وأن الميكروبات المعوية للأفراد الريفيين كانت أكثر تشابهاً مع بعضها البعض من الميكروبيومات للأفراد الحضريين.

الكلمات الرئيسية: جودة المياه، تسلسل 16S، تسلسل ITS، تسلسل « Shotgun »، ميكروبيوم الأمعاء، المجتمع البكتيري، المجتمع الفطري، المواضيع الريفية والحضرية، السكان المغاربة والأفارقة.

Sommaire

Déclaration d'originalité	2
Remerciement	3
Résumé	4
Sommaire	8
Liste des communications et publications	11
Liste des abréviations	12
Liste des figures	13
Liste des tableaux	16
Liste des annexes	17
Introduction	18
Chapitre 1 : Revue bibliographique	20
1.1. Eau et Santé	20
1.1.1. L'eau: une exigence universelle pour la survie	20
1.1.2. Politique et réglementation au Maroc	21
1.1.3. Source de pollution de la nappe phréatique.....	21
1.1.4. Maladies transmissibles associées à l'eau.....	22
1.2. Santé et Microbiote	23
1.2.1. Microbiote intestinal	23
1.2.1.1. Acquisition du microbiote intestinal	25
1.2.1.2. Diversification du microbiote intestinal	26
1.2.2. Microbiote intestinal sain	26
1.2.3. Microbiote et maladies.....	26
1.3. Techniques de séquençage	32
1.3.1. Séquençage Sanger.....	32
1.3.2. Séquençage de nouvelle génération NGS	32
1.3.2.1. Séquençage rRNA 16S/ITS	37

1.3.2.2. Séquençage « Shotgun ».....	38
Chapitre 2 : Matériel et Méthodes	39
2.1. Zone d'étude « Sidi Ayache ».....	39
2.2. Enquête via questionnaires	40
2.3. Collecte des échantillons des eaux de puits	41
2.4. Collecte des échantillons des selles des habitants.....	41
2.5. Traitement des échantillons des eaux de puits	42
2.6. Traitement des échantillons des selles	43
2.7. Extractions de l'ADN des microorganismes par ZymoKits ...	43
2.8. Quantification de l'ADN par Quant-iT Pico Green.....	46
2.8.1. Le logiciel Gen5 pour la lecture de microplaques et l'analyse des données.....	46
2.8.2. Intégration au système de détection BioTek	47
2.9. Détermination de la longueur de l'ADN	47
2.9.1. Gel d'agarose:	47
2.9.2. Bioanalyseur:	47
2.10. Amplification par Réaction en Chaîne par Polymérase PCR des bactéries 16S et champignons ITS	48
2.11. Protocole d'alimentation des amorces résiduelles	49
2.11.1. Nextera Library Prep	49
2.11.2. AMPure XP beads.....	50
2.12. Séquençage de l'ARNr 16S/ITS	50
2.13. Séquençage métagénomique « Shotgun ».....	52
2.14. Analyse bioinformatique	53
Chapitre 3 : Résultats.....	57
3.1. Résultats de l'enquête par questionnaires.....	57
3.1.1. Sexe de la population interrogée.....	Error! Bookmark not defined.
3.1.2. Puits : couvert/non couvert.....	57
3.1.3. Profondeur des puits.....	57
3.1.4. Type de construction des fosses septiques	58

3.1.5.	Profondeur des fosses septiques	58
3.1.6.	Distance entre fosse septique et puits	59
3.1.7.	Appréciation de l'eau de puits	59
3.1.8.	Changement au goût de l'eau des puits	60
3.1.9.	Problèmes de santé.....	60
3.1.10.	La cause des maladies	60
3.2.	Résultats des analyses des échantillons des eaux de puits	61
3.2.1.	Quantification de l'ADN des microorganismes dans les filtres et la matière lyophilisée	61
3.2.2.	Traitement des résultats de séquençage « Shotgun ».....	62
3.2.2.1.	Analyse d'abondance relative.....	62
3.2.2.2.	Analyse alpha diversité Shannon.....	77
3.2.2.3.	Analyse bêta diversité NMDS.....	77
3.3.	Résultats des analyses du microbiome intestinal	78
3.3.1.	Quantification de l'ADN extraits des échantillons des selles.....	78
3.3.1.1.	Résultats du logiciel Gen5.....	78
3.3.1.2.	Résultats du Gel d'agarose	80
3.3.1.3.	Traces de bio-analyseur	81
3.3.2.	Résultats du séquençage « Shotgun ».....	82
3.3.2.1.	Analyse d'abondance relative	82
3.3.2.2.	Analyse des champignons et d'un bactériophage.....	84
3.3.3.	Résultats du séquençage de l'ARNr des bactéries 16S et des champignons ITS	85
3.3.3.1.	Analyse d'abondance relative 16S/ITS par individu	85
3.3.3.2.	Analyse d'abondance relative 16S/ITS par cohorte.....	86
3.3.3.3.	Analyse alpha diversité 16S/ITS.....	88
3.3.3.4.	Analyse beta diversité 16S/ITS	89
Discussion.....	91
Conclusion	96
Bibliographie	97
Annexes	105

Liste des communications et publications

- 1) S. El Halas, B. Charaf, M. Ouhssine, The Impact of Septic Tanks on the Groundwater: Survey Results in a Rural Area of Morocco, Africa, Poll Res. 39 (November Suppl. Issue) : S192-S195 (2020), Copyright © EM International, ISSN 0257–8050
- 2) Soumaya El Halas, Matthew Gargus, Julio Avelar-Barragan, Andrew Oliver, Jason Rothman, Bahya Charaf, Abdelaali Balahbib, Mohammed Ouhssine, Katrine Whiteson, SIGNIFICANT DIFFERENCES IN HUMAN GUT BACTERIAL AND FUNGAL COMMUNITIES IN PEOPLE LIVING IN URBAN AND RURAL AREAS IN MOROCCO, the 5th Annual Gut Microbiome Conference, 21-22 September 2019 at Huntington Beach, California, USA.
- 3) EL HALAS S. , Pr. CHARAF B. Pr. OUHSSINE M. , IMPACT OF SEPTIC TANK ON THE GROUNDWATER : SURVEY ON LAHMASSISS REGION, KENITRA, MOROCCO , the 7th Global Conference on Global Warming, 24-28 June 2018 at Izmir, Turkey.
- 4) EL HALAS S. , Pr. CHARAF B. Pr. OUHSSINE M. , IMPACT OF SEPTIC TANK ON THE GROUNDWATER : SURVEY ON LAHMASSISS REGION, KENITRA, MOROCCO , the 9th International Conference on Sustainable Water Resources Management , 18-20 July 2017 at Prague, Czech Republic.
- 5) EL HALAS S. , Pr. CHARAF B. Pr. OUHSSINE M. , ETUDE D'IMPACT DES FOSSES SEPTIQUES SUR LA NAPPE PHREATIQUE : RESULTATS DES ANALYSES SUR LES EAUX DES PUIITS DE LA REGION LAHMASSISS, KENITRA, MAROC , Symposium National sur les premières journées de la valorisation des ressources naturelles, 08 juin 2017 , Faculté des Sciences de université Ibn Tofail, Kénitra, Maroc.
- 6) EL HALAS S. , Pr. CHARAF B. Pr. OUHSSINE M. , ETUDE D'IMPACT DES FOSSES SEPTIQUES SUR LA NAPPE PHREATIQUE : ENQUÊTE SUR LA REGION LAHMASSISS, KENITRA, MAROC, 3ème édition des journées doctorales sciences et techniques, 11 et 12 Mai 2017 , Faculté des Sciences et Techniques FST Errachidia de l'université Moulay Ismail sous le thème : La recherche Scientifique & la Valorisation des Potentialités Régionales.

Liste des abréviations

ADN	<i>Acide DésoxyriboNucléique</i>
ADNdb.....	<i>ADN double-brin</i>
ARN.....	<i>Acide RiboNucléique</i>
EDTA.....	<i>ÉthylèneDiamineTétraAcétique</i>
GPF.....	<i>Green Fluorescent Protein</i>
TAE	<i>Tris, Acétate, EDTA</i>
UPW	<i>Ultrapure Water Treatment</i>

Liste des figures

Figure 1: Relation Microbiote-Santé.....	29
Figure 2: Schéma de préparation de la bibliothèque Illumina. Les adaptateurs sont ligaturés à des molécules d'ADN fragmentées au hasard.	34
Figure 3: Schéma de génération de cluster. Les fragments monocaténares se lient à la Flowcell. Les grappes sont ensuite générées par amplification pontée. Les clusters finaux contiennent jusqu'à 1 000 copies du fragment initial.	35
Figure 4: Schéma de séquençage d'Illumina par synthèse. Quatre désoxynucléosides triphosphates (dNTP) marqués différemment, des amorces et une ADN polymérase sont ajoutés et le dNTP approprié est ajouté à la chaîne d'acide nucléique.	35
Figure 5: Schéma de séquençage par paires. En mode apparié, chaque fragment d'ADN est lu aux deux extrémités.	36
Figure 6 : Description géographique de la zone d'étude[71].....	39
Figure 7 : Nappe Phréatique[72].....	40
Figure 8: Traitement des échantillons d'eau.....	42
Figure 9: Traitement des échantillons de selles	43
Figure 10: Les étapes des extractions de l'ADN	44
Figure 11: Tube de lyse contenant 150mg de matière fécale et 750 µl de solution de lyse	44
Figure 12: L'ADN résultant adapté à la PCR.....	45
Figure 13 : Exemple de lecture et analyse de microplaques sur le logiciel Gen5.....	46
Figure 14 : Le protocole de Beckman.....	50
Figure 15: Sexe de la population interrogée	Error! Bookmark not defined.
Figure 16: Couverture ou non des puits.....	57
Figure 17: Profondeur des puits	58
Figure 18: Type de construction des fosses septiques	58
Figure 19: Profondeur des fosses septiques	59
Figure 20: Distance entre fosse septique et puits.....	59
Figure 21: Appréciation de l'eau de puits	59
Figure 22: Existence de changement au goût de l'eau des puits	60

Figure 23: Existence de problèmes de santé	60
Figure 24: La cause des maladies.....	61
Figure 25: Les abondances relatives des 10 principales espèces dans les échantillons déterminées par séquençage « Shotgun ».....	63
Figure 26: Shannon alpha diversité	77
Figure 27: NMDS bêta diversité	78
Figure 28: Image des résultats du gel de six échantillons	81
Figure 29: Les traces du bio-analyseur	82
Figure 30: 10 premières classifications taxonomiques des communautés dans les échantillons	83
Figure 31: Abondances de champignons et d'un bactériophage pour les trois cohortes déterminées par séquençage « Shotgun ».....	84
Figure 32: Distribution des abondances relatives par cohorte des 10 principales espèces bactériennes significatives	85
Figure 33: Distribution des abondances relatives par cohorte des 10 principaux champignons.....	86
Figure 34: Abondance relative au niveau du phylum	86
Figure 35 : a) Abondances relatives moyennes des 10 genres bactériens les plus abondants dans les trois cohortes. «Other» indique les genres qui ne figuraient pas dans le top 10 des plus abondants. b) Un arbre 16S montrant la diversité taxonomique au sein de la cohorte rurale. L'abondance relative de chaque taxon est indiquée par la couleur comme indiqué dans la clé.	87
Figure 36: a) Abondances relatives moyennes des 10 genres de champignons les plus abondants dans les trois cohortes. «Other» indique les genres qui ne figuraient pas dans le top 10 des plus abondants. b) Un arbre ITS montrant la diversité taxonomique au sein de la cohorte rurale.....	88
Figure 37: Diagrammes en boîtes montrant la diversité alpha de chaque cohorte estimée à l'aide de l'indice de Shannon. a) montre des estimations pour les amplicons 16S et b) montre des estimations pour les amplicons ITS. Chaque case montre la médiane (ligne horizontale) et l'intervalle interquartile. Les différences n'étaient pas statistiquement significatives (ANOVA; p _{16S} = 0,387; p _{ITS} = 0,527).....	88

Figure 38: Analyse alpha diversité de 16S à gauche et ITS à droite par le Microbiome Anayst Website	89
Figure 39: Tracés NMDS des différences de Bray-Curtis pour a) les amplicons 16S et b) les amplicons ITS. Les ellipses représentent les intervalles de confiance à 95% pour chaque cohorte	90
Figure 40: Analyse alpha diversité de 16S à gauche et ITS à droite par le Microbiome Anayst Website	90

Liste des tableaux

Tableau 1: Modifications du microbiote intestinal associées à la maladie.....	30
Tableau 2 : Programme de PCR sur la machine BioRad C1000	49
Tableau 3: Quantification de l'ADN des échantillons des eaux de puits	62
Tableau 4: Détection de <i>Vibrio_rotiferianus</i>	63
Tableau 5: Détection de <i>Salmonella</i>	64
Tableau 6: Détection des espèces du gène <i>Mycobacterium</i>	65
Tableau 7: Détection de <i>Legionella</i>	66
Tableau 8: Détection de <i>Bacillus</i>	67
Tableau 9: Détection de <i>Pseudomonas_aeruginosa</i>	68
Tableau 10: Détection de deux espèces du gène <i>Acinetobacter</i>	69
Tableau 11: Détection du <i>Stenotrophomonas</i>	70
Tableau 12: Détection des espèces de <i>Sphingomonas</i>	71
Tableau 13: Détection des espèces de <i>Flavobacterium</i>	72
Tableau 14: Détection de <i>Caulobacter_vibrioides</i>	74
Tableau 15: Détection de <i>Nitrosomonas</i> et <i>Nitrospira</i>	75
Tableau 16: Détection des <i>Streptomyces</i>	76
Tableau 17 : Tableau des concentrations d'ADN dans les échantillons	78

Liste des annexes

Annexe 1: Protocole du PicoGreen.....	105
Annexe 2: Protocole de Beckman.....	106
Annexe 3: Metadata.....	107
Annexe 4: Barcodes 16S.....	109
Annexe 5: Barcodes ITS.....	111
Annexe 6: Résultats du Gel.....	113
Annexe 7: Résultats de Shotgun.....	117

Introduction

Le Maroc est confronté à un défi de plus en plus difficile pour fournir de l'eau à la bonne quantité et qualité pour répondre à la demande en raison de l'urbanisation et de la croissance rapide de la population [1]. La plupart des régions rurales ne disposent pas d'installations sanitaires adéquates, ce qui oblige les habitants à recourir à des pratiques dangereuses, notamment l'utilisation de toilettes à fosse [2]. Cette pratique, compte tenu des conditions environnementales qui prévalent, peut polluer la nappe phréatique et empêcher le public d'accéder à une eau potable.

L'OMS a estimé que les maladies diarrhéiques, souvent résultant de la consommation d'eau contaminée microbiologiquement, causent environ un milliard de maladies et 2,2 millions de décès par an dans le monde [3]. Le Maroc n'en fait pas exception !

Les conditions dans les zones rurales du Maroc engendrent de mauvaises pratiques d'assainissement [4], principalement en raison d'un manque d'investissement public. Face à un assainissement médiocre, certaines zones rurales du Maroc n'ont pas accès à l'eau potable, et obtiennent plutôt de l'eau directement du sol à travers des puits contaminés par des fosses septiques à proximité. La majorité des puits ne sont pas construits à l'aide de méthodes standardisées ou ne sont pas entretenus correctement. Boire de ces puits peut causer des effets néfastes importants sur la santé communautaire et l'environnement.

La mauvaise qualité de l'eau est un problème courant dans les régions rurales au Maroc. Sidi Ayache en est un exemple. Les habitants de cette région rurale n'ont pas accès à une eau potable et n'ont pas accès à des installations sanitaires sûres fournies et réglementées par des agences gérées par le gouvernement. Ils ont recours à d'autres sources d'eau, y compris les eaux de puits, qui restent mal réglementés, en grande partie non traités et potentiellement contaminés par des micropolluants et agents pathogènes.

La région rurale Sidi Ayache se développe rapidement et sa population ne dispose pas de réseau de distribution d'eau potable ni de réseau d'assainissement (figure

1). Cette population utilise l'eau provenant de la nappe phréatique à des fins de boisson et d'utilisation ménagères.

À ce jour, il n'y a pas eu de suivi régulier de la qualité de l'eau des puits de cette région rurale, ni de programmes de recherche menés pour évaluer les qualités microbiennes et physico-chimiques des ressources en eau.

Boire à partir de puits contaminés pourrait affecter probablement le microbiome intestinal de chaque individu, ce qui est particulièrement important car différentes compositions de microbiome ont été liées à certaines maladies (Cancers, diabètes, obésité, ...).

L'objectif de notre travail de recherche, consiste dans un premier temps à effectuer une enquête auprès des habitants de Sidi Ayache pour évaluer la qualité des eaux de puits. Et dans un deuxième temps, à identifier par séquençage moléculaire les microorganismes dans 10 puits d'eau. Ensuite, nous avons analysé par séquençage le microbiome intestinal de trois cohortes urbaines et rurales.

Chapitre 1 : Revue bibliographique

1.1. Eau et Santé

1.1.1. L'eau: une exigence universelle pour la survie

La survie de l'humanité sur terre dépend de trois ressources de base: l'eau, l'air et le sol. L'eau est sans aucun doute la ressource naturelle la plus essentielle qui existe sur notre planète. L'eau douce ne représente que 2,5% des ressources en eau de la terre, dont seulement 1,2% est facilement disponible comme eau de surface à des fins de boisson et à usages multiples [5]. La croissance de la population humaine, l'urbanisation rapide et les changements rapides de l'utilisation des terres ont rendu les écosystèmes naturels de plus en plus vulnérables aux changements environnementaux et aux événements extrêmes [6].

Le droit de l'homme à garantir l'accès à l'eau potable et à l'assainissement est essentiel à la durabilité d'une bonne santé et d'une bonne qualité de vie. Malheureusement, le droit de l'homme à l'eau potable et à l'assainissement n'a pas été explicitement reconnu comme un droit de l'homme autonome dans les traités internationaux [7]. Cependant, le droit international des droits de l'homme implique des obligations spécifiques liées à l'accès à l'eau potable, qui oblige les états à assurer à tous les citoyens l'accès à une quantité suffisante d'eau potable pour les usages personnels et domestiques. En outre, ces obligations exigent également que les pays garantissent progressivement l'accès à un assainissement adéquat en tant qu'élément fondamental de la vie humaine, et comme garantie de la qualité des ressources en eau potable [7], [8].

Le problème est encore exacerbé par les défis sociaux et environnementaux de l'accélération de l'urbanisation, le changement climatique, l'augmentation de la pollution et l'épuisement des ressources en eau.

Le changement climatique menace de plus en plus le bien-être humain en raison d'impacts négatifs sur de nombreuses facettes de l'environnement, notamment la

perte de biodiversité, la dégradation de la qualité de l'eau et la propagation de maladies [9].

1.1.2. Politique et réglementation au Maroc

Le Maroc a rédigé des objectifs nationaux ambitieux dans sa politique d'eau, d'assainissement et d'hygiène (WASH) 2015-2030. La loi sur les contaminants environnementaux contient des procédures et des permis de pesticides. Il définit de multiples objets tels que; promouvoir la gestion judicieuse des ressources naturelles, en protégeant l'environnement tout en permettant un développement d'une manière qui améliore la qualité de la vie et maintient les processus écologiques dont dépend la vie [10]; appliquer une approche de précaution à l'évaluation des risques de dommages environnementaux et veiller à ce que tous les aspects de la qualité de l'environnement affectés par les dommages environnementaux soient pris en compte dans les décisions relatives à l'environnement [11].

Cependant, le pays n'a jusqu'à présent pas été en mesure de relever le défi du développement. Le gouvernement est en retard par rapport à une approche globale du développement durable qui est cruciale pour accroître la compréhension et le respect des droits de l'homme et leur valeur pour fournir un cadre d'action clair et internationalement accepté. Selon le rapport du Programme des Nations Unies pour le développement (PNUD), les causes de la crise actuelle de l'eau et de l'assainissement peuvent être attribuées à la pauvreté et aux luttes de pouvoir entre les différents niveaux au sein des agences gérées par le gouvernement et de la communauté [12].

1.1.3. Source de pollution de la nappe phréatique

La nappe phréatique est menacée par plusieurs sources de contamination, notamment des sources chimiques et microbiennes. Les rejets de déchets domestiques sont les principales causes de pollution de l'eau. La principale source de pollution de l'eau des puits comprend:

- a) Déversement aléatoire de déchets.
- b) Non respect de distances entre fosses septiques et puits.

- c) Ruissellement dû au mauvais assainissement.
- d) Fosses non construites dans les normes (sable au lieu du gravier).
- e) Effets des changements géologiques et géomorphologiques dans le bassin.

L'augmentation de la population crée de nombreux problèmes, notamment en tant que facteur majeur de pollution de l'eau. Elle entraîne une augmentation de la production des excréments humains et donc accélère la contamination de l'eau.

Dans l'eau contaminée, on trouve un grand nombre de microorganismes nocifs pour la santé humaine. Le gouvernement est incapable de répondre aux besoins essentiels des citoyens en raison du nombre croissant de population. Les installations sanitaires sont disponibles dans les zones urbaines mais font défaut dans les zones rurales. La surpopulation et l'eau potable insalubre sont des problèmes de santé majeurs dans les zones rurales [13].

La qualité de l'eau est une mesure de l'état de l'eau par rapport à tout besoin ou objectif humain [14]. La qualité de l'eau est largement déterminée à la fois par les processus naturels, y compris la lithologie du bassin, les apports atmosphériques et les conditions climatiques, et par les apports anthropiques. Les mesures de la qualité de l'eau comprennent une variété de paramètres physiques, chimiques et biologiques (Température, pH, alcalinité, conductivité, TSD Total Solides Dissous, turbidité, dureté et métaux lourds) [15].

1.1.4. Maladies transmissibles associées à l'eau

Les problèmes de santé transmissibles associés à l'eau comprennent la transmission d'infections respiratoires, de maladies d'origine alimentaire et des maladies transmises par le sang. Par exemple, plusieurs épidémies telle que celle de choléra en 1990, ont été liées à l'eau [16].

Bien que la gastro-entérite et les infections diarrhéiques représentent une proportion importante des diagnostics parmi les patients admis dans les hôpitaux des régions rurales, on en sait peu sur la nature et l'ampleur de ce problème [17].

En plus de choléra, il existe d'autres maladies infectieuses qui se propagent par une eau polluée, comme la fièvre typhoïde, la diarrhée, les vomissements, les problèmes de peau et de reins.

1.2. Santé et Microbiote

1.2.1. Microbiote intestinal

La collection de micro-organismes cohabitant dans notre corps humain a fait l'objet d'études approfondies au cours des 10 dernières années, ce qui est le résultat direct du développement et de la mise en œuvre de technologies avancées de séquençage de l'ADN et de l'ARN [18]. Le microbiome, ou la totalité des génomes compris dans ces communautés, représente un ensemble diversifié d'organismes qui peuvent bénéficier ou, dans certains cas, nuire à l'hôte [19]. Ces bactéries, champignons, virus, archées et autres microbes eucaryotes (protozoaires) habitent sur toutes les surfaces du corps humain, y compris le tube digestif et le système respiratoire [20]. Il y a jusqu'à 10^{14} cellules individuelles et plus de 500 espèces de bactéries seules dans l'intestin [21]. La recherche sur le microbiome humain qui a été menée jusqu'à présent peut être décomposée en sous-catégories spécifiques basées sur le système physiologique avec lequel les microbes interagissent: nasal, oral, gastro-intestinal (intestin), urogénital et cutané. Dans le cas sain, les microbes travaillent en harmonie avec les tissus environnants pour acquérir les produits chimiques nécessaires à leur survie et fournir des nutriments, en tant que sous-produit dans de nombreux cas, qui seraient pour la plupart indisponibles autrement pour les humains à ingérer [22]. Les changements dans la structure communautaire et le contenu génétique peuvent modifier considérablement le phénotype de la communauté dans son ensemble et avoir des conséquences sur la santé humaine [23]. Cependant, bon nombre de ces perturbations du microbiome sont le résultat direct du stress induit par l'hôte humain. Les variations du régime alimentaire, l'utilisation d'antibiotiques, l'inflammation, le changement de localisation géographique et essentiellement tout facteur qui influencerait normalement la survie d'un microbe, peuvent modifier la présence et l'abondance globales des microbes communautaires [24]. Bien qu'il existe de nombreuses preuves montrant que les antibiotiques peuvent créer un microbiome dysbiotique, une étude récente a

également révélé que plus de 1000 médicaments non antibiotiques éloignent le microbiome de son état «sain » [25].

La nutrition est l'un des facteurs environnementaux les plus importants compte tenu de la stabilité et de la composition du microbiome. Plusieurs études ont rapporté des biomolécules utiles pour maintenir un microbiome qui contribue à la santé de l'individu [26], [27]. Les préférences alimentaires étant normalement basées sur la situation géographique et la disponibilité des aliments, plusieurs études ont pu comparer le microbiome intestinal d'individus ayant un apport nutritionnel très différent [28]. Les enquêtes de cette nature comparent généralement le régime «occidental», composé de sucres de faible complexité et riche en graisses, à un régime «traditionnel», caractérisé par une teneur élevée en fibres [29]. L'idée selon laquelle les régimes riches en fibres fournissent un vaste assortiment de biomolécules qui peuvent être traitées et transmises par le microbiote aux cellules intestinales humaines semble faire l'unanimité parmi la communauté scientifique [30]. Ces glucides alimentaires d'origine végétale et animale sont la source préférée d'énergie et de carbone pour les microbes intestinaux, ce qui explique précisément pourquoi le régime «traditionnel» produit de meilleurs résultats pour la santé [31]. Un régime riche en fibres est également supposé aider à pousser rapidement le bol alimentaire dans le tube digestif, empêchant ainsi l'accumulation de sous-produits microbiens dangereux. Le régime «occidental» et ses effets sur le microbiote intestinal ont été liés de manière causale à des maladies métaboliques, au cancer colorectal, au diabète, aux maladies inflammatoires de l'intestin, à l'asthme et aux allergies [32]. Le manque de complexité et de diversité du régime alimentaire peut littéralement affecter certaines souches de microbes qui fournissent les nutriments nécessaires à la subsistance humaine.

Des facteurs supplémentaires, au-delà de la mauvaise nourriture, peuvent perturber le microbiote, tels que les voyages et le stress mental et émotionnel. Bien que peu de recherches aient été menées sur le «microbiome ambulant», une étude s'est penchée sur des personnes ayant reçu un diagnostic de diarrhée du voyageur en Amérique centrale ou du Sud ou en Inde après leur arrivée aux États-Unis d'Amérique [33]. Il est intéressant de noter que les tripes des voyageurs en bonne santé ressemblaient plus à celles des patients souffrant de diarrhée qu'aux témoins sains du

projet sur le microbiome humain, ce qui indique que le stress et les changements environnementaux liés aux voyages peuvent avoir un effet dysbiotique.

La source directe de stress peut être difficile à cerner, car l'alimentation, le rythme circadien et le stress émotionnel peuvent tous jouer un rôle [34]. L'effet du cerveau sur le microbiome intestinal a été un sous-domaine passionnant qui s'est récemment développé. Ces études décrivent l'interaction chimique qui se produit entre les microbes du tractus gastro-intestinal et les systèmes nerveux central et périphérique, que l'on pense être largement médiée par les cellules immunitaires [35]. Des rapports récents ont montré l'utilisation de souches de bactéries particulières pour aider à réguler le comportement émotionnel et aider au développement du cerveau chez les souris [36], [37]. Dans l'autre sens, le corps humain peut créer une réponse inflammatoire liée au stress, poussant les cellules immunitaires résidant dans le tractus intestinal à s'engager avec le microbiote [38]. Bien que les voies soient tracées entre d'autres systèmes physiologiques, la communication entre le microbiote intestinale et le cerveau semble exister. Les rythmes circadiens, régulés par le noyau suprachiasmatique de l'hypothalamus dans le cerveau des mammifères, se sont avérés avoir un impact sur la composition du microbiome intestinal par des moyens indirects : disponibilité de la nourriture, changements environnementaux, concentration sanguine humaine d'anticorps, peptides antimicrobiens innés de cellules immunitaires et auto-anticorps [30].

1.2.1.1. Acquisition du microbiote intestinal

Le nouveau-né acquiert le microbiote à partir du microbiote vaginal, fécal et cutané de sa mère lors de l'accouchement, cette colonisation est facilitée par l'immaturation de son système immunitaire et un environnement tolérogène au sein des muqueuses, mais également par différents composés du lait maternel. Ce microbiote évolue, avec l'âge et en fonction de l'environnement et de l'alimentation. D'une manière qualitative et quantitative. Le nouveau-né acquiert sa flore adulte entre 2 et 4 ans [39].

1.2.1.2. Diversification du microbiote intestinal

Après l'âge de 6 mois Le nouveau-né est capable de consommer les aliments solides, donc le passage vers une alimentation solide est un facteur majeur de diversification du microbiote intestinal.

Le microbiote intestinal est composé de bactéries commensales, de champignons et de virus qui se font concurrence pour une niche dans un écosystème très dense.

Chez l'adulte, la variation de densité en bactéries dépend du segment de l'intestin, elle accroît en fonction de la diminution de teneur en oxygène le long du tractus gastro-intestinal [39].

1.2.2. Microbiote intestinal sain

Les bactéries et leurs bactériophages rivaux représentent la grande majorité des résidents intestinaux, tandis que les grands eucaryotes, comme les champignons, sont beaucoup moins abondants [40]. Les microbes produisent les métabolites nécessaires à l'hôte et possèdent ensemble une quantité de gènes qui dépasse le nombre du génome humain [41]. Ces gènes sont responsables de la production de métabolites pour l'hôte. La surabondance de la recherche sur le microbiome au cours de la dernière décennie s'est principalement concentrée sur la composition du microbiote, les métabolites qu'ils produisent et l'immunité, qu'elle soit adaptative ou innée. Cette recherche a conduit certains à considérer l'intestin comme le «deuxième cerveau» du corps et plus généralement comme un bon indicateur de la santé globale [42].

1.2.3. Microbiote et maladies

La composition unique de micro-organismes dans l'intestin de chaque individu a été au centre des diagnostics reliant certaines espèces ou profils taxonomiques à une pathologie clinique particulière. Bien qu'il y ait très peu de consensus sur la composition d'un microbiome sain, la plupart des études visant à lier le microbiome à un état pathologique ont plutôt choisi une conception d'étude cas-témoins, définissant la

dysbiose spécifique à la maladie de la composition intestinale par rapport à la non-maladie (groupes de contrôle). Il est prouvé que la dysbiose intestinale peut affecter la santé et la fonction d'organes distants, notamment les poumons [43], [44], le cœur [45] et le cerveau [46]. Il a été démontré que l'intestin influence la fonction de ces organes distants par le biais de métabolites produits, des cellules immunitaires circulantes et de l'inflammation [46], [47].

L'une des fonctions les plus importantes du microbiote est son rôle de barrière protectrice [48]. Les muqueuses de micro-organismes dans nos intestins comprennent des espèces extrêmement diverses de bactéries et de champignons qui impactent notre santé. Pourtant, dans des conditions différentes, certaines espèces pathogènes qui vivent (en petit nombre) parmi les commensaux, également appelées pathobiontes, peuvent atteindre un nombre qui commence à avoir un effet négatif sur l'hôte [49]. Normalement, le microbiote peut maintenir le nombre de pathobiontes à un faible niveau en devançant les ressources et / ou en stimulant la réponse immunitaire humaine [50], [51]. Comme il est généralement enregistré dans les études écologiques sur le microbiote, les microorganismes intestinaux sécrètent leurs propres peptides et toxines antibiotiques pour se protéger contre les parasites - ainsi que pour aider à éliminer les espèces concurrentes [52], [53]. Un mécanisme à trois volets est l'excrétion d'acides gras à chaîne courte des bactéries bénéfiques qui 1) abaisse le pH intestinal, empêchant la croissance de bactéries nocives, 2) sert de métabolite humain important et 3) régule à la baisse l'expression des gènes de virulence [54]. Plusieurs événements, cependant, peuvent perturber l'harmonie entre le microbe bénéfique et l'hôte, comme l'administration d'antibiotiques. Des expériences sur des souris ont montré que l'effet général des antibiotiques sur le microbiote peut augmenter considérablement le nombre de pathobiontes - y compris *Salmonella enterica*, *Escherichia coli* et *Clostridium difficile* [55]. Intelligemment, les agents pathogènes peuvent également métaboliser les sous-produits du microbiote à leur avantage. Par exemple, il a été découvert que l'acide biliaire sécrété par le microbiote que l'on trouve couramment dans les intestins dysbiotiques favorise la croissance des spores de *C. difficile* [56]. Afin d'influencer directement la communauté microbienne, les pathobiontes peuvent sécréter à la fois des toxines bactéricides et des facteurs de virulence, ce qui peut déclencher des réponses

inflammatoires de l'hôte humain, réduisant ainsi l'abondance globale du microbiote dans les intestins [57].

Les agents pathogènes bactériens intestinaux humains les plus répandus:

- Bactéries: *Campylobacter jejuni*, *Salmonella enterica*, différentes espèces de *Shigella*, *Vibrio cholerae*, *Escherichia coli*, *Clostridium difficile*

- Virus: Norovirus, Rotavirus, Adénovirus

- Protozoaires: *Giardia*, *Entamoeba*, *Cryptosporidium*, *Blastocystis*

Cette grande diversité d'agents pathogènes (virus, bactéries, protistes) présente un défi pour les communautés microbiennes bénéfiques et le corps humain. Une croissance suffisante pour provoquer une infection peut entraîner des symptômes de diarrhée, de douleurs à l'estomac, de nausées et de fièvre. La diarrhée, résultante d'une infection, peut être causée par la sécrétion de toxines par les parasites qui incitent les entérocytes à sécréter des cytokines, modifiant l'équilibre osmotique dans la lumière intestinale et provoquant l'expulsion de l'eau des tissus intestinaux [58]. L'extrusion rapide et aqueuse de ces bactéries dégage davantage le terrain de jeu pour que les agents pathogènes colonisent, s'ils sont bactériens, ou entrent en contact avec l'épithélium intestinal. Peu d'études se sont penchées sur le microbiome intestinal des patients diarrhéiques [59].

Comme ces agents pathogènes doivent pénétrer par la cavité buccale, les infections proviennent généralement de la consommation d'eau ou d'aliments contaminés. Les virus causant la diarrhée se propagent principalement par les mains, peut-être en raison du non-rinçage après l'utilisation de la salle de bain, et des facteurs de virulence à *E. coli* ont été trouvés principalement dans des sources d'eau stockées insalubres [60]. Étant donné que l'assainissement, et les infrastructures en général, sont un mécanisme clé de prévention contre la propagation des maladies infectieuses, les pays qui n'ont pas été en mesure d'investir dans des mesures de santé publique appropriées sont généralement vulnérables aux flambées.

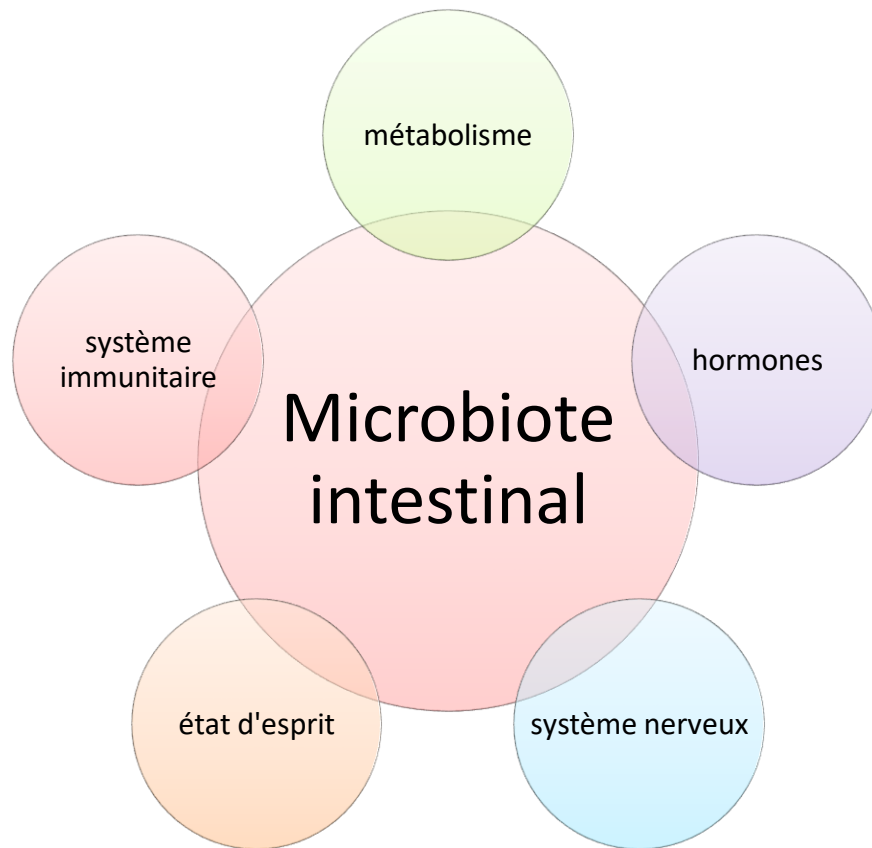


Figure 1: Relation Microbiote-Santé

Un certain nombre d'études suggèrent que l'altération du microbiote intestinal était liée à une infection virale (par exemple, le VHB et le VIH) et à *Clostridium difficile*, les maladies inflammatoires de l'intestin (MII), les maladies métaboliques comme le diabète de type 2 (DT2), l'obésité, les maladies cardiovasculaires et même les maladies auto-immunes (par exemple la PR Polyarthrite Rhumatoïde) et les maladies mentales (par exemple l'autisme). À la différence de la théorie classique «une bactérie - une maladie», les études actuelles sur les maladies et le microbiote intestinal ont commencé à utiliser des technologies indépendantes de la culture telles que les technologies de séquençage de nouvelle génération (NGS) et la bioinformatique pour l'analyse de données à grande échelle. Le principe est de construire des connexions entre le microbiote intestinal et des maladies spécifiques, et ce pour identifier des marqueurs bactériens spécifiques qui sont en corrélation avec des traits phénotypiques, puis de valider ces connexions dans des modèles animaux.

Tableau 1: Modifications du microbiote intestinal associées à la maladie

Microbiote impliqué^a	Modifications de la présence / Fonction du microbiote
Allergies	
<i>Lactobacillus</i> spp. ↓	colonisation précoce par <i>Lactobacillus</i> / diminution des allergies
<i>Bifidobacterium adolescentis</i> ↓	une colonisation précoce avec un microbiote plus diversifié pourrait prévenir les allergies
<i>Clostridium difficile</i> ↓	
<i>Helicobacter pylori</i> ↓	Tolérance à <i>H. pylori</i> médiée par les Treg qui suppriment l'asthme
Maladie cœliaque	
<i>Bacteroides vulgatus</i> ↑	diversité plus élevée chez les patients atteints de la maladie cœliaque par rapport aux témoins
<i>Escherichia coli</i> ↓	
<i>Clostridium coccoïdes</i> ↓	
Cancer de l'estomac	
<i>H. pylori</i> ↑	élément important de la voie cancérogène pour le développement d'adénocarcinomes gastriques
Autisme	
Bacteroidetes ↑	augmentation de la diversité bactérienne dans les selles des enfants autistes par rapport aux témoins
<i>Proteobacteria</i> ↑	
<i>Actinobacteria</i> ↓	
Firmicutes ↓	
Obésité	
Bacteroidetes ↓	des changements importants dans le microbiote intestinal sont associés à une augmentation de l'obésité
<i>Lactobacillus</i> ↑	
Firmicutes/Bacteroidetes ratio ↓	
<i>Methanobrevibacter smithii</i> ↓	

Anorexie	
<i>Methanobrevibacter smithii</i> ↑	<i>Bacteroidetes</i> , <i>Firmicutes</i> , et <i>Lactobacillus</i> similaires aux patients maigres, bien que <i>M. smithii</i> ait augmenté de manière significative
IBD—Maladie de Crohn	
<i>Bacteroides ovatus</i> ↑	moins de diversité chez les patients atteints de la maladie de Crohn par rapport aux patients en bonne santé
<i>Bacteroides vulgatus</i> ↑	
<i>Bacteroides uniformis</i> ↓	
IBD (Général)	
Bacteroidetes ↓	IBD associée à une dysbiose communautaire globale plutôt qu'à une seule espèce bactérienne causale
<i>Lachnospiraceae</i> ↓	
<i>Actinobacteria</i> ↑	
<i>Proteobacteria</i> ↑	
<i>Clostridium leptum</i> ↓	
<i>Clostridium coccooides</i> ↓	
<i>Faecalibacterium prasnitzii</i> ↓	
Firmicutes/Bacteroidetes ratio ↓	
<i>Bifidobacteria</i> ↓	
Type 2 Diabète	
Firmicutes ↓	modifications du microbiote intestinal associées à une augmentation des concentrations plasmatiques de glucose
<i>Clostridia</i> ↓	
<i>Bacteroides-Prevotella</i> ↑ versus <i>Clostridia coccooides-Eubacterium rectale</i> ↓	
<i>Betaproteobacteria</i> ↑	
Bacteroidetes/Firmicutes ratio ↑	

a Changements par rapport aux sujets sains. Augmentation: ↑. Diminution: ↓.

1.3. Techniques de séquençage

1.3.1. Séquençage Sanger

Le processus de détermination de l'ordre des quatre nucléotides différents Adénosine, Guanine, Cytosine et Thymine d'une molécule d'acide désoxyribonucléique (ADN) est appelé séquençage d'ADN. En 1977, Frederick Sanger et ses collègues ont publié une technique de séquençage appelée méthode de terminaison de chaîne qui est connue sous le nom de séquençage de Sanger [61]. Pour cette méthode, l'ADN à séquencer est dénaturé en ADN simple brin et divisé en quatre réactions. L'ADN polymérase et les quatre désoxynucléotides différents (dNTP) (dATP, dGTP, dCTP et dTTP) sont ajoutés aux quatre réactions. Un des quatre didésoxynucléotides (ddNTP) est ajouté à chaque réaction, c'est-à-dire ddATP à la première réaction, ddGTP à la seconde, ddCTP à la troisième et ddTTP à la quatrième. L'ADN polymérase synthétise de nouveaux brins d'ADN complémentaires de l'ADN simple brin d'entrée en incorporant les dNTP appropriés. Il incorpore également aléatoirement les ddNTP spécifiques à chaque réaction. Si un ddNTP est incorporé, l'élongation du brin s'arrête car les molécules de ddNTP n'ont pas de groupe 3'-OH qui empêche d'autres dNTP de se lier. Cela conduit à des molécules d'ADN de différentes longueurs. Les molécules sont ensuite triées en fonction de leur poids qui correspond à leur longueur.

Parce qu'un seul type de ddNTP était présent dans chaque réaction, le dernier nucléotide de chaque molécule est connu, ce qui permet la reconstruction de la séquence d'ADN.

Pendant 30 ans, le séquençage Sanger a été la technologie de séquençage la plus utilisée et a évolué progressivement. Il a été parallélisé et automatisé pour une meilleure qualité et une réduction des coûts de séquençage. Par exemple, le séquençage automatisé de Sanger a été utilisé pour séquencer le premier génome humain complet.

1.3.2. Séquençage de nouvelle génération NGS

La composition du microbiome intestinal est déterminée par deux méthodes de séquençage de nouvelle génération. La méthode de séquençage d'amplicon ciblé

(16S/ITS) est bien établie dans les études sur le microbiome et il existe de nombreuses bases de données sur lesquelles les données acquises peuvent être alignées. Le gène de l'ARNr 16S est utilisé pour identifier les bactéries, tandis que l'une des deux régions d'espacement interne transcrite (ITS) chez les eucaryotes est utilisée pour identifier les champignons. Le séquençage « Shotgun » produit de nombreux fragments génomiques différents et nécessite de grandes bases de données génomiques pour identifier ces séquences - l'inconvénient étant que de nombreux phylums n'ont pas de génomes séquencés et que pour les micro-organismes, comme les champignons, aucune base de données significative n'existe. D'autre part, le séquençage « Shotgun » offre une profondeur dans la variété d'informations, y compris les données métabolomiques et les abondances relatives des micro-organismes dans l'environnement.

Les technologies de séquençage de nouvelle génération (NGS), ou séquençage de deuxième génération, sont des méthodes de séquençage massivement parallèle de courts fragments d'ADN [62]. Ces technologies ont réduit les coûts de séquençage par paire de bases séquencée (pb) d'environ cinq ordres de grandeur par rapport au séquençage automatisé de Sanger, c'est-à-dire le séquençage de première génération. De 2005 à 2007, les premiers instruments NGS ont été introduits par Roche / 454 [63], Illumina / Solexa [64] et LifeTechnologies / ABI [65]. Ces trois technologies utilisent des biochimies et des méthodes de séquençage différentes pour l'amplification de l'ADN d'entrée, ce qui conduit à différents avantages et inconvénients en termes de longueur de lecture, de qualité et de débit. Cependant, ils partagent un flux de travail de base [66]:

- Premièrement, l'ADN d'entrée est fragmenté au hasard, suivi de la ligation des séquences d'adaptateur communes. Ce processus s'appelle la préparation de la bibliothèque.
- Dans une étape suivante, les molécules de la bibliothèque unique sont amplifiées de manière à ce que les molécules d'origine et toutes leurs copies restent regroupées dans la même position.
- Le séquençage proprement dit est ensuite effectué en alternant des cycles d'addition de nucléotides marqués par fluorescence et d'imagerie.

La préparation de la bibliothèque pour Illumina NGS commence par une fragmentation aléatoire de l'ADN génomique (figure 2). Par la suite, les adaptateurs sont ligaturés aux fragments d'ADN. Les adaptateurs permettent une liaison covalente de l'ADN à la Flowcell. Les Flowcells sont des lames de verre qui contiennent 8 canaux dans lesquels le séquençage a lieu.

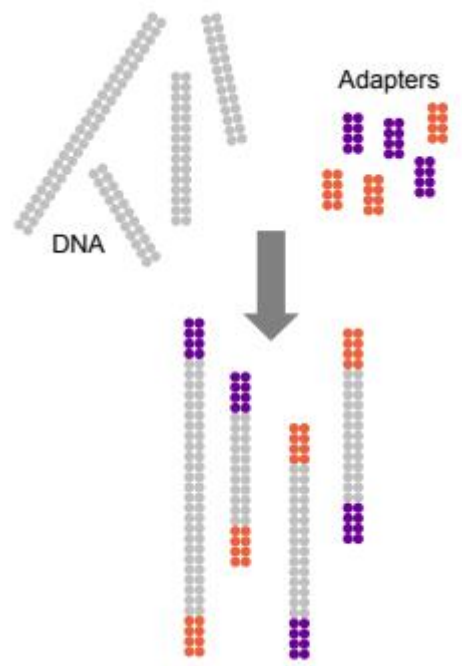


Figure 2: Schéma de préparation de la bibliothèque Illumina. Les adaptateurs sont ligaturés à des molécules d'ADN fragmentées au hasard.

Les fragments d'ADN simple brin sont ensuite liés au hasard à la surface à l'intérieur des voies (figure 3). Le séquençage proprement dit est effectué en mesurant la fluorescence des dNTP marqués incorporés. Étant donné que le signal d'un seul marqueur serait trop faible pour être détecté, une PCR dite pontée est effectuée pour générer des grappes de copies identiques autour des fragments d'ADN initiaux. Après plusieurs étapes de PCR pontée, jusqu'à 1000 copies de chaque fragment sont présentes.

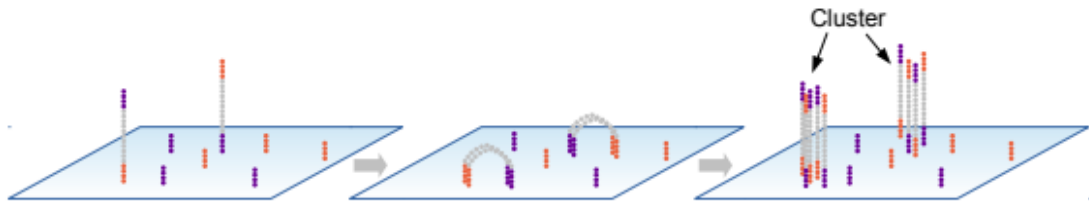


Figure 3: Schéma de génération de cluster. Les fragments monocaténaire se lient à la Flowcell. Les grappes sont ensuite générées par amplification pontée. Les clusters finaux contiennent jusqu'à 1 000 copies du fragment initial.

Le séquençage est effectué en cycles (Figure 3). Dans chaque cycle, quatre dNTP marqués différemment, des amorces et une polymérase sont ajoutés à la cuve à circulation. Les dNTP appropriés se lient aux chaînes d'acide nucléique. Le marqueur fluorescent sert de terminateur. Ainsi, un seul dNTP est ajouté à chaque chaîne d'acide nucléique par cycle. Les dNTP, les amorces abondantes sont ensuite lavés de la cuve à circulation. Après excitation laser, les signaux fluorescents sont mesurés pour chaque cluster. Ensuite, l'étiquette fluorescente est clivée et le cycle suivant commence. Après chaque cycle, le logiciel Illumina effectue un appel de base, c'est-à-dire qu'il attribue A, C, G ou T à chaque cluster en fonction des signaux fluorescents mesurés.

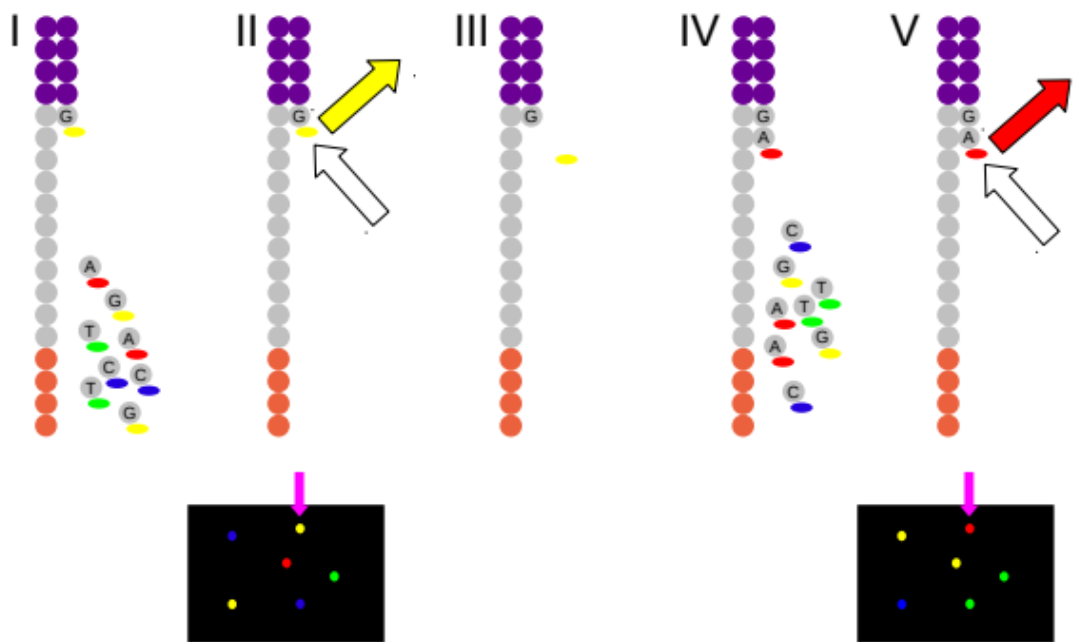


Figure 4: Schéma de séquençage d'Illumina par synthèse. Quatre désoxynucléosides triphosphates (dNTP) marqués différemment, des amorces et une ADN polymérase sont ajoutés et le dNTP approprié est ajouté à la chaîne d'acide nucléique.

Le nombre de cycles par cycle Illumina NGS est passé de 35 pb dans les premières expériences à 300 pb avec la dernière version de l'instrument MiSeq. Les instruments Illumina modernes offrent également la possibilité d'effectuer un séquençage par paires, c'est-à-dire un séquençage des deux extrémités de chaque fragment (Figure 4). En plus d'allonger la longueur de lecture, Illumina a également augmenté le débit en augmentant la densité des grappes sur la surface de la cuve à circulation. Ceci pourrait être réalisé en améliorant le séquençage ainsi que le système optique. Les instruments HiSeq 2500 ont désormais une sortie allant jusqu'à 1 terabase (To) par analyse. La longueur maximale de lecture d'Illumina est principalement limitée par un processus appelé déphasage: toutes les molécules d'un cluster n'incorporent pas un dNTP dans chaque cycle. Ces molécules incorporent ensuite le dNTP manqué dans le cycle suivant tandis que toutes les autres molécules incorporent déjà le dNTP suivant. Cet effet s'accumule au fil du temps, conduisant à des intensités de signal plus faibles et à un rapport signal sur bruit plus élevé. Ainsi, la qualité moyenne des bases séquencées diminue avec la longueur de lecture.

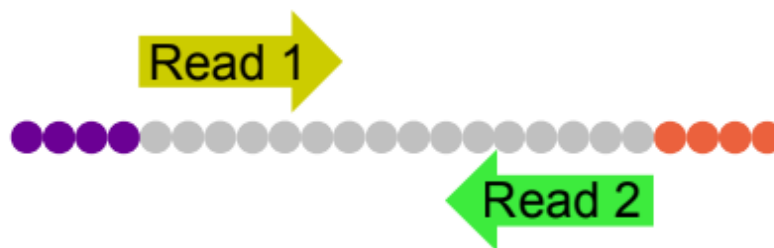


Figure 5: Schéma de séquençage par paires. En mode apparié, chaque fragment d'ADN est lu aux deux extrémités.

Dans le domaine de la génétique humaine, la NGS est principalement utilisée pour identifier des facteurs de risque dans des maladies complexes. Le séquençage du génome entier des individus affectés est désormais techniquement possible à des fins de détection de variantes. Cependant, le séquençage d'un génome humain entier avec une couverture élevée, c'est-à-dire ≥ 30 fois, reste coûteux.

Ainsi, diverses approches de séquençage ciblé sont utilisées pour la détection de variantes: Le séquençage d'amplicon est utilisé pour identifier des variantes dans de petites régions, telles que des gènes uniques, dans un grand nombre d'échantillons. Pour

la détection de variants dans un groupe de gènes connus associés à une maladie, des panels de gènes spécifiques à une maladie sont utilisés.

Si une grande attention a été accordée ces dernières années aux progrès réalisés dans le séquençage du génome humain, le séquençage de nouvelle génération a également conduit à une explosion du séquençage utilisé pour étudier les microbiomes. Il existe deux méthodes courantes de séquençage effectuées pour étudier le microbiome: le séquençage de l'ADNr 16S/ITS et la métagénomique « Shotgun ».

1.3.2.1. Séquençage rRNA 16S/ITS

L'ARN ribosomal 16S est un composant de la sous-unité 30S du ribosome bactérien, qui est un grand complexe protéique nécessaire à la traduction des séquences d'ARN messenger en protéines.

Dans les années 1970, Carl Woese et George Fox, microbiologistes américains, ont remarqué que les séquences 16S pouvaient être utilisées pour classer phylogénétiquement les microbes [67]. Ceci est possible car certaines régions de la séquence 16S ont été conservées pendant des milliers d'années, tandis que d'autres régions, dites hypervariables, ont divergé avec le temps, donnant à chaque organisme une signature génétique. L'ensemble du gène 16S couvre environ 1 600 paires de nucléotides. Les bactéries et les archées peuvent désormais être identifiées directement grâce à leurs séquences, ce que l'on a appelé la détection microbienne indépendante de la culture [68]. Traditionnellement, l'identification des microbes dépend du contrôle visuel de la croissance et de la morphologie de l'organisme sur un milieu riche en nutriments spécifiques. Une explosion de la recherche en cours d'études sur le microbiome a été possible grâce au séquençage de nouvelle génération des séquences 16S collectives dans une communauté. Pourtant, en raison de la longueur limitée du gène, par rapport à l'ensemble du génome (millions de paires de bases), la résolution d'identification peut être limitée par rapport au séquençage du génome entier [69].

On pense que le gène ribosomal 16S existe dans toutes les bactéries, mais il a encore des régions qui sont très variables entre les espèces. Pour cette raison, des amorces ont été créées pour amplifier les régions conservées qui entourent des régions variables, permettant aux chercheurs de cibler les zones des gènes qui sont similaires

pour observer les zones distinctes. Parce que cette approche nous permet d'observer des régions très spécifiques du génome, nous pouvons abandonner le séquençage nécessaire par échantillon de manière spectaculaire, ne nécessitant que 50 000 à 100 000 lectures pour identifier différentes espèces bactériennes dans un échantillon. Le principal inconvénient de cette technique est qu'elle ne peut identifier que les bactéries pour 16S ou les champignons pour ITS et n'identifie pas les virus.

1.3.2.2. Séquençage « Shotgun »

La métagénomique du « Shotgun » étudie l'ensemble des génomes de tous les microorganismes présents dans l'échantillon, par opposition aux séquences 16S uniquement. L'un des principaux avantages de ce séquençage par rapport au 16S est qu'il peut capturer des séquences de tous les organismes, y compris les virus et les champignons, qui ne peuvent pas être capturés avec le séquençage 16S. Les études 16S peuvent être extrêmement utiles pour la comparaison entre différents échantillons. En outre, la métagénomique « Shotgun » peut être utilisée pour identifier des organismes rares ou nouveaux dans la communauté, ce que 16S ne peut pas faire [69].

Chapitre 2 : Matériel et Méthodes

2.1. Zone d'étude « Sidi Ayache »

Rabat-Salé-Kénitra est une région administrative, créée par la dernière division territoriale (en 2015), entourant la capitale du Royaume du Maroc (Rabat). Elle couvre une superficie de 17569 km² représentant 2,4% de la superficie totale du Maroc (2015) [70], et elle est divisé en trois préfectures (Rabat, Salé et Skhirat-Témara) et quatre provinces (Sidi Kacem, Kenitra, Khemisset et Sidi Slimane)) (Figure 6).

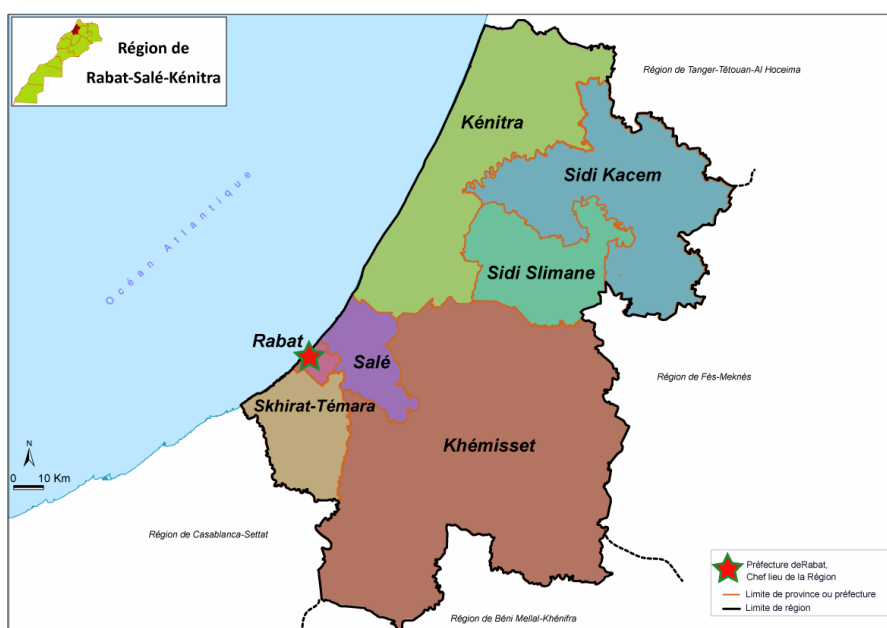


Figure 6 : Description géographique de la zone d'étude[71]

Sa population est estimée à 4 580 866 personnes correspondant à 13,5% de la population marocaine [70].

Le plateau de la Maâmora, délimité au Sud par le massif de la Meseta au Nord par la vaste plaine du Gharb, à l'Ouest par l'océan atlantique entre Rabat et Kénitra et à l'Est par l'oued Beht, recèle une nappe libre très importante d'une superficie de 1500 km². Cette nappe circule dans les grès calcaires, sables et sables argileux du Plio-quaternaire, reposant sur un substratum imperméable de marnes du Miocène.

La profondeur de la nappe est généralement comprise entre quelques mètres et 40 m et peut dépasser localement 80m.

Les eaux de la Maâmora sont extrêmement douces 250 à 500 mg/l de sels totaux ; leur faciès est bicarbonaté calcique. En zone littorale, la salinité est relativement élevée à cause de l'invasion marine entraînée par l'intensification des pompages dans cette zone.

Pour la nappe de la Maâmora, la recharge est assurée exclusivement par l'infiltration des eaux des précipitations. Les sorties sont constituées par l'écoulement vers l'océan, par le sous écoulement vers la nappe du Gharb, par le drainage de l'oued Beht et des autres oueds du plateau (Fouarat, Tiflet,...) ainsi que par les prélèvements pour l'AEPI (captages ONEP, RAK) et l'agriculture.

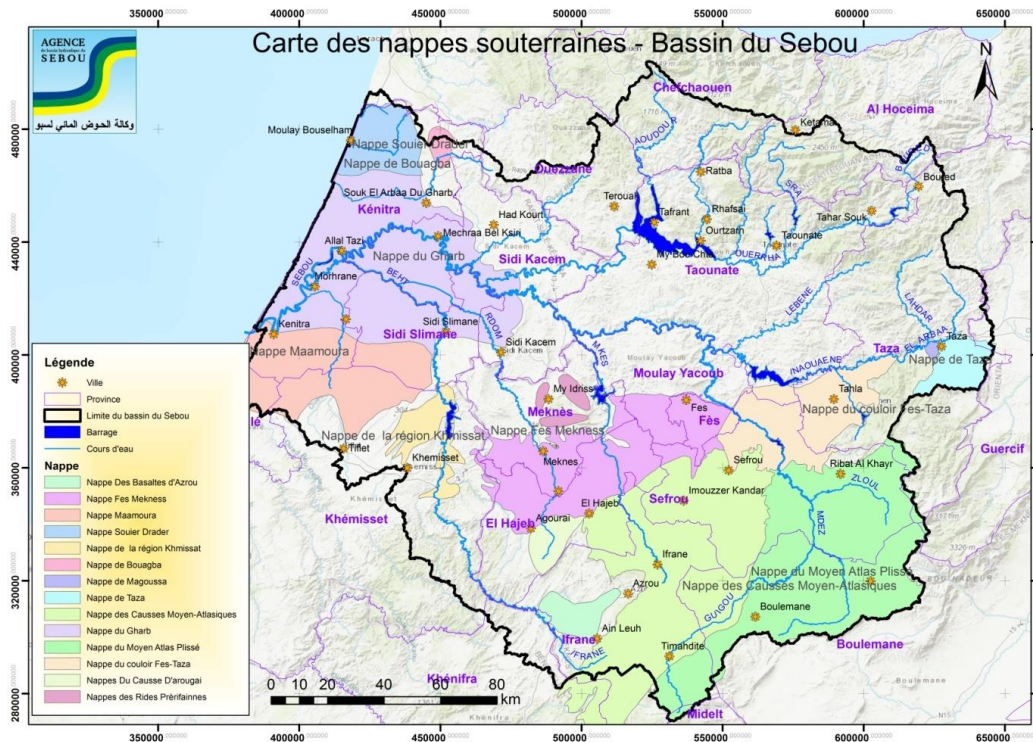


Figure 7 : Nappe Phréatique[72]

2.2. Enquête via questionnaires

Une enquête par questionnaires a été établie pour évaluer la qualité des eaux de puits dans la zone rurale d'étude, à travers le témoignage des habitants, pour avoir un

retour sur les différents usages des eaux de puits. 413 ménages ont été concernés par cette enquête.

Des questionnaires d'enquête alimentaire ont été également demandés à chaque individu. Tous les participants ont consommé un régime faible en gras et riche en fibres, avec peu d'aliments transformés et 25 g de fibres par jour, y compris des lentilles, des haricots, des sardines, des fruits et des légumes.

Nous avons également collecté des informations relatives à l'âge, au sexe et à l'indice de masse corporelle. Nous avons constaté que l'IMC était inférieur à trente (IMC <30) pour tous les participants.

2.3. Collecte des échantillons des eaux de puits

Nous avons échantillonné 10 puits dans la région Sidi Ayache. Le prélèvement d'un échantillon d'eau est une opération délicate à laquelle le plus grand soin doit être apporté ; il conditionne les résultats analytiques et l'interprétation qui en sera donnée. L'échantillon doit être homogène, représentatif et obtenu sans modifier les caractéristiques physico-chimiques de l'eau [73].

Au moment du prélèvement, on rince les bouteilles de 5L plus de trois fois avec l'eau à analyser. Ensuite, on les remplit et on place les bouchons d'une manière qu'il ne soit pas éjecté au cours du transport.

Nous avons transporté les échantillons d'eau dans une glacière à une température de 4°C jusqu'au laboratoire de Biologie, Santé et Environnement à la Faculté des Sciences et Techniques de Settat.

2.4. Collecte des échantillons des selles des habitants

Trente échantillons de selles ont été prélevés sur des sujets ruraux buvant de l'eau de puits. Quinze échantillons de selles ont été prélevés sur des sujets urbains. Chaque sujet a reçu un récipient stérilisé de 150 ml avec des instructions verbales pour excréter ses matières fécales directement dans l'ouverture du récipient, fermer rapidement le couvercle et placer l'échantillon au réfrigérateur jusqu'au lendemain où il

a été récupéré. Tous les échantillons ont été conservés dans une glacière jusqu'à ce qu'ils atteignent le laboratoire de l'Institut National d'Hygiène de Rabat. Dans les 6 heures après la collecte de tous les échantillons, chaque échantillon a été divisé en 3 aliquotes dans des tubes de 1,5 ml.

Trente-cinq échantillons de selles ont été prélevés sur des sujets vivant en milieu urbain; ils étaient tous des étudiants d'autres pays d'Afrique (Guinée, Gambie, Zambie, Nigéria, Soudan du Sud et Tanzanie). Tous ces sujets migrants vivant dans la zone urbaine Rabat buvaient de l'eau du robinet.

2.5. Traitement des échantillons des eaux de puits

La filtration de l'eau a été réalisée en utilisant un collecteur de filtration à quatre emplacements sur mesure comme réservoirs d'échantillon connectés aux unités de filtration sous pression Sterivex-GP (EMD Millipore SVGPL10RC). Les corps de seringue ont été remplis à plusieurs reprises avec de l'eau d'échantillon jusqu'à ce que 5 litres d'eau passent à travers le filtre. Aucun des échantillons d'eau de citerne de 5 litres n'a pu traverser les unités de filtration Sterivex en 1 heure. L'extension supplémentaire du temps de filtration n'a fait qu'augmenter légèrement le volume de filtration, indiquant que le filtre était colmaté. Après filtration, les unités Sterivex ont été placées dans des tubes de 15mLs et transporté dans une glacière à l'Université de Californie à Irvine et congelées instantanément au Whiteson lab.

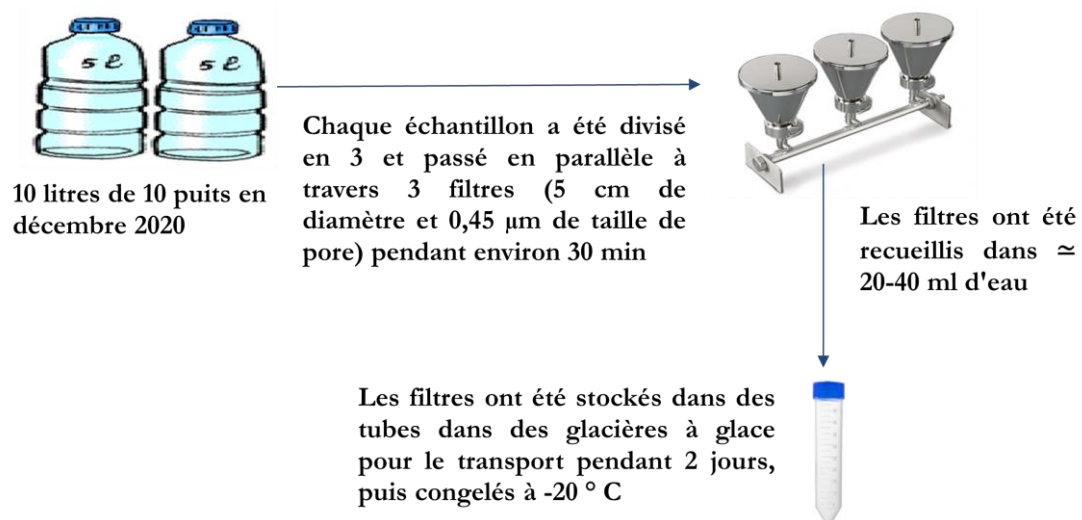


Figure 8: Traitement des échantillons d'eau

2.6. Traitement des échantillons des selles

Les échantillons ont été stockés à -20 °C au laboratoire de l'Institut National d'Hygiène de Rabat au Maroc jusqu'au transport au laboratoire Whiteson de l'Université de Californie Irvine aux Etats Unis. Les échantillons ont été décongelés sur de la glace et presque 150 mg de contenu fécal ont été retirés stérilement pour l'extraction de l'ADN.



Figure 9: Traitement des échantillons de selles

2.7. Extractions de l'ADN des microorganismes par ZymoKits

Au laboratoire Whiteson de l'Université de Californie Irvine, la biomasse microbienne sur les filtres Sterivex (stockés au congélateur à -20 °C) a été lyophilisée et extraite à l'aide du kit Quick-DNATM Fecal / Soil Microbe Miniprep en suivant les protocoles recommandés par le fabricant.

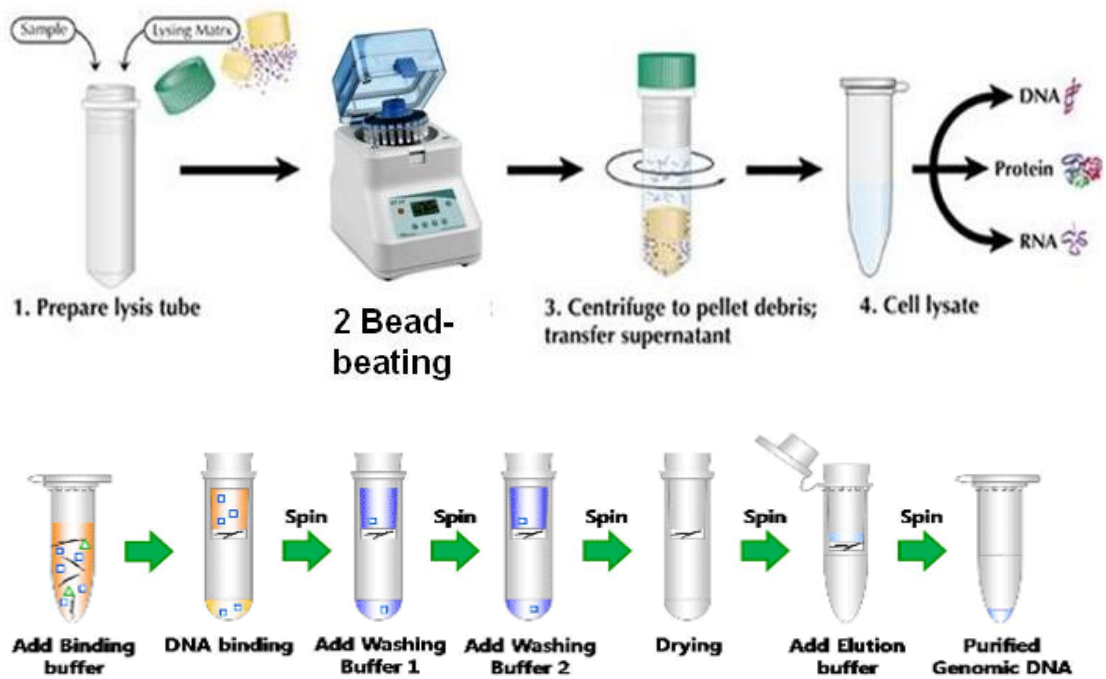


Figure 10: Les étapes des extractions de l'ADN

Pour obtenir de l'ADN à partir des échantillons de matières fécales, le kit Quick-DNA™ Fecal / Soil Microbe Miniprep a été sélectionné [74].

Tout d'abord, environ 150 mg de l'échantillon fécal sont placés dans un tube de 1,5 ml contenant des billes en plastique avec 750 µl de solution tampon de lyse (ZR BashingBead).



Figure 11: Tube de lyse contenant 150mg de matière fécale et 750 µl de solution de lyse

Les tubes de lyse sont agités à vitesse maximale sur le batteur à billes pendant 5 répétitions de 60 secondes chacune (soit un total de 5 minutes).

Les tubes d'échantillons sont ensuite centrifugés à $10000 \times g^1$ pendant 1 minute.

Jusqu'à $400 \mu\text{l}$ de surnageant est transféré dans un filtre Zymo-Spin (dessus orange) dans un tube de collecte et centrifugé à $8000 \times g$ pendant 1 minute pour éliminer les particules plus grosses.

Ensuite, $1200 \mu\text{l}$ d'une autre solution de tampon de lyse génomique, dissous avec du bêta-mercaptoéthanol pour aider à dénaturer les protéines, a été ajouté.

Seulement $800 \mu\text{l}$ de solution est transférée après avoir mélangé le tout par pipetage. Par la suite, une étape de centrifugation à $10000 \times g$ pendant 1 minute a été réalisée à travers un autre filtre qui lie l'ADN. Ce filtre est ensuite rincé deux fois, avec deux solutions de nettoyage différentes, tampon de pré-lavage d'ADN ($200 \mu\text{l}$) et tampon de lavage g-DNA ($500 \mu\text{l}$).

Enfin, le filtre a été élué dans $\sim 60 \mu\text{l}$ de tampon d'élué d'ADN à base de sel, approprié pour stabiliser l'ADN à long terme.

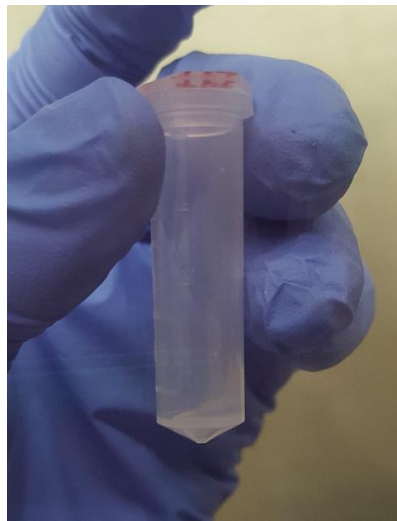


Figure 12: L'ADN résultant adapté à la PCR

¹ Values of **RCF** in units of times gravity ($\times g$)

2.8. Quantification de l'ADN par Quant-iT Pico Green

Les concentrations d'ADN ont été mesurées à l'aide du test d'ADNdb Quant-iT Pico Green de Thermo-Fisher (Waltham, MA), selon le protocole du fabricant [75].

2.8.1. Le logiciel Gen5 pour la lecture de microplaques et l'analyse des données

Gen5™ offre de puissantes fonctionnalités de lecture de microplaques, permettant des applications allant de la quantification d'acide nucléique en micro-volume aux dosages cinétiques cellulaires à long terme [76]. L'utilitaire de spectres de fluorophore offre plus de 100 spectres de fluorophore, ce qui rend la correspondance des cubes de filtre rapide et facile.

La logique conditionnelle améliorée ajoute des fonctionnalités puissantes pour la transformation des données, les seuils et la validation avec rejet automatique des valeurs aberrantes et opérateurs de formule étendus. Il offre également une amélioration de la taille des fichiers, du tri et des outils de recherche.

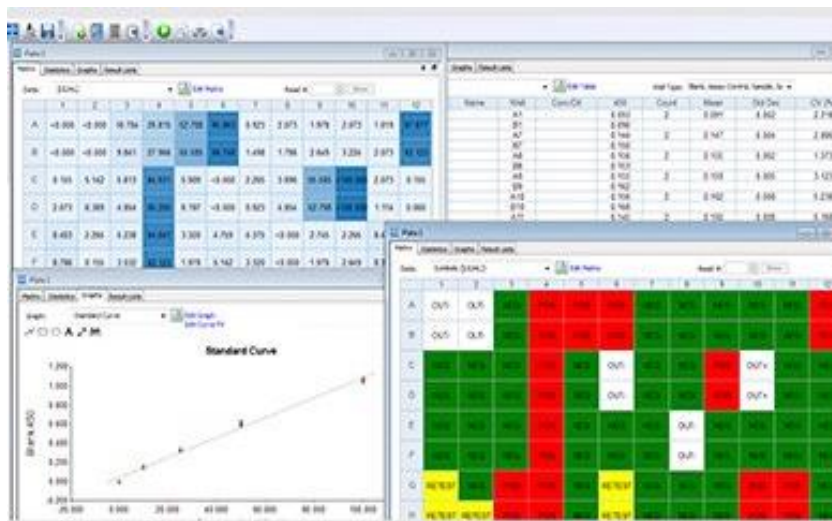


Figure 13 : Exemple de lecture et analyse de microplaques sur le logiciel Gen5

2.8.2. Intégration au système de détection BioTek

Le logiciel de lecture et d'imagerie de microplaques Gen5 s'intègre à tous les systèmes de détection BioTek : C'est un outil intégré de collecte et d'analyse de données, compatible avec tous les lecteurs multimodes et monomodes de BioTek. Il est facile de démarrer avec plus de 150 protocoles préparés dans Gen5 et en ligne. Gen5 prend en charge les modes de lecture de point final, cinétique, de zone de puits et de balayage spectral. L'exportation en un seul bouton envoie vos données directement vers Excel®, sans nécessiter d'étapes supplémentaires.

2.9. Détermination de la longueur de l'ADN

Deux méthodes de détermination de la longueur de l'ADN ont été utilisées: gel d'agarose et bioanalyseur.

2.9.1. Gel d'agarose:

Un gel d'agarose est fait à partir du mélange de 0,5 g d'agarose dans 50 ml de tampon SB (Sodium Borate or Sodium Boric Acid). La solution dissoute est chauffée au micro-ondes pendant 1 min à 70% de puissance jusqu'à effervescence. Après refroidissement, 4 µl d'une solution de bromure d'éthidium à 1% a été ajouté aux 40 ml de gel d'agarose liquide (voir protocole complet en annexe).

2.9.2. Bioanalyseur:

«Le bioanalyseur Agilent 2100 est une plateforme basée sur la microfluidique pour le dimensionnement, la quantification et le contrôle qualité de l'ADN, de l'ARN, des protéines et des cellules» [77], comme l'indique le site Web Agilent Genomics [78]. Sans entrer dans les détails, le site précise que cette technologie Lab-on-a-Chip utilise «des réseaux de canaux et de puits gravés sur une puce de verre ou de polymère». Le gel est inséré dans la puce pendant le protocole, avec 1 µL de chaque échantillon, jusqu'à 12 échantillons. Les résultats sont affichés sur l'écran de l'ordinateur, où un graphique, décrivant la distribution de longueur d'ADN, peut être vu et lu.

Les matériaux et la technologie pour créer et visualiser un gel, respectivement, sont bon marché, mais la résolution est légèrement inférieure à celle du Bioanalyzer Agilent, qui utilise la technologie à l'échelle nanométrique pour lire la longueur de la séquence. En guise de contrôle final avant le séquençage, l'utilisation du bioanalyseur peut donner des données de longueur de lecture presque exactes, malgré l'augmentation du coût, garantissant à l'utilisateur que les séquences ressemblent à la taille et à la distribution attendues.

Système Alpha-Innotech Fluorchem :

Ce système dispose d'une transillumination UV, d'une transillumination de lumière blanche et d'un éclairage de lumière blanche réfléchi pour une variété de besoins d'imagerie. Lors de l'utilisation de la transillumination UV, les filtres pour le bromure d'éthidium, le vert SYBR², le GPF et le bleu Hoechst³ peuvent être sélectionnés. Le Fluorchem utilise une caméra 8 à 16 bits à faible luminosité et haute sensibilité pour l'acquisition d'images à l'aide d'un PC Windows. Cette unité est capable de mesurer la chimioluminescence et l'imagerie de la GFP.

2.10. Amplification par Réaction en Chaîne par Polymérase PCR des bactéries 16S et champignons ITS

Découverte en 1983 par Kary Mullis, la réaction en chaîne par polymérase (PCR) est un protocole qui permet l'amplification de séquences d'ADN spécifiques [79]. Essentiellement, cette amplification est réalisée par séquençage répétitif de régions ciblées d'ADN. Tout d'abord, l'ADN est chauffé à plus de 90 ° C, ce qui fait disparaître la liaison hydrogène et sépare les brins d'ADN. La température est réduite à environ 60 ° C, ce qui permet aux amorces, petites séquences d'ADN ajoutées à la solution, de correspondre à des sites spécifiques sur l'ADN génomique. En fonction de l'ADN

² Le SYBR Green I est un composé organique aromatique de formule chimique C₃₂H₃₇N₄S faisant partie des cyanines asymétriques (fluorophores)

³ Les colorants Hoechst font partie d'une famille de colorants fluorescents bleus utilisés pour colorer l'ADN.

polymérase (qui dépend généralement de l'organisme auprès duquel elle a été acquise), la température est élevée à une température plus élevée (inférieure à 90 ° C), où les nucléotides dans la solution peuvent être utilisés pour compléter l'amorce partiellement double brin molécules d'ADN liées. Après plusieurs cycles (généralement 20 à 40), la solution résultante contient une concentration élevée de molécules d'ADN dont la séquence (en utilisant l'ADN génomique comme point de référence) se situe entre les positions où les amorces se lient.

L'amplification a été réalisée en utilisant le programme sur le thermocycleur. Tout d'abord, l'ADN est chauffé à 72 ° C pendant 3 minutes. Ensuite, la température est élevée jusqu'à 95 ° C pendant 30 secondes. Et ensuite pour 15 cycles, 95 ° C pendant 10 secondes, 55 ° C pendant 30 secondes et 72 ° C pendant 30 secondes. Enfin, la température est réduite à 72 ° C pendant 5 minutes.

Tableau 2 : Programme de PCR sur la machine BioRad C1000

94°C	3	min	
94°C	45	sec	30 cycles
55°C	30	sec	
72°C	20	sec	
72°C	10	min	
12°C	forever		

2.11. Protocole d'alimentation des amorces résiduelles

2.11.1. Nextera Library Prep

En utilisant l'ADN extrait, des bibliothèques de séquençage ont été préparées en utilisant le kit de bibliothèque Nextera d'Illumina (San Diego, CA) [80] avec le protocole décrit dans Baym et al. 2015 [81]. L'enzyme de tagmentation a été utilisée

pour attacher des codes-barres uniques à chaque échantillon. Les codes-barres ont été commandés auprès d'IDT [82].

2.11.2. AMPure XP beads

Des billes magnétiques Ampure XP (catalogue n ° A63880) de Beckman Coulter (Brea, CA) ont été utilisées pour éliminer toutes les amorces résiduelles restant libres dans les échantillons, selon le protocole du fabricant [83].

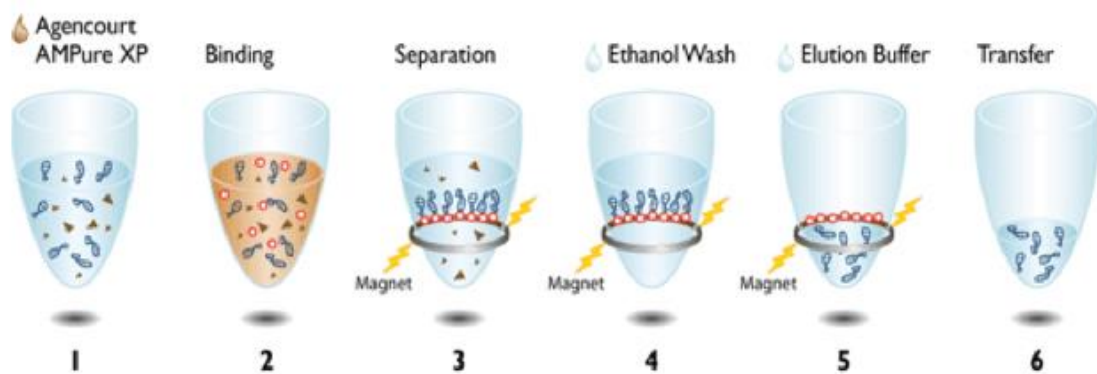


Figure 14 : Le protocole de Beckman

2.12. Séquençage de l'ARNr 16S/ITS

Les bibliothèques générées ont été chargées sur un Illumina Next-Seq à une concentration de 1,8 picomolaire à l'aide du kit à mi-sortie d'Illumina pour un séquençage par paires de 75 pb.

Pour étudier la structure de microbiome intestinale des échantillons de selles, le séquençage de l'ARNr 16S à l'aide du Next-Seq pour le système Illumina MiSeq a été utilisé [84].

Parmi les neuf régions hypervariables (V1-V9), nous avons choisi la région V3-V4, qui couvre environ 450 paires de bases. Les régions V3-V4 sont couramment utilisées dans les études médicales et environnementales classant les communautés microbiennes [85]. Les amorces des régions V3-V4, suggérées par le protocole Nextera XT, ont été conçues pour se lier au bord des régions conservatrices, permettant aux amorces de cibler toutes les espèces bactériennes.

Amplicon 16S V3-V4 (5' → 3')

F:

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG

R:

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTA
ATCC

Chaque lettre des séquences ci-dessus représente l'un des quatre nucléotides (A pour l'adénine, G pour la guanine, T pour la thymine et C pour la cytosine) et d'autres lettres représentent la présence de toute combinaison spécifique des nucléotides. Par exemple, W peut remplacer un A ou un T.

Le protocole Nextera XT commence par une réaction PCR utilisant les amorces de la région V3-V4 à partir de l'ADN extrait des échantillons. L'ADN est vérifié avec un gel d'agarose pour déterminer si les amplicons, ou séquences amplifiées, sont de la bonne taille. Si une PCR réussie s'est produite, l'ADN est nettoyé avec de l'éthanol et une autre réaction de PCR avec huit adaptateurs Nextera avant et douze adaptateurs inverses, donnant 96 combinaisons d'adaptateurs uniques, avec un autre nettoyage après la PCR. Les échantillons sont ensuite vérifiés sur le bioanalyseur Agilent pour déterminer les données haute résolution pour la taille de l'amplicon. Étant donné que les adaptateurs servent d'étiquette unique pour chaque échantillon, 96 échantillons peuvent être regroupés dans une seule solution.

L'échantillon a été dilué avec du chlorure de sodium et de l'eau (pour séparer les brins) à une concentration appropriée pour le séquenceur MiSeq. La machine Illumina MiSeq des KAUST Bioscience Core Labs a été utilisée. Le kit de réactifs MiSeq v3 (600 cycles) a été utilisé, ce qui signifie que 300 paires de bases de chaque extrémité de la molécule d'environ 500 paires de bases ont été lues, également connu sous le nom de séquençage par paires.

2.13. Séquençage métagénomique « Shotgun »

Bien que le séquençage de l'ARNr 16S permette des analyses globales, le séquençage « Shotgun » du génome entier est destiné à produire des ensembles de données qui peuvent être utilisés pour comprendre le répertoire global de gènes représenté dans un échantillon. Cette approche séquence l'intégralité de l'ADN dans un échantillon. De nombreux ensembles de gènes nous intéressent dans le cadre de cette étude: gènes de résistance aux antimicrobiens, facteurs de virulence, gènes marqueurs liés à l'origine géographique et gènes associés à des caractéristiques métaboliques uniques. Ainsi, pour obtenir plus d'informations sur les caractéristiques génétiques du microbiome intestinal, un séquençage « Shotgun » du génome entier sur l'Illumina HiSeq 4000 avec le kit de préparation de bibliothèque d'ADN NEBNext® Ultra™ II a été réalisé.

Le kit de préparation de bibliothèque d'ADN NEBNext® Ultra™ II pour Illumina® a été sélectionné en fonction de la disponibilité de l'ADN. La première étape du protocole nécessite le cisaillement de l'ADN afin de créer une longueur uniforme pour le séquençage. Un ultrasoniseur Covaris a été utilisé sur les échantillons pour briser l'ADN en morceaux d'environ 500 paires de bases. Les extrémités cassées sont réparées à l'aide du mélange d'enzymes NEBNext End Prep, qui comprend une ADN polymérase qui remplit les nucléotides manquants. Une solution contenant des adaptateurs d'acide nucléique et des enzymes de ligature est ajoutée à la solution. Une solution de billes magnétiques est ensuite ajoutée, qui se lie à l'ADN, l'isolant de la solution elle-même. Une solution 0,9X des billes est ajoutée afin de sélectionner les brins de longueur proche de 500 paires de bases. Une solution d'éthanol à 80%, qui précipite l'ADN, est ensuite utilisée pour nettoyer les billes, deux fois. L'ADN attaché aux billes est libéré dans un tampon d'élution (à base de sel). Un mélange NEBNext PCR a ensuite été ajouté afin d'amplifier l'ADN dans la solution, permettant une meilleure couverture de séquençage. Deux amorces, une universelle et une contenant une séquence unique, ont été ajoutées à chaque échantillon. L'ADN amplifié est nettoyé deux fois avec 80% d'éthanol et les lectures sont vérifiées sur le bioanalyseur pour déterminer la distribution de longueur.

2.14. Analyse bioinformatique

Les données de séquençage Fastq ont été analysées à l'aide du cluster High-Performance Computer (HPC) fourni par l'Université de Californie à Irvine. L'exécution initiale d'Illumina Nextseq a donné 193 millions de lectures avec une moyenne de 1,4 million de lectures par échantillon. La deuxième série Illumina Nextseq a donné 166 millions de lectures avec une moyenne de 2 millions de lectures par échantillon. Une combinaison des outils Fastqc et Prinseq a été utilisée pour filtrer et découper les lectures appariées de séquences brutes. La longueur minimale des lectures était de 70 pb. Un total de 16% des lectures ont été perdues.

Les séquences générées, ou lectures, ont ensuite été soumises à plusieurs niveaux de nettoyage et de formatage avant d'entrer dans l'analyse.

En utilisant le kit Nextera XT, un fichier fastq a été généré pour chacune des séquences inverse et directe de chaque échantillon groupé. Pour fusionner les régions qui se chevauchent des lectures avant et derrière chaque fichier pour créer une lecture plus longue, l'outil bbmerge a été utilisé. Selon les auteurs des outils, BBMerge est la fusion de lecture basée sur le chevauchement la plus rapide et la plus précise. Cependant, les lectures non fusionnées ont été soumises à un ajustement de qualité supplémentaire et ont été remergées, en utilisant respectivement bbdup et bbmerge. Les lectures des fichiers remergés ont été concaténées avec les données du fichier de lectures fusionné d'origine. Les lectures ont ensuite été corrigées en utilisant l'algorithme BayesHammer. Cet outil regroupe essentiellement des sections de lectures pour déterminer la qualité de lecture, rejetant potentiellement les lectures rares supposées être erronées [86], [87]. En moyenne, environ 25% à 60% des lectures ont été laissées dans le fichier échantillon, en fonction de la qualité des lectures, après le prétraitement.

Les lectures nettoyées ont ensuite été introduites dans le protocole bioinformatique Mothur MiSeq pour la classification de l'ARNr 16S [88]. Tout d'abord, les lectures ont été criblées en fonction de la longueur (entre 200 et 500 paires de bases). Puisque pas plus de huit homopolymères ont été enregistrés dans une séquence d'ARNr 16S, toute lecture avec ce numéro ou plus du même nucléotide a été exclue. Les fichiers

sont ensuite traités par une fonction qui produit un fichier avec les valeurs de lecture uniques et un autre fichier avec le numéro, le nom et le contenu de chaque lecture. Le fichier contenant uniquement les séquences uniques est plus facile, d'un point de vue informatique, à gérer pour la classification taxinomique. Les lectures uniques ont été alignées sur la base de données ARNr 16S, qui contient des séquences d'ARNr 16S de microbes génétiquement caractérisés jusqu'à présent [89]. Les lectures passent ensuite par une étape de « pré-clustering », une autre méthode de correction d'erreur, qui consiste en un algorithme qui classe les séquences dans l'ordre de leur abondance, puis recherche les séquences plus rares qui se trouvent dans un pourcentage spécifique de similitude de la séquence originale. Les séquences à l'intérieur du seuil de similarité sont fusionnées, réduisant les données et, par conséquent, le temps de calcul dans les étapes ultérieures.

Notre flux de travail original comprenait les trois étapes suivantes:

- 1) pré-clustering,
- 2) création d'une matrice de distance (fournissant des valeurs quantitatives pour la similitude entre les séquences) et
- 3) clustering à l'aide de l'algorithme de mothur.

La matrice de distance est construite en effectuant des comparaisons de similitude quantitative par paires entre chacune des séquences uniques. Cependant, avec plus de 500 échantillons, le nombre de séquences uniques devrait augmenter, ce qui conduit à une augmentation de la quantité globale de données stockées (temporaires). Le calcul des distances entre le grand nombre de séquences dépassait systématiquement les limites de mémoire. Après avoir approfondi ce problème, un article de l'auteur correspondant du manuscrit présentant le mothur, le professeur Patrick Schloss déclarait ce qui suit [90]:

Pour l'étape suivante, le regroupement, utilise également la similarité de séquence pour définir des unités taxinomiques opérationnelles, ou des regroupements de lectures qui sont supposées être dans la même division taxinomique bactérienne. Le calcul requis dans l'étape de clustering est également exigeant en termes de calcul et nécessite une quantité importante de RAM (Random Access Memory) et de nœuds

informatiques afin de regrouper efficacement et correctement les lectures. Afin d'atténuer le problème de mémoire rencontré avec la commande 'dist.seqs', créant la matrice de distance, l'algorithme UCLUST a été utilisé sur les échantillons, qui a identifié avec succès et rapidement les clusters (ou OTU) auxquels les lectures pouvaient être attribuées sans avoir pour créer une matrice de distance [91].

Par la suite, les lectures ont été scannées et identifiées avec la commande «classify.seqs». Les séquences mitochondriales, eucaryotes et autres séquences non identifiées ont été supprimées à l'aide de la commande «remove.lineage» de mothur. Les trois fichiers résultants suivants ont été reformatés pour être analysés plus en détail à l'aide du package phyloseq du langage R Statistical Computing [92], [93].

1) Table OTU - informations concernant le nombre de OTU (identifiants uniquement) présents par échantillon

2) Fichier de taxonomie - informations correspondant à la classification taxonomique de chaque OTU

3) Échantillon de métadonnées - données pour chaque individu (par exemple, âge, origine géographique, ...)

R est un environnement logiciel gratuit pour le calcul statistique et les graphiques. Il compile et fonctionne sur une grande variété de plates-formes UNIX, Windows et MacOS.

Essentiellement, R est un langage informatique qui fournit plusieurs fonctions statistiques utiles, en particulier pour l'analyse de grands ensembles de données. RStudio est un environnement de développement intégré open-source pour R qui est couramment utilisé en raison de sa facilité à visualiser et à manipuler les données, par rapport à la ligne de commande seule. R et RStudio ont été utilisés pour l'analyse quantitative des OTU d'ARNr 16S/ITS trouvés dans les échantillons.

Le package Phyloseq R associe l'interface graphique du package ggplot2 (l'un des packages de visualisation de données les plus populaires écrits pour R) aux analyses statistiques couramment utilisées pour les études sur le microbiome. En plus des calculs

plus avancés, le package phyloseq est un moyen pratique de produire des graphiques à barres simples à partir de métadonnées à grande échelle.

En considérant comment définir et mesurer la diversité dans les échantillons métagénomiques, les deux représentations de données les plus courantes faites dans les études d'ARNr 16S/ITS sont les graphiques de diversité alpha et bêta. La mesure de la diversité alpha représente une estimation de la richesse, qui est le nombre total d'espèces dans un échantillon, et de la régularité, autrement appelée l'abondance relative des espèces. La diversité alpha est utilisée pour avoir une idée de la diversité du microbiome dans un échantillon individuel. Plus précisément, les indices Chao1 et Shannon ont été utilisés pour évaluer la diversité des échantillons. La diversité bêta est définie comme la proportion d'espèces trouvées dans un échantillon par rapport à l'environnement dans son ensemble, ou, en d'autres termes, la diversité entre les individus.

Ce calcul est basé sur la distance déterminée quantitativement entre les échantillons, en utilisant les lectures et les métadonnées. Plus précisément, les paramètres de dissimilarité de Bray-Curtis ont été utilisés pour vérifier les mesures de diversité bêta. L'abondance différentielle des espèces et l'analyse multivariée de la variance (MANOVA) ont également été calculées en utilisant le phyloseq. Les valeurs p inférieures à 0,05 étaient considérées comme significatives.

N.B. L'analyse des données du « Shotgun » est semblable au pipeline bioinformatique d'ARNr 16S/ITS.

Chapitre 3 : Résultats

3.1. Résultats de l'enquête par questionnaires

413 réponses au questionnaire nous ont permis d'obtenir des avis sur la qualité des eaux de puits et sur les souhaits des usagers des eaux des puits pour améliorer l'ensemble des services a savoir la qualité des fosses septiques de la région. L'échantillon enquêté représente près de 58% d'hommes et 42% de femmes.

3.1.1. Puits : couvert/non couvert

Près de 63 % des personnes interrogées prennent soin de leur puits et les couvrent alors qu'environ 37% ne le font pas. Ceci dit, plus de la moitié de la population interrogée accorde de l'attention par rapport à la protection des eaux de leurs puits alors que le reste qui constitue 37% ne prête pas d'attention quant à la protection des eaux de puits principalement à cause de leur ignorance.

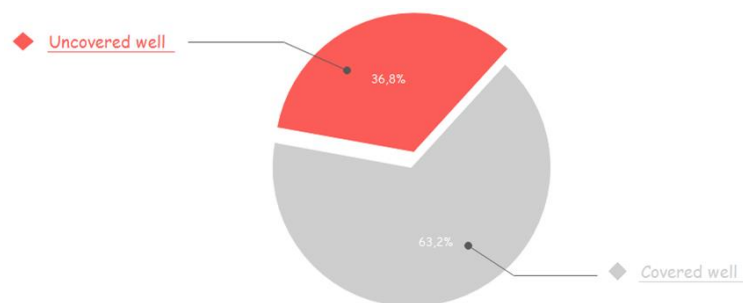


Figure 15: Couverture ou non des puits

3.1.2. Profondeur des puits

Le pourcentage des puits qui ont une profondeur entre 21-23m occupe près de 58%. Les mesures de profondeur entre 18-20 et 24-26 partagent le reste du cercle. Il aurait été souhaitable d'avoir un résultat avec une forte dominance des puits avec une profondeur 24-26m parce qu'il est fortement recommandé vu qu'il est plus proche de la nappe phréatique à cette profondeur.

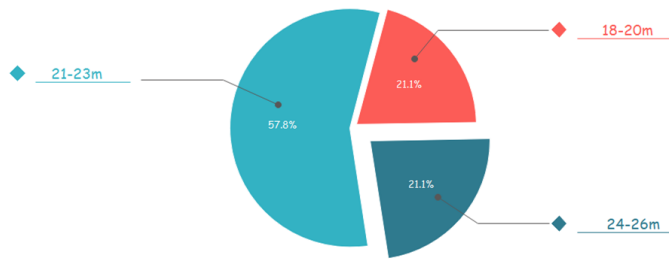


Figure 16: Profondeur des puits

3.1.3. Type de construction des fosses septiques

La majorité des fosses septiques est construite par des roches, alors qu'elles doivent être construites à base de béton afin de répondre aux normes internationales de construction des fosses septiques.

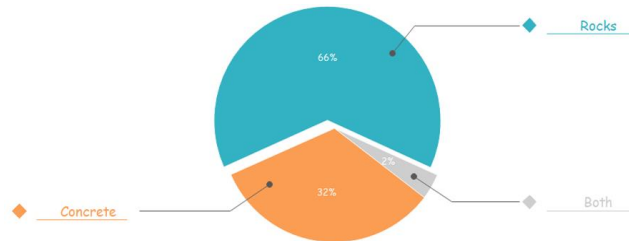


Figure 17: Type de construction des fosses septiques

3.1.4. Profondeur des fosses septiques

68 fosses septiques ont une profondeur qui dépasse les 5 mètres alors qu'elle devrait être comprise entre 3 et 5 mètres pour assurer le maximum de distance entre les fosses septiques et la nappe phréatique sinon les premières risquent d'impacter les deuxièmes.

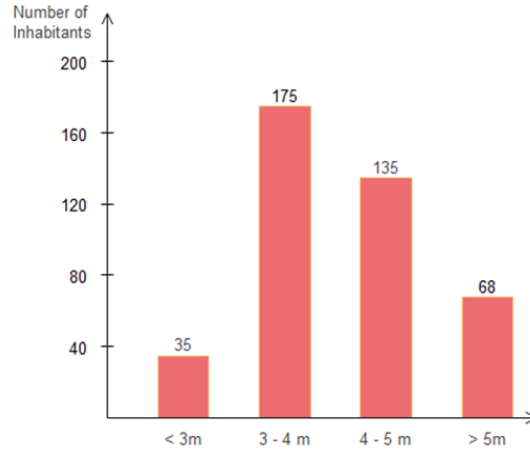


Figure 18: Profondeur des fosses septiques

3.1.5. Distance entre fosse septique et puits

Seulement 18.3% des habitats laissent une distance suffisante entre les fosses et les puits qui varie entre 20 et 25m.

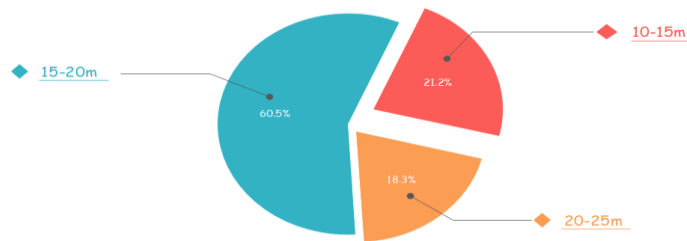


Figure 19: Distance entre fosse septique et puits

3.1.6. Appréciation de l'eau de puits

Les résultats de l'enquête ont révélé que la plupart des habitants doutent de la qualité des eaux de puits.

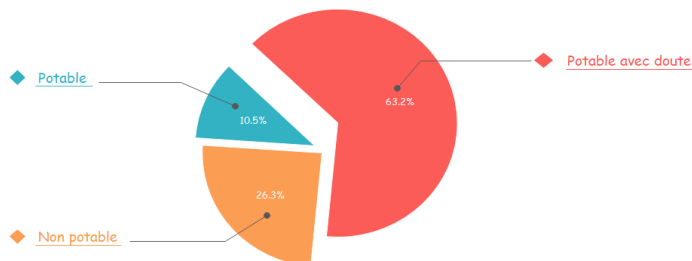


Figure 20: Appréciation de l'eau de puits

3.1.7. Changement au goût de l'eau des puits

73% ont remarqué un changement au goût des eaux de puits tandis que le quart n'a rien sentis.

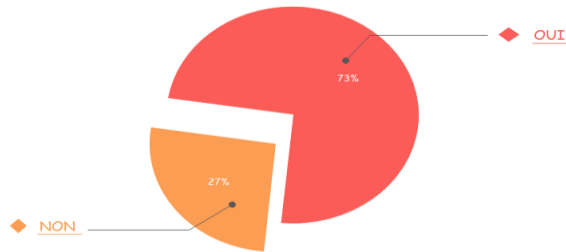


Figure 21: Existence de changement au goût de l'eau des puits

3.1.8. Problèmes de santé

Parmi les maladies les plus courantes dans la zone d'étude sont la diarrhée, la fièvre et les allergies. On a remarqué que la diarrhée était la plus répétitive.

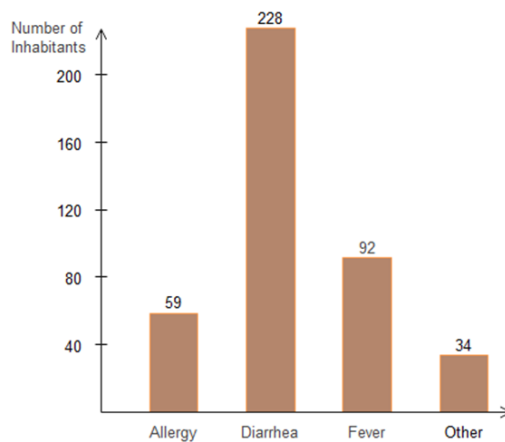


Figure 22: Existence de problèmes de santé

3.1.9. La cause des maladies

La majorité des habitants prétendent que l'ensemble des maladies observées dans la région pourraient provenir de la mauvaise qualité des eaux des puits ayant un mauvais goût et par conséquent n'est pas potable.

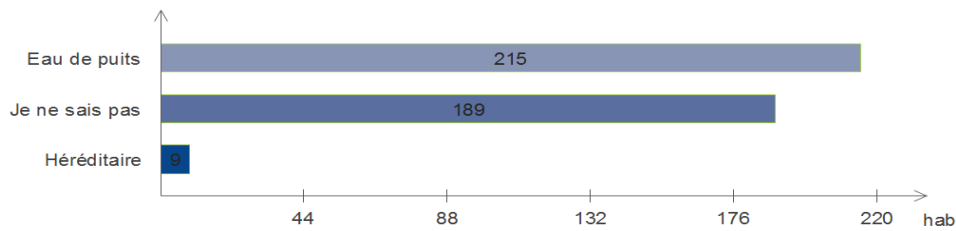


Figure 23: La cause des maladies

Les maladies observées dans la région pourraient provenir de la mauvaise qualité des eaux des puits ayant un mauvais goût et par conséquent non potable.

On remarque un non respect de distance entre les fosses septiques et les puits. Il y a même des ménages qui ont transformé le puits en fosse septique.

Les résultats de l'enquête menée prouvent certainement qu'il faut faire face aux problèmes de pollution de la nappe phréatique.

3.2. Résultats des analyses des échantillons des eaux de puits

3.2.1. Quantification de l'ADN des microorganismes dans les filtres et la matière lyophilisée

La quantité d'ADN dans chaque échantillon a donné des concentrations différentes mais a rapporté de bons résultats (tableau 3).

Tableau 3: Quantification de l'ADN des échantillons des eaux de puits

Sample	Sample Name	RFU	Concentration from curve ug/mL	Sample Stock Concentration ng/UI
1	1L	3177.05	0.094	11.728
	1F	175.66	0.005	0.648
2	2L	831.40	0.025	3.069
	2F	1763.65	0.052	6.511
3	3L	16648.18	0.492	61.458
	3F	3128.47	0.092	11.549
4	4L	3192.44	0.094	11.785
	4F	3976.99	0.117	14.681
5	5L	2295.27	0.068	8.473
	5F	441.60	0.013	1.630
6	6L	977.93	0.029	3.610
	6F	2789.78	0.082	10.299
7	7L	1329.15	0.039	4.907
	7F	575.11	0.017	2.123
8	8L	793.57	0.023	2.930
	8F	2875.50	0.085	10.615
9	9L	3118.30	0.092	11.511
	9F	1713.21	0.051	6.324
10	10L	3390.67	0.100	12.517
	10F	9797.47	0.289	36.168

Si un ADN est élevé dans le matériel lyophilisé, nous en trouvons moins dans les filtres et vice-versa, nous remarquons également des quantités qui se complètent dans les filtres et le matériel lyophilisé. Cela signifie que si l'ADN manque dans les filtres, il se retrouvera dans le matériel lyophilisé. Il est donc logique que la quantité d'ADN dans chaque échantillon soit distribuée à la fois dans les filtres et dans le matériel lyophilisé.

3.2.2. Traitement des résultats de séquençage « Shotgun »

3.2.2.1. Analyse d'abondance relative

Nous avons présenté la distribution des abondances relatives des 10 principales espèces importantes dans la figure 25.

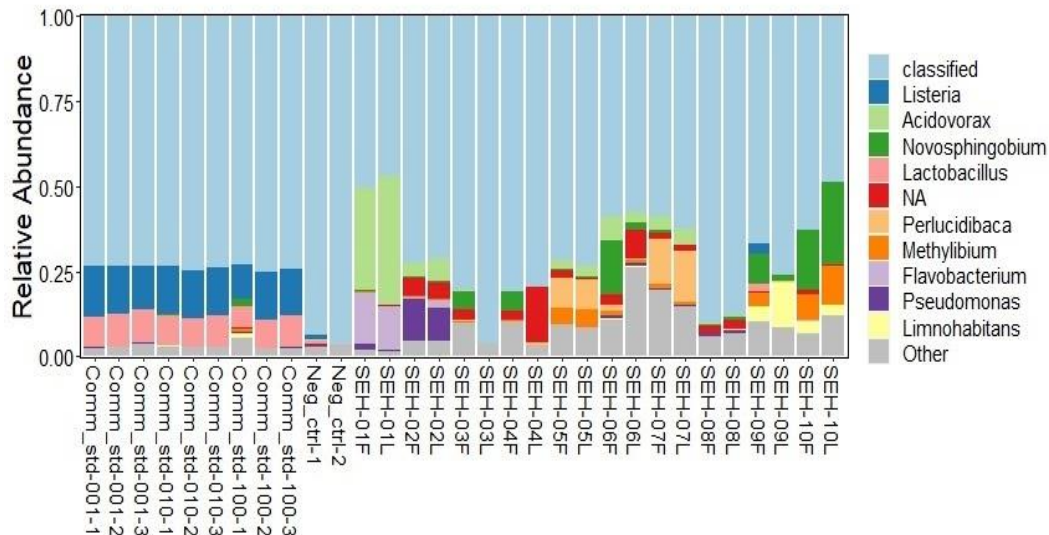


Figure 24: Les abondances relatives des 10 principales espèces dans les échantillons déterminées par séquençage « Shotgun ».

Nous avons constaté que l'*Acidovorax* était présent en abondance dans le premier échantillon. Le *novosphingobium* était présent dans tous les échantillons. Le *perlicidibaca* était présent en abondance dans trois échantillons. Le *méthylibium* était présent dans tous les échantillons. *Flavobacterium* était présent en abondance dans le premier échantillon. *Pseudomonas* était présent en abondance dans le deuxième échantillon. Les *limnohabitans* étaient présents en abondance dans les deux derniers échantillons. *Polaromonas* était présent en abondance dans le dernier échantillon. Enfin *Macondimonas* était présent en abondance dans l'échantillon numéro 7.

Les espèces pathogènes qui n'ont pas été classés parmi les 10 premières espèces abondantes mais qui ont été détectés dans les échantillons sont les suivants :

- *Vibrio_rotiferianus*

Vibrio_rotiferianus était présent dans l'échantillon 5. C'est un pathogène marin capable de provoquer des maladies chez divers organismes aquatiques [94].

Tableau 4: Détection de *Vibrio_rotiferianus*

taxonomy	<i>Vibrio_rotiferianus</i>
Comm_std-001-1	0
Comm_std-001-2	0
Comm_std-001-3	0
Comm_std-010-1	0

Comm_std-010-2	0
Comm_std-010-3	0
Comm_std-100-1	0
Comm_std-100-2	0
Comm_std-100-3	0
Neg_ctrl-1	0
Neg_ctrl-2	0
SEH-01F	0
SEH-01L	0
SEH-02F	0
SEH-02L	0
SEH-03F	0
SEH-03L	0
SEH-04F	0
SEH-04L	0
SEH-05F	118
SEH-05L	172
SEH-06F	0
SEH-06L	0
SEH-07F	2
SEH-07L	2
SEH-08F	0
SEH-08L	0
SEH-09F	0
SEH-09L	0
SEH-10F	0
SEH-10L	0

- *Salmonella*

Salmonella a été trouvée dans tous les échantillons de puits, à l'exception de l'échantillon 1. La salmonelle est considérée comme un indicateur de l'eau qui a été contaminée par les excréments d'humains ou d'animaux infectés. C'est l'une des principales causes de maladies intestinales dans le monde ainsi que l'agent étiologique de maladies systémiques plus graves telles que la fièvre typhoïde et paratyphoïde [95].

Tableau 5: Détection de Salmonella

taxonomy	Salmonella
Comm_std-001-1	20859
Comm_std-001-2	25008

Comm_std-001-3	11803
Comm_std-010-1	1320
Comm_std-010-2	1339
Comm_std-010-3	3021
Comm_std-100-1	162
Comm_std-100-2	70
Comm_std-100-3	305
Neg_ctrl-1	0
Neg_ctrl-2	0
SEH-01F	0
SEH-01L	0
SEH-02F	3
SEH-02L	10
SEH-03F	0
SEH-03L	2
SEH-04F	4
SEH-04L	1
SEH-05F	2
SEH-05L	1
SEH-06F	0
SEH-06L	2
SEH-07F	0
SEH-07L	2
SEH-08F	6
SEH-08L	11
SEH-09F	8
SEH-09L	18
SEH-10F	13
SEH-10L	29

- *Mycobacterium*

Des espèces du gène *Mycobacterium* ont été détectées dans les échantillons.

Tableau 6: Détection des espèces du gène *Mycobacterium*

taxonomy	<i>Mycobacterium</i> sp. NAZ190054	<i>Mycolicibacterium</i> brisbanense	<i>Mycolicibacterium</i> novocastrense	<i>Mycolicibacterium</i> wolinskyi
Comm_std-001-1	0	0	0	0
Comm_std-001-2	0	0	0	0
Comm_std-001-3	0	0	0	0
Comm_std-010-1	0	0	0	0

Comm_std-010-2	0	0	0	0
Comm_std-010-3	0	0	0	0
Comm_std-100-1	0	0	0	0
Comm_std-100-2	0	0	0	0
Comm_std-100-3	0	0	0	0
Neg_ctrl-1	0	0	0	0
Neg_ctrl-2	0	0	0	0
SEH-01F	0	4	0	2
SEH-01L	0	15	0	0
SEH-02F	69	0	0	44
SEH-02L	54	0	0	35
SEH-03F	51	3	0	37
SEH-03L	2	0	0	2
SEH-04F	66	0	0	44
SEH-04L	0	2	0	0
SEH-05F	12	8	0	2
SEH-05L	4	7	0	2
SEH-06F	2	0	53	0
SEH-06L	2	0	31	0
SEH-07F	76	100	18	0
SEH-07L	61	45	4	0
SEH-08F	14	0	0	0
SEH-08L	0	0	0	0
SEH-09F	0	0	0	0
SEH-09L	0	0	0	0
SEH-10F	10	0	0	0
SEH-10L	0	0	0	0

- *Legionella*

Legionella_londiniensis et *Legionella_waltersii* ont été trouvées dans l'échantillon 6. *Legionella* est un exemple de microorganisme pathogène connu pour sa réplication dans les amibes qui sont protégés et structurellement résistants à l'action des bactéricides, permettant ainsi le transport et la survie des pathogènes d'origine hydrique [96].

Tableau 7: Détection de Legionella

taxonomy	<i>Legionella_londiniensis</i>	<i>Legionella_waltersii</i>
Comm_std-001-1	0	0
Comm_std-001-2	0	0

Comm_std-001-3	0	0
Comm_std-010-1	0	0
Comm_std-010-2	0	0
Comm_std-010-3	0	0
Comm_std-100-1	0	0
Comm_std-100-2	0	0
Comm_std-100-3	0	0
Neg_ctrl-1	0	0
Neg_ctrl-2	0	0
SEH-01F	0	0
SEH-01L	0	0
SEH-02F	0	0
SEH-02L	0	0
SEH-03F	0	0
SEH-03L	0	0
SEH-04F	0	0
SEH-04L	0	0
SEH-05F	0	0
SEH-05L	0	0
SEH-06F	36	31
SEH-06L	112	88
SEH-07F	0	0
SEH-07L	0	0
SEH-08F	0	0
SEH-08L	0	0
SEH-09F	0	0
SEH-09L	0	0
SEH-10F	0	0
SEH-10L	0	0

- *Bacillus*

Bacillus a été trouvé dans les échantillons 3, 4, 9 et 10. Les membres de *Bacillus* sont capables de produire des spores hautement résistantes aux traitements de désinfection [97].

Tableau 8: Détection de Bacillus

taxonomy	Bacillus
Comm_std-001-1	7141
Comm_std-001-2	7824

Comm_std-001-3	3433
Comm_std-010-1	422
Comm_std-010-2	408
Comm_std-010-3	932
Comm_std-100-1	99
Comm_std-100-2	32
Comm_std-100-3	99
Neg_ctrl-1	0
Neg_ctrl-2	0
SEH-01F	0
SEH-01L	0
SEH-02F	0
SEH-02L	0
SEH-03F	0
SEH-03L	2
SEH-04F	2
SEH-04L	0
SEH-05F	0
SEH-05L	0
SEH-06F	0
SEH-06L	0
SEH-07F	0
SEH-07L	0
SEH-08F	0
SEH-08L	0
SEH-09F	10
SEH-09L	2
SEH-10F	2
SEH-10L	10

- *Pseudomonas_aeruginosa*

Pseudomonas_aeruginosa a été trouvé dans tous les échantillons. C'est une bactérie pathogène d'origine hydrique selon l'OMS (2008).

Tableau 9: Détection de *Pseudomonas_aeruginosa*

taxonomy	<i>Pseudomonas_aeruginosa</i>
Comm_std-001-1	342
Comm_std-001-2	393

Comm_std-001-3	177
Comm_std-010-1	26
Comm_std-010-2	26
Comm_std-010-3	32
Comm_std-100-1	4
Comm_std-100-2	2
Comm_std-100-3	4
Neg_ctrl-1	0
Neg_ctrl-2	0
SEH-01F	8
SEH-01L	0
SEH-02F	2
SEH-02L	0
SEH-03F	8
SEH-03L	0
SEH-04F	2
SEH-04L	3
SEH-05F	0
SEH-05L	2
SEH-06F	27
SEH-06L	39
SEH-07F	5
SEH-07L	3
SEH-08F	984
SEH-08L	1813
SEH-09F	0
SEH-09L	3150
SEH-10F	176
SEH-10L	282

- *Acinetobacter*

Deux espèces du gène *Acinetobacter* ont été identifiées dans tous les échantillons. C'est un pathogène commun avec une résistance élevée aux antibiotiques [98].

Tableau 10: Détection de deux espèces du gène *Acinetobacter*

taxonomy	<i>Acinetobacter;Acinetobacter_johnsonii</i>	<i>Acinetobacter;not_classified</i>
Comm_std-001-1	0	0

Comm_std-001-2	0	0
Comm_std-001-3	0	0
Comm_std-010-1	0	0
Comm_std-010-2	0	0
Comm_std-010-3	0	0
Comm_std-100-1	0	0
Comm_std-100-2	0	0
Comm_std-100-3	0	0
Neg_ctrl-1	0	0
Neg_ctrl-2	0	0
SEH-01F	104	1875
SEH-01L	103	1736
SEH-02F	4	80
SEH-02L	10	181
SEH-03F	0	6
SEH-03L	0	4
SEH-04F	0	5
SEH-04L	2	14
SEH-05F	22	316
SEH-05L	23	354
SEH-06F	65	750
SEH-06L	14	195
SEH-07F	0	18
SEH-07L	2	45
SEH-08F	7	86
SEH-08L	12	165
SEH-09F	0	0
SEH-09L	4	106
SEH-10F	0	3
SEH-10L	0	4

- *Stenotrophomonas*

Stenotrophomonas, trouvé dans l'échantillon 2, est le quatrième agent pathogène le plus courant parmi les bactéries gram-négatives non fermentantes (après *Pseudomonas aeruginosa*, *Acinetobacter* spp et le complexe *Burkholderia cepacia*) [99].

Tableau 11: Détection du *Stenotrophomonas*

taxonomy	<i>Stenotrophomonas</i> ;not_classified
----------	---

Comm_std-001-1	0
Comm_std-001-2	0
Comm_std-001-3	0
Comm_std-010-1	0
Comm_std-010-2	0
Comm_std-010-3	0
Comm_std-100-1	0
Comm_std-100-2	0
Comm_std-100-3	0
Neg_ctrl-1	0
Neg_ctrl-2	0
SEH-01F	0
SEH-01L	0
SEH-02F	42
SEH-02L	67
SEH-03F	0
SEH-03L	0
SEH-04F	0
SEH-04L	0
SEH-05F	0
SEH-05L	0
SEH-06F	0
SEH-06L	0
SEH-07F	0
SEH-07L	0
SEH-08F	0
SEH-08L	0
SEH-09F	0
SEH-09L	0
SEH-10F	0
SEH-10L	0

- *Sphingomonas*

Des espèces de *Sphingomonas* ont été identifiées dans tous les échantillons.

Tableau 12: Détection des espèces de *Sphingomonas*

taxonomy	<i>Sphingomonas</i> not_classified	<i>Sphingomonas</i> sp._35-24ZXX	<i>Sphingomonas</i> sp._IC081	<i>Sphingomonas</i> ;sp._PP-CE-1A- 5	<i>Sphingomonas</i> wittichii
----------	---------------------------------------	-------------------------------------	----------------------------------	--	----------------------------------

Comm_std-001-1	0	0	0	0	0
Comm_std-001-2	0	0	0	0	0
Comm_std-001-3	0	0	0	0	0
Comm_std-010-1	0	0	0	0	0
Comm_std-010-2	0	0	0	0	0
Comm_std-010-3	0	0	0	0	0
Comm_std-100-1	0	0	0	0	0
Comm_std-100-2	0	0	0	0	0
Comm_std-100-3	0	0	0	0	0
Neg_ctrl-1	0	0	0	2	0
Neg_ctrl-2	0	0	0	0	0
SEH-01F	2	14	7	0	2
SEH-01L	0	25	4	0	0
SEH-02F	5	51	102	1	2
SEH-02L	15	53	62	0	12
SEH-03F	151	302	10	6	28
SEH-03L	19	19	0	0	4
SEH-04F	198	295	11	7	34
SEH-04L	0	8	0	0	0
SEH-05F	69	16	0	0	6
SEH-05L	107	21	0	0	10
SEH-06F	679	55	1114	5	0
SEH-06L	118	58	59	0	0
SEH-07F	19	23	20	0	11
SEH-07L	20	63	10	2	6
SEH-08F	7	8311	303	107	2
SEH-08L	40	11072	518	196	10
SEH-09F	0	0	0	0	0
SEH-09L	10	13	2	0	0
SEH-10F	4	21	2	0	2
SEH-10L	2	1	0	0	0

- *Flavobacterium*

Plusieurs espèces de *Flavobacterium* étaient présentes dans les échantillons.

Tableau 13: Détection des espèces de *Flavobacterium*

taxonomy	<i>Flavobacterium</i> flevene	<i>Flavobacterium</i> fluvatile	<i>Flavobacterium</i> hydatis	<i>Flavobacterium</i> sp._GA093	<i>Flavobacterium</i> not_classified
----------	----------------------------------	------------------------------------	----------------------------------	------------------------------------	---

Comm_std-001-1	0	0	0	0	0
Comm_std-001-2	0	0	0	0	0
Comm_std-001-3	0	0	0	0	0
Comm_std-010-1	0	0	0	0	0
Comm_std-010-2	0	0	0	0	0
Comm_std-010-3	0	0	0	0	0
Comm_std-100-1	0	0	0	0	0
Comm_std-100-2	0	0	0	0	0
Comm_std-100-3	0	0	0	0	0
Neg_ctrl-1	0	0	0	0	0
Neg_ctrl-2	0	0	0	0	0
SEH-01F	14902	1543	417	503	50089
SEH-01L	24205	2615	645	968	78417
SEH-02F	580	58	14	18	1869
SEH-02L	2018	184	52	66	6842
SEH-03F	27	0	0	2	110
SEH-03L	12	0	0	0	42
SEH-04F	28	2	2	0	101
SEH-04L	28	8	0	0	115
SEH-05F	25	2	0	0	123
SEH-05L	304	28	4	6	910
SEH-06F	6	20	0	1	393
SEH-06L	10	206	0	17	1654
SEH-07F	32	3	0	0	91
SEH-07L	12	0	0	0	59
SEH-08F	22	0	0	0	53
SEH-08L	15	2	0	0	79
SEH-09F	2	0	0	0	6
SEH-09L	50	3	0	4	645
SEH-10F	1	0	1	0	122
SEH-10L	0	2	0	0	26

Flavobacterium est un genre de bactéries Gram-négatives, non mobiles et mobiles, en forme de bâtonnet, qui se compose de 130 espèces reconnues [100].

- *Caulobacter*

Différentes espèces de *Caulobacter* ont été identifiées dans les échantillons, en particulier *Caulobacter_vibrioides*.

Tableau 14: Détection de *Caulobacter_vibrioides*

taxonomy	Caulobacter_vibrioides
Comm_std-001-1	0
Comm_std-001-2	0
Comm_std-001-3	0
Comm_std-010-1	0
Comm_std-010-2	0
Comm_std-010-3	0
Comm_std-100-1	0
Comm_std-100-2	0
Comm_std-100-3	0
Neg_ctrl-1	0
Neg_ctrl-2	0
SEH-01F	1
SEH-01L	0
SEH-02F	5
SEH-02L	1
SEH-03F	1240
SEH-03L	23
SEH-04F	1485
SEH-04L	0
SEH-05F	3
SEH-05L	2
SEH-06F	260
SEH-06L	251
SEH-07F	4
SEH-07L	1
SEH-08F	2
SEH-08L	11
SEH-09F	0
SEH-09L	16
SEH-10F	3
SEH-10L	0

- *Nitrosomonas et Nitrospira*

Nitrosomonas et *Nitrospira* ont été identifiés dans les échantillons 3, 4 et 7. Ce sont des bactéries nitrifiantes considérées comme des microorganismes nuisibles [101], [102].

Tableau 15: Détection de Nitrosomonas et Nitrospira

taxonomy	Nitrospira Nitrospira sp._CG24E	Nitrospira Not classified	Nitrosomonas mobilis	Nitrosomonas sp._Nm143	Nitrosomonas not_classified
Comm_std-001-1	0	0	0	0	0
Comm_std-001-2	0	0	0	0	0
Comm_std-001-3	0	0	0	0	0
Comm_std-010-1	0	0	0	0	0
Comm_std-010-2	0	0	0	0	0
Comm_std-010-3	0	0	0	0	0
Comm_std-100-1	0	0	0	0	0
Comm_std-100-2	0	0	0	0	0
Comm_std-100-3	0	0	0	0	0
Neg_ctrl-1	0	0	0	0	0
Neg_ctrl-2	0	0	0	0	0
SEH-01F	0	0	0	0	2
SEH-01L	0	0	0	0	0
SEH-02F	0	0	0	0	0
SEH-02L	0	0	0	0	0
SEH-03F	57	102	0	74	121
SEH-03L	0	0	0	0	14
SEH-04F	55	118	0	81	206
SEH-04L	186	444	0	335	0
SEH-05F	14	6	0	4	0
SEH-05L	0	0	0	0	0
SEH-06F	0	2	0	0	0
SEH-06L	0	0	0	0	2
SEH-07F	18	24	56	9	0
SEH-07L	7	28	92	7	0
SEH-08F	0	4	0	0	2
SEH-08L	0	4	0	0	1
SEH-09F	0	0	0	0	0
SEH-09L	1	12	0	0	0
SEH-10F	0	0	0	0	0
SEH-10L	0	0	0	0	0

- *Streptomyces*

Streptomyces, trouvé dans l'échantillon 5, est un microorganisme nuisible parmi les actinomycètes [103].

Tableau 16: Détection des Streptomyces

taxonomy	Bacteria;Actinobacteria;Actinobacteria;Streptomycetales; Streptomycetaceae;Streptomyces;Streptomyces_sp._SM14
Comm_std-001-1	0
Comm_std-001-2	0
Comm_std-001-3	0
Comm_std-010-1	0
Comm_std-010-2	0
Comm_std-010-3	0
Comm_std-100-1	0
Comm_std-100-2	0
Comm_std-100-3	0
Neg_ctrl-1	0
Neg_ctrl-2	0
SEH-01F	0
SEH-01L	0
SEH-02F	0
SEH-02L	0
SEH-03F	0
SEH-03L	0
SEH-04F	0
SEH-04L	0
SEH-05F	73
SEH-05L	72
SEH-06F	0
SEH-06L	0
SEH-07F	6
SEH-07L	0
SEH-08F	0
SEH-08L	0
SEH-09F	0
SEH-09L	0
SEH-10F	0
SEH-10L	0

En plus des 4 espèces de virus non identifiées suivantes:

Virus; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; not_classified;
not_classified

Virus; Uroviricota; Caudoviricetes; Caudovirales; Podoviridae; Kafunavirus;
not_classified

Virus; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; not_classified;
not_classified

Les virus identifiés dans les échantillons d'eau sont:

- *Prokaryotic_dsDNA_virus_sp*
- *Klebsiella_virus_RaK2*
- *Achromobacter_phage_Motura*
- *Sphingobium_virus_Lacusarx*
- *Croceibacter_phage_P2559Y*
- *Xanthomonas_phage_Xp15*

3.2.2.2. Analyse alpha diversité Shannon

Nous présentons la diversité alpha de Shannon dans la figure suivante :

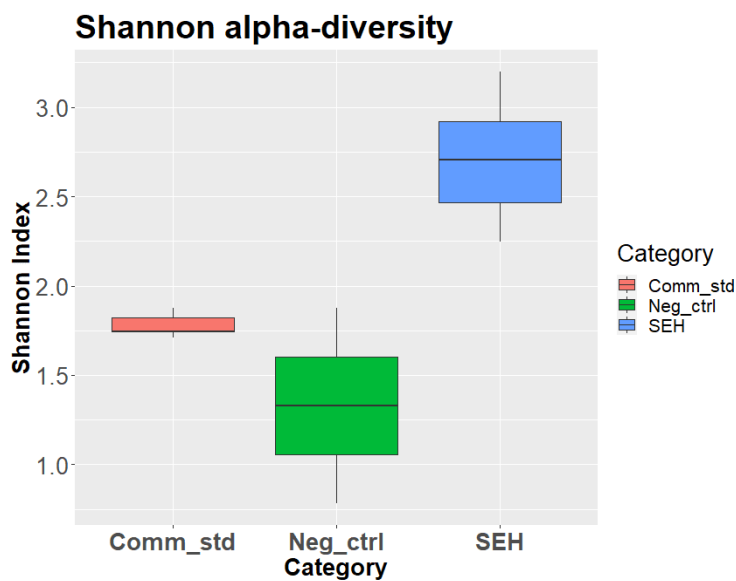


Figure 25: Shannon alpha diversité

Parce que la valeur $p = 0,000148$ qui est inférieure au niveau de signification 0,05, nous avons suffisamment de preuves pour rejeter l'hypothèse nulle que les moyennes de la population sont toutes égales.

3.2.2.3. Analyse bêta diversité NMDS

Le graphique NMDS (Non Metric Multidimensional Scaling) de la diversité bêta est présenté dans la figure suivante. Dans le graphique NMDS, chaque point représente le microbiome d'un échantillon.

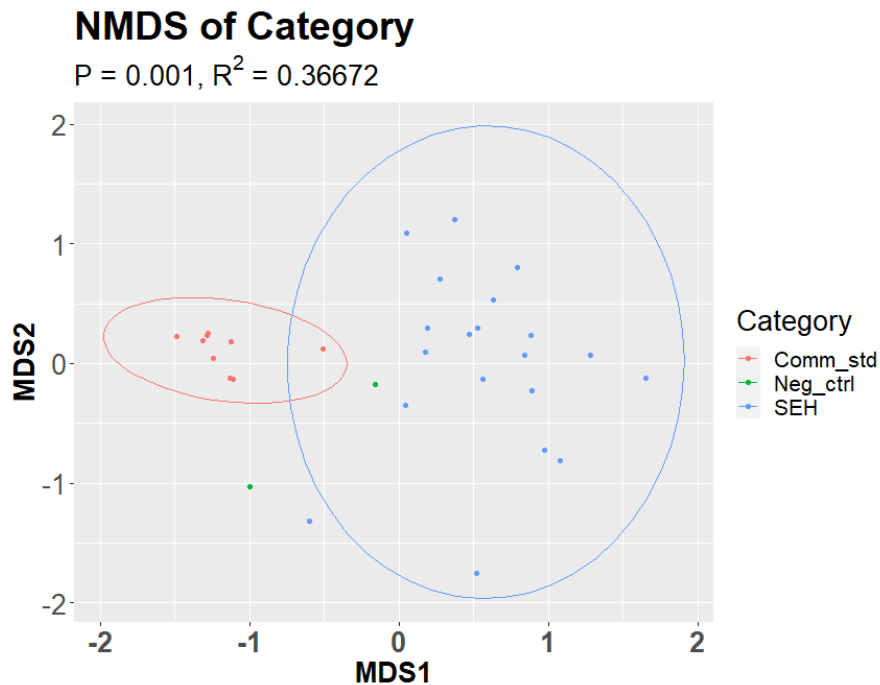


Figure 26: NMDS bêta diversité

3.3. Résultats des analyses du microbiome intestinal

3.3.1. Quantification de l'ADN extraits des échantillons des selles

3.3.1.1. Résultats du logiciel Gen5

Les concentrations de l'ADN sont reportées sur Excel et représentés dans le tableau suivant :

Tableau 17 : Tableau des concentrations d'ADN dans les échantillons

Sample #	RFU	RFU - blank	Volume of Sample Used (μL)	Well conc (ng/ml)	Actual Conc (ng/ml)
1	8295	8281	2	792.69806	99087.2575
2	6763	6749	2	644.55366	80569.2075

3	8614	8600	2	823.54536	102943.17
4	8477	8463	2	810.29746	101287.1825
5	5218	5204	2	495.15216	61894.02
6	3899	3885	2	367.60486	45950.6075
7	3387	3373	2	318.09446	39761.8075
8	4156	4142	2	392.45676	49057.095
9	3611	3597	2	339.75526	42469.4075
10	3967	3953	2	374.18046	46772.5575
11	3651	3637	2	343.62326	42952.9075
12	2062	2048	2	189.96696	23745.87
13	5	-9	2	-8.94494	-1118.1175
14	2979	2965	2	278.64086	34830.1075
15	3016	3002	2	282.21876	35277.345
16	4791	4777	2	453.86126	56732.6575
17	5004	4990	2	474.45836	59307.295
18	7	-7	2	-8.75154	-1093.9425
19	3706	3692	2	348.94176	43617.72
20	5478	5464	2	520.29416	65036.77
21	5	-9	2	-8.94494	-1118.1175
22	6	-8	2	-8.84824	-1106.03
23	4450	4436	2	420.88656	52610.82
24	2597	2583	2	241.70146	30212.6825
25	4921	4907	2	466.43226	58304.0325
26	3659	3645	2	344.39686	43049.6075
27	1004	990	2	87.65836	10957.295
28	3042	3028	2	284.73296	35591.62
29	470	456	2	36.02056	4502.57
30	5	-9	2	-8.94494	-1118.1175
31	3296	3282	2	309.29476	38661.845
32	1624	1610	2	147.61236	18451.545
33	742	728	2	62.32296	7790.37
34	6	-8	2	-8.84824	-1106.03
35	713	699	2	59.51866	7439.8325
36	859	845	2	73.63686	9204.6075
37	5	-9	2	-8.94494	-1118.1175
38	4138	4124	2	390.71616	48839.52
39	45	31	2	-5.07694	-634.6175
40	123	109	2	2.46566	308.2075
41	40	26	2	-5.56044	-695.055
42	33	19	2	-6.23734	-779.6675

43	5	-9	2	-8.94494	-1118.1175
44	9	-5	2	-8.55814	-1069.7675
45	8	-6	2	-8.65484	-1081.855
46	7	-7	2	-8.75154	-1093.9425
47	5480	5466	2	520.48756	65060.945
48	2318	2304	2	214.72216	26840.27
49	33	19	2	-6.23734	-779.6675
50	3983	3969	2	375.72766	46965.9575
51	6	-8	2	-8.84824	-1106.03
52	85	71	2	-1.20894	-151.1175
53	15	1	2	-7.97794	-997.2425
54	5245	5231	2	497.76306	62220.3825
55	457	443	2	34.76346	4345.4325
56	20	6	2	-7.49444	-936.805
57	227	213	2	12.52246	1565.3075
58	7	-7	2	-8.75154	-1093.9425
59	2	-12	2	-9.23504	-1154.38
60	6	-8	2	-8.84824	-1106.03
61	6	-8	2	-8.84824	-1106.03
62	317	303	2	21.22546	2653.1825
63	2737	2723	2	255.23946	31904.9325
64	104	90	2	0.62836	78.545
65	4852	4838	2	459.75996	57469.995
66	4424	4410	2	418.37236	52296.545
67	2760	2746	2	257.46356	32182.945
68	4672	4658	2	442.35396	55294.245
69	4698	4684	2	444.86816	55608.52
70	479	465	2	36.89086	4611.3575
71	1804	1790	2	165.01836	20627.295
72	4526	4512	2	428.23576	53529.47
73	19	5	2	-7.59114	-948.8925
74	1087	1073	2	95.68446	11960.5575
75	3272	3258	2	306.97396	38371.745
76	1995	1981	2	183.48806	22936.0075
77	1661	1647	2	151.19026	18898.7825
78	493	479	2	38.24466	4780.5825
79	540	526	2	42.78956	5348.695
80	5429	5415	2	515.55586	64444.4825

3.3.1.2. Résultats du Gel d'agarose

Voir toutes les images du gel en annexe.

Les concentrations d'ADN ont été uniformément ajustées à 2 ng / ul.

(1 ng / ul pour High, 2 ng / ul pour Medium et 3 ng / ul pour Low)

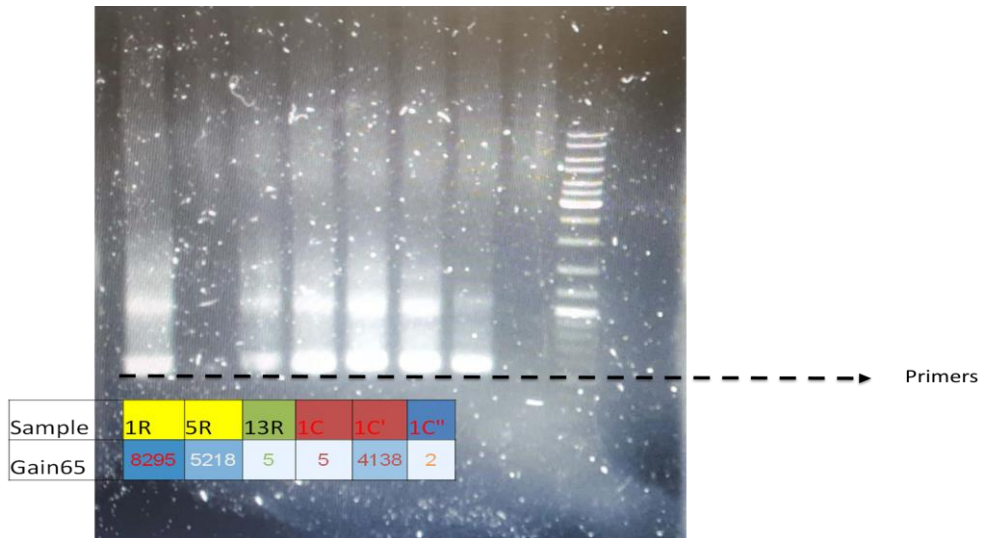
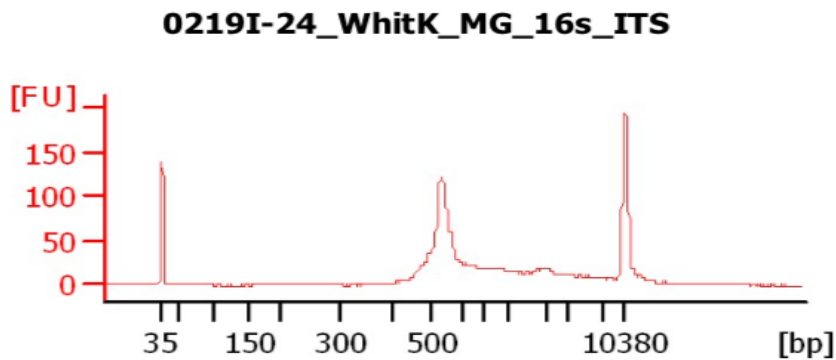


Figure 27: Image des résultats du gel de six échantillons

3.3.1.3. Traces de bio-analyseur

Les réactions de PCR groupées ont été analysées par un bio-analyseur pour confirmer une taille de bande d'environ 500 pb.

Les bibliothèques générées ont été chargées sur un Illumina Next-Seq à une concentration de 1,8 picomolaire à l'aide du kit à mi-sortie d'Illumina pour un séquençage par paires de 75 pb.



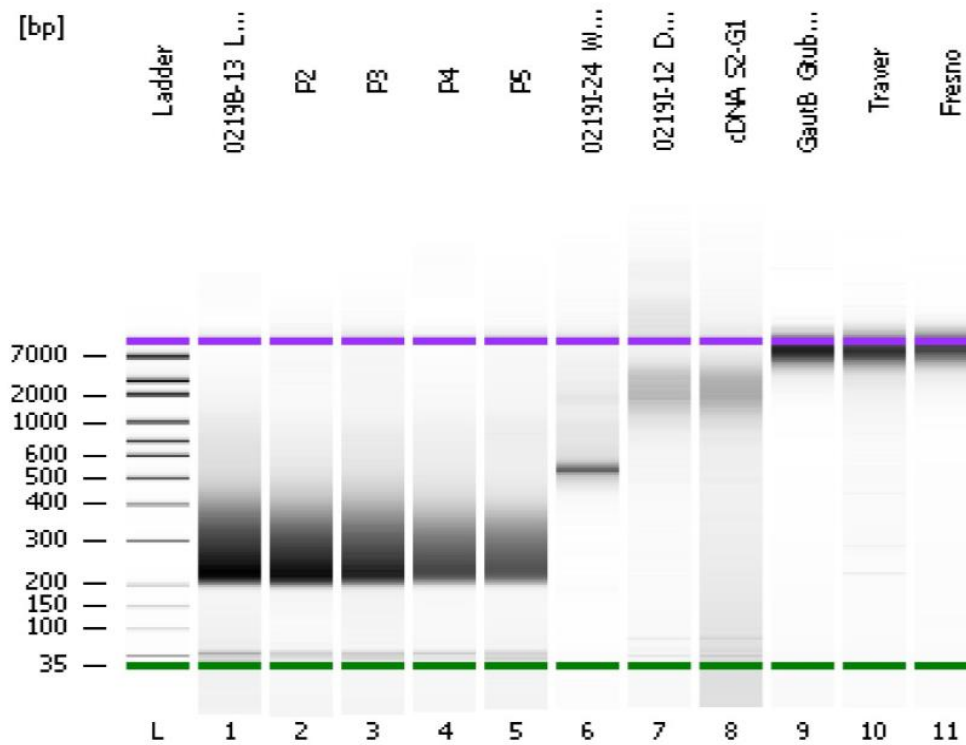


Figure 28: Les traces du bio-analyseur

3.3.2. Résultats du séquençage « Shotgun »

3.3.2.1. Analyse d'abondance relative

Les abondances relatives de la figure suivante montrent les 10 premières classifications taxonomiques des communautés dans les échantillons, évaluées par Metaphlan, dans les trois groupes des habitants marocains. Nous pouvons voir sur le graphique les différences significatives entre les trois cohortes.

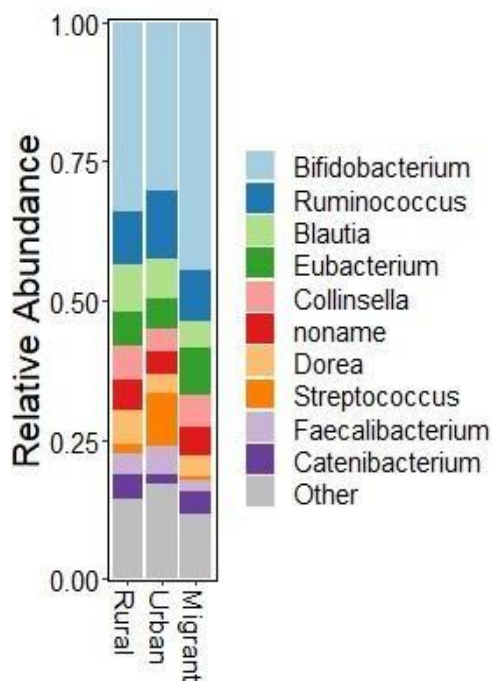


Figure 29: 10 premières classifications taxonomiques des communautés dans les échantillons

Nous avons constaté que *Bifidobacterium* suivi par *Ruminococcus* était le plus abondant dans tous les échantillons.

Blautia et *Faecalibacterium* étaient abondants chez les participants marocains ruraux et urbains.

Eubacterium était plus abondante dans les échantillons de migrants et *Dorea* dans les échantillons ruraux.

Les individus ruraux et migrants avaient des portions plus élevées de *Collinsella* et de *Catenibacterium*. Des études montrent que *Catenibacterium* est lié à l'emplacement géographique, car nous considérons que les participants migrants viennent de différentes zones rurales [104].

De plus, nous pouvons voir une abondance significative de *Streptococcus* dans les échantillons urbains.

3.3.2.2. Analyse des champignons et d'un bactériophage

Les abondances des champignons et d'un bactériophage pour les trois cohortes telles que déterminées par séquençage « Shotgun » sont représentées dans la figure suivante :

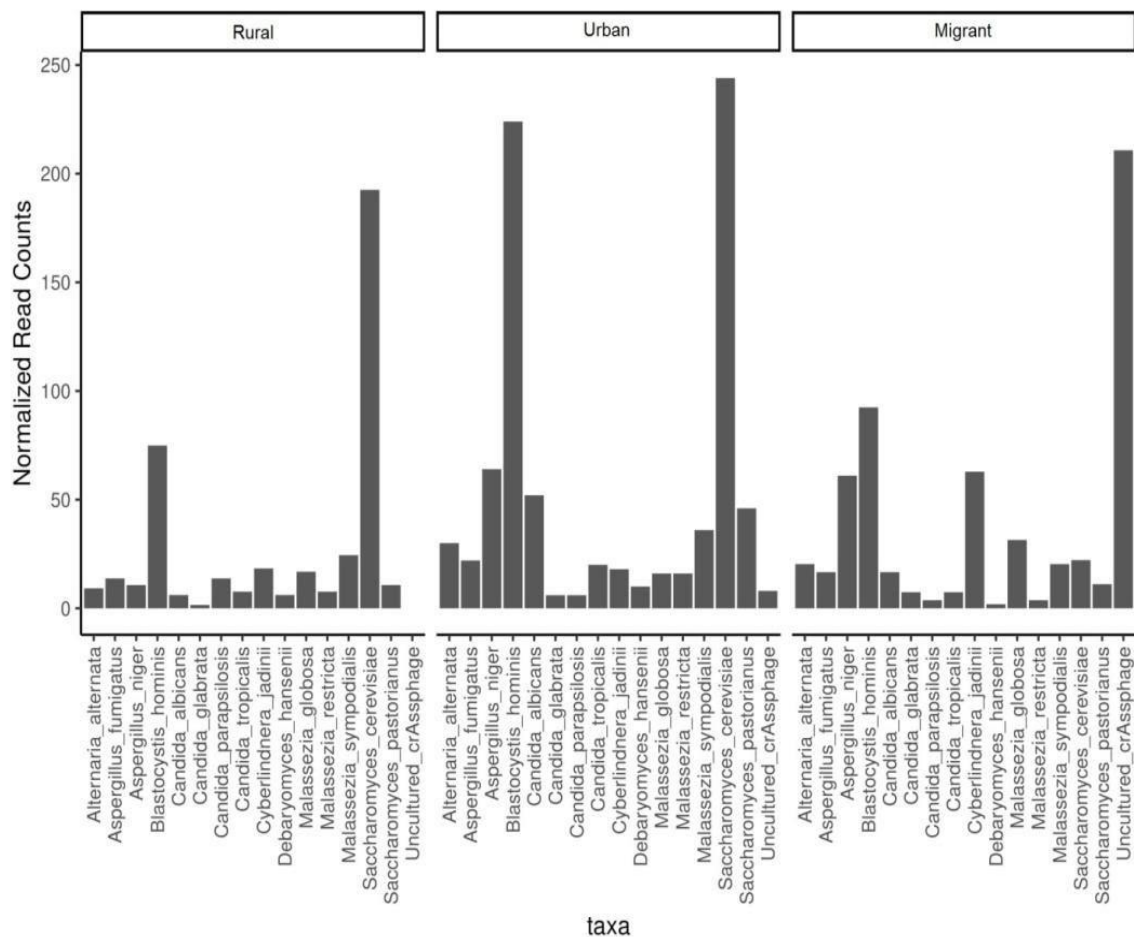


Figure 30: Abondances de champignons et d'un bactériophage pour les trois cohortes déterminées par séquençage « Shotgun ».

Nous avons identifié les 3 espèces suivantes:

La *blastocystis* était plus abondante dans les échantillons urbains.

Saccharomyces était présent chez les sujets marocains des zones rurales et urbaines.

Le *crAssphage* était présent en abondance dans les échantillons de migrants.

Ces résultats reflètent la difficulté de trouver des agents pathogènes dans l'environnement dense et complexe des matières fécales.

3.3.3. Résultats du séquençage de l'ARNr des bactéries 16S et des champignons ITS

Pour mieux comprendre le microbiome intestinal des habitants du Maroc, nous avons réalisé un séquençage 16S et ITS pour tous les échantillons.

Après séquençage d'amplicon 16S et ITS de l'ADN extrait d'échantillons de selles, les séquences ont été regroupées à ~ 97% de similitude pour identifier les OTU (unités taxonomiques opérationnelles).

Dans les figures 32 et 33, nous avons présenté la distribution des abondances relatives par cohorte (30 participants ruraux, 15 urbains et 35 migrants) des 10 principales espèces bactériennes significatives pour 16S et des champignons pour les ITS.

3.3.3.1. Analyse d'abondance relative 16S/ITS par individu

Analyse d'abondance relative 16S par individu :

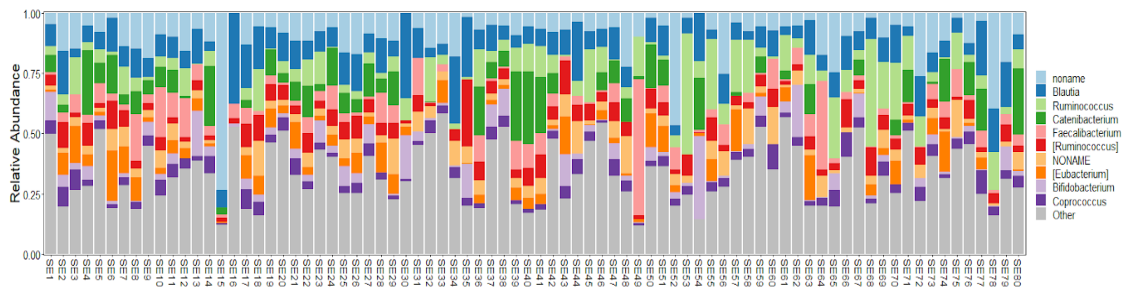


Figure 31: Distribution des abondances relatives par cohorte des 10 principales espèces bactériennes significatives

Analyse d'abondance relative ITS par individu :

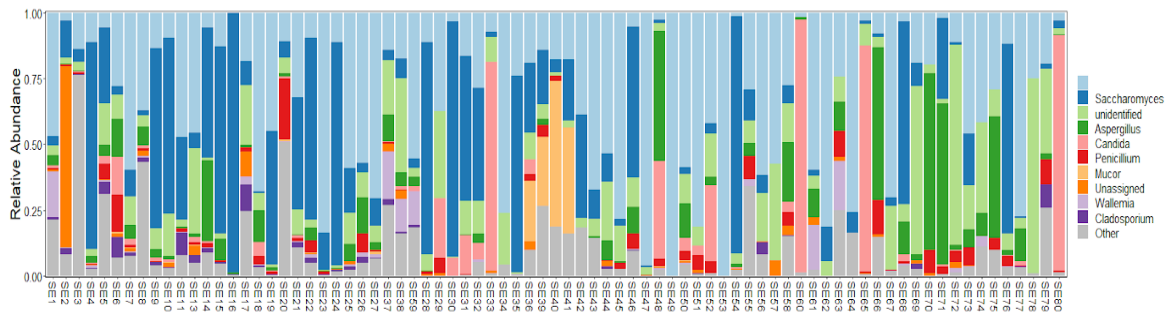


Figure 32: Distribution des abondances relatives par cohorte des 10 principaux champignons

3.3.3.2. Analyse d'abondance relative 16S/ITS par cohorte

- Analyse d'abondance relative 16S par cohorte :

Les compositions microbiennes globales des individus vivant au Maroc, représentées par 35 échantillons rurales et 45 échantillons urbaines, ont été analysées avec les séquences d'ARNr 16S disponibles.

Les graphiques à barres initiaux comparant les abondances relatives des espèces microbiennes entre les individus témoins (c'est-à-dire urbains) et les individus ruraux n'ont montré aucune différence significative au niveau du phylum (Figure 34).

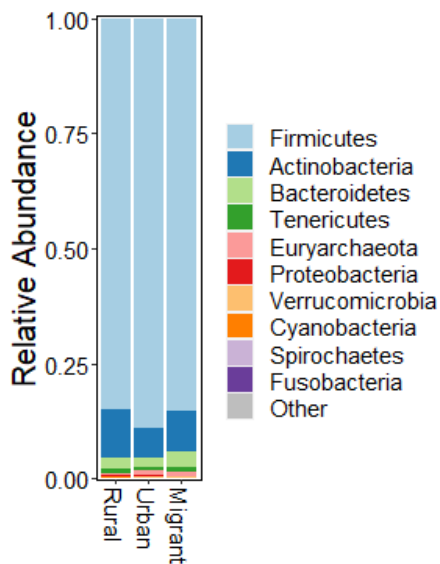


Figure 33: Abondance relative au niveau du phylum

Fait intéressant, les phyla Firmicutes, Actinobacteria et Bacteriodes étaient tous en abondance significative par rapport aux autres phylums

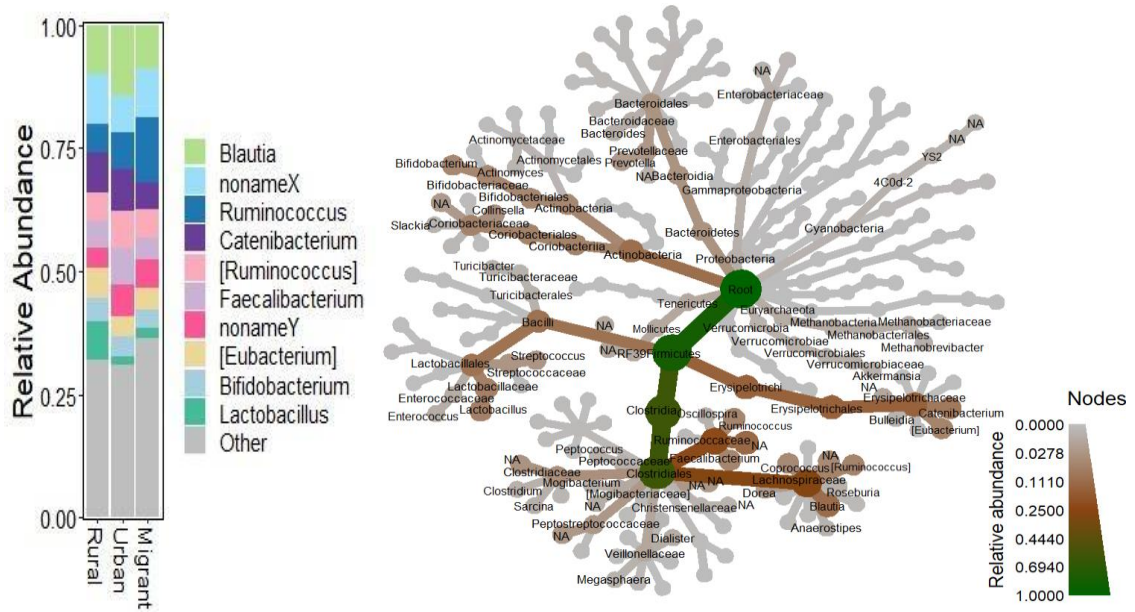


Figure 34 : a) Abondances relatives moyennes des 10 genres bactériens les plus abondants dans les trois cohortes. «Other» indique les genres qui ne figuraient pas dans le top 10 des plus abondants. b) Un arbre 16S montrant la diversité taxonomique au sein de la cohorte rurale. L'abondance relative de chaque taxon est indiquée par la couleur comme indiqué dans la clé.

- Analyse d'abondance relative ITS par cohorte

Les champignons présents dans les échantillons de selles des individus vivant au Maroc, représentées par 35 échantillons rurales et 45 échantillons urbaines, ont été analysés avec les séquences d'ARNr ITS disponibles.

Les graphiques à barres initiaux comparant les abondances relatives des champignons entre les individus témoins (c'est-à-dire urbains) et les individus ruraux ont montré des différences significatives.

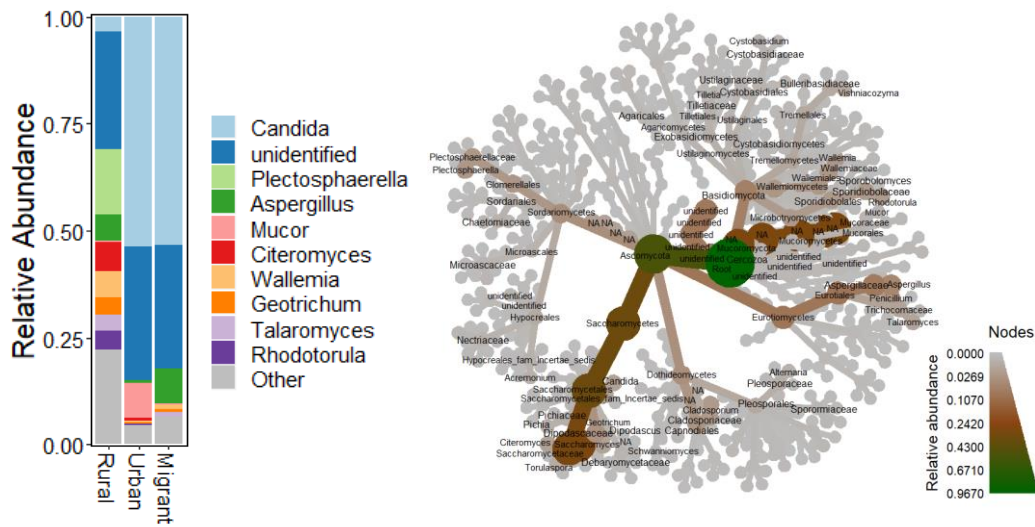


Figure 35: a) Abondances relatives moyennes des 10 genres de champignons les plus abondants dans les trois cohortes. «Other» indique les genres qui ne figuraient pas dans le top 10 des plus abondants. b) Un arbre ITS montrant la diversité taxonomique au sein de la cohorte rurale

3.3.3.3. Analyse alpha diversité 16S/ITS

Nous présentons la diversité alpha de Shannon dans la figure 37, où nous ne voyons pas de différences significatives entre les trois cohortes. Une richesse plus élevée est observée dans les échantillons ruraux.

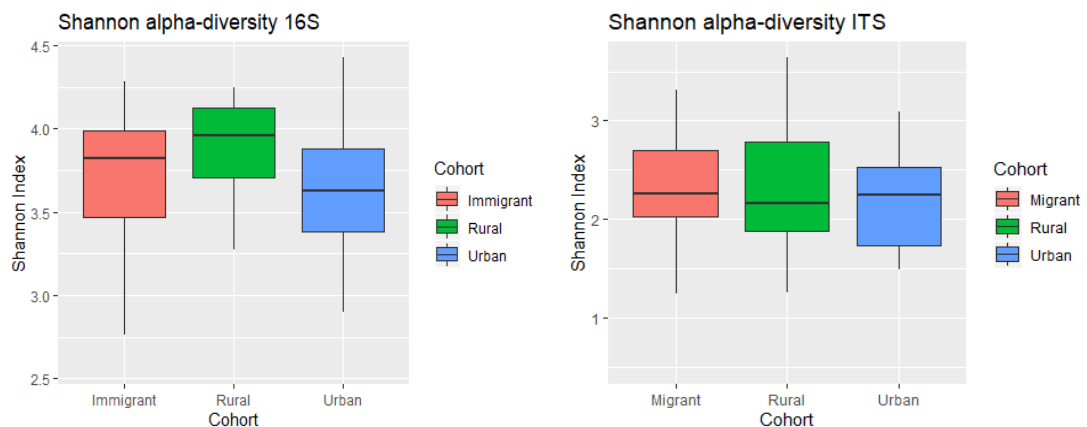


Figure 36: Diagrammes en boîtes montrant la diversité alpha de chaque cohorte estimée à l'aide de l'indice de Shannon. a) montre des estimations pour les amplicons 16S et b) montre des estimations pour les amplicons ITS. Chaque case montre la médiane (ligne horizontale) et l'intervalle interquartile. Les différences n'étaient pas statistiquement significatives (ANOVA; $p_{16S} = 0,387$; $p_{ITS} = 0,527$).

Les échantillons de selles n'étaient pas clairement distingués par les valeurs de diversité alpha, comme le montre la figure 37.

Les mesures de la diversité alpha n'ont montré aucune différence notable entre les trois groupes.

Pour confirmer les résultats, le Microbiome Analyst Website a été utilisé et sont représentés dans la figure 38 suivante.

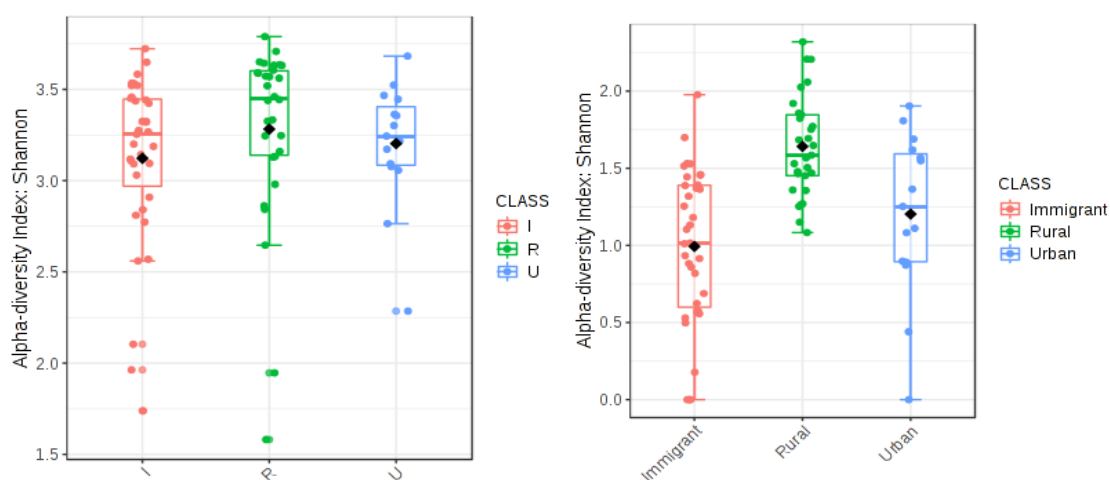


Figure 37: Analyse alpha diversité de 16S à gauche et ITS à droite par le Microbiome Analyst Website

3.3.3.4. Analyse beta diversité 16S/ITS

Les graphiques de diversité bêta ont montré les regroupements des trois groupes. Des graphiques à échelle multidimensionnelle non métrique (NMDS) sont présentés aux figures 39 et 40 pour déterminer la diversité bêta. Dans les graphiques NMDS, chaque point représente le microbiome intestinal d'un individu. Des différences statistiques ont été observées entre les groupes avec une valeur p significative <0,05 dans l'analyse 16S et ITS.

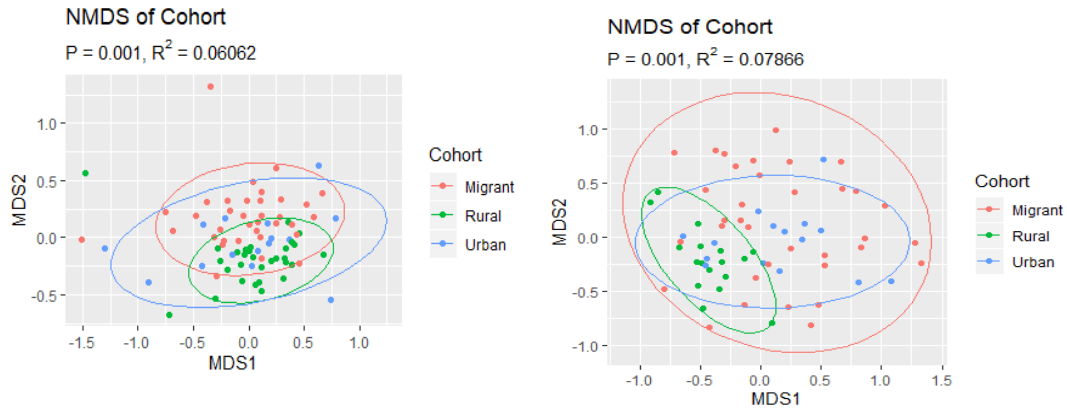


Figure 38: Tracés NMDS des différences de Bray-Curtis pour a) les amplicons 16S et b) les amplicons ITS. Les ellipses représentent les intervalles de confiance à 95% pour chaque cohorte

Pour confirmer ces résultats, le Microbiome Analyst Website a été utilisé.

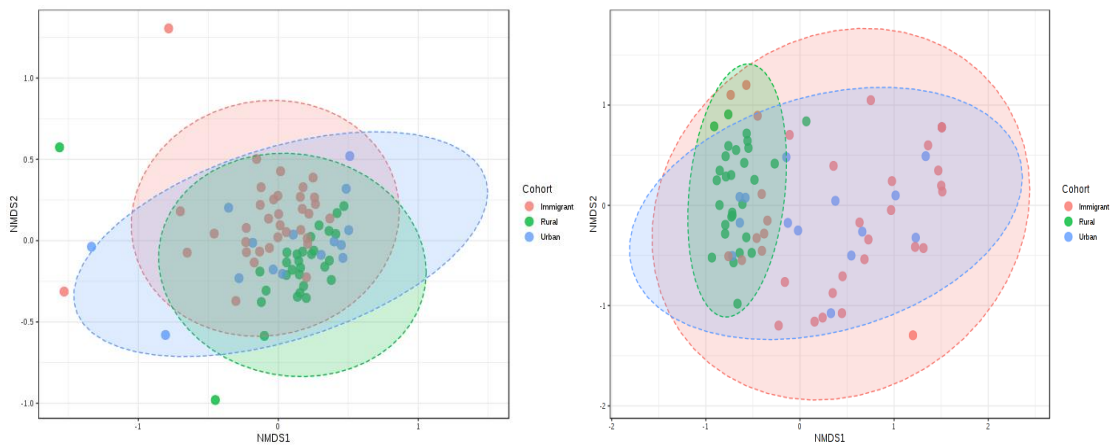


Figure 39: Analyse alpha diversité de 16S à gauche et ITS à droite par le Microbiome Anayst Website

On observe que le microbiome intestinal des individus vivant dans les zones rurales est plus étroitement regroupé que les échantillons urbains.

Discussion

Impact des fosses septiques sur la nappe phréatique :

Les fosses septiques contribuent à la pollution diffuse par l'introduction directe de contaminants tels que les bactéries pathogènes, les nutriments et la matière organique, et indirectement par les apports d'eau douce, ce qui peut entraîner une augmentation du transport des contaminants et/ou de la survie des agents microbiens.

Les faibles profondeurs de la nappe phréatique et les substrats sableux perméables des sites d'étude représentent un environnement à haut risque de contamination des eaux souterraines par l'évacuation des eaux usées domestiques. De nombreuses études ont signalé une contamination microbienne des eaux souterraines peu profondes et le transport de bactéries coliformes dans l'environnement souterrain saturé [105]. Les différentes espèces pathogènes identifiées dans les échantillons des eaux des puits confirment la non potabilité des ces eaux.

Contamination des eaux des puits :

Nous constatons d'après cette étude que les eaux des puits étudiées présentent une contamination bactérienne. La nature des germes rencontrés, rendent l'eau des puits étudiés impropre à la consommation. Cette situation pourrait provoquer la propagation de nombreuses maladies infectieuses et parasitaires à transports hydriques notamment : le choléra, la typhoïde, l'hépatite, dysenterie, gastro-entérite, maladies diarrhéiques, la bilharziose et le paludisme [106], [107]. Ce travail s'est déroulé pendant la saison des pluies. Quelle soit chimique ou bactériologique, la pollution est accentuée pendant la saison pluvieuse. Cela s'explique par la montée générale du niveau de la nappe par une plus forte dilution à cause de la recharge de la nappe d'une part, et d'autre part à cause de l'exposition directe des puits aux eaux de ruissellements qui viennent y tomber chargées de nombreuses substances polluantes. Les causes de cette pollution sont multiples; parmi lesquelles nous pouvons citer: la mauvaise protection des puits, la non application des mesures d'hygiène élémentaires, la mauvaise conception des latrines et des fosses septiques, la mauvaise évacuation des eaux usées et la présence des sources de pollution, tel que les dépôts des ordures à proximité des puits étudiés.

Solutions proposées :

Pour remédier à cette situation et afin de mettre les populations à l'abri des affections liées à l'eau de consommation, nous recommandons aux autorités locales les solutions suivantes :

- Sensibiliser la population utilisatrice de cette eau et l'inciter à traiter l'eau des puits avant consommation avec des pastilles d'hypochlorite de sodium, et à adopter des comportements favorables à la préservation et à la potabilisation de l'eau depuis la source d'approvisionnement jusqu'à sa consommation.

- Une éducation sanitaire au profit des populations en encourageant la sensibilisation par les Medias.

- Un contrôle permanent des puits et la réalisation d'une étude spatio-temporelle de la qualité hygiénique de l'ensemble des puits recensés et utilisés par la population comme eau de boisson et/ou pour des activités domestiques.

- Exiger le respect de la distance entre puits et fosse (une distance minimum de Vingt cinq mètres) et appliquer les mesures et les règles relatives à une meilleure conception des fosses septiques et des puisards. Ce qui diminuera la pollution de la nappe superficielle.

- Mettre en place un système d'évacuation correcte des eaux usées.

- Exiger le respect des normes de protection des puits notamment le revêtement intérieur du puits qui doit être étanche et le revêtement extérieur qui comprend généralement : Une margelle : c'est un équipement fait en surélévation de l'ouverture du puits. Il a pour but d'arrêter tout ce qui doit passer par l'ouverture du puits pour contaminer l'eau. La dalle de couverture : elle est en béton armé et de dimensions variables suivant le diamètre du puits. Elle repose sur la margelle et est conforme à ses dimensions.

La plate-forme : cet aménagement est réalisé tout au tour du puits. Cette aire doit être inclinée du puits vers l'extérieure. Une rigole doit la ceinturer et drainer toutes les eaux qui s'écoulent vers un puisard à quelques mètres du puits.

L'implantation, la construction et la surveillance d'un système d'alimentation en eau doivent être de rigueur afin de minimiser tout risque de pollution et d'atteinte à la santé humaine.

Le microbiome intestinal des sujets vivant au Maroc :

Le microbiome d'un intestin sain est dominé par : les Bacteroidetes, les Firmicutes, les Proteobacteria et les Actinobacteria. Dans cette analyse, nous avons comparé des échantillons de selles d'habitants ruraux avec ceux de deux cohortes de sujets urbains. Les comparaisons de la diversité alpha ont indiqué que les trois cohortes avaient des richesses en espèces similaires.

Il s'agit d'une découverte intéressante qui signifie que les zones urbaines et rurales et les différences dans les sources d'eau potable ne modifient pas de manière significative la composition taxonomique du microbiome intestinal. Cela pourrait s'expliquer par la qualité de vie et l'alimentation similaires pour tous les participants.

Les analyses du NMDS ont indiqué que les microbiomes intestinaux des individus vivant dans les zones rurales se regroupaient plus étroitement les uns avec les autres que les microbiomes des échantillons urbains. Cependant, la cohorte n'expliquait que 6 % de la variation dans la composition de la communauté et toutes les ellipses de l'intervalle de confiance se chevauchaient. Les similitudes entre les microbiomes des individus ruraux étaient inattendues car les individus des zones rurales consommaient de l'eau provenant de différentes sources non traitées, ce qui, selon nous, aurait des influences personnalisées sur leurs microbiomes. Les régimes alimentaires des individus ruraux différaient également davantage les uns des autres que ceux des individus ruraux et migrants.

Le microbiome intestinal et la diversité des populations :

De nombreuses études ont démontré que le microbiote intestinal a des impacts importants sur la santé humaine et sur les processus favorisant le développement de la maladie. Cependant, nous manquons encore d'une compréhension globale de l'ensemble des informations génétiques et de la diversité fonctionnelle du microbiome intestinal [19], [23], [108]. Contrairement aux méthodes conventionnelles d'isolement et de culture de clones microbiens purs, les approches d'amplicon d'ARNr 16S et de

séquençage métagénomique se concentrent sur l'ensemble de la communauté microbienne et leurs interactions.

Les résultats affirment que la population au Maroc a un microbiome intestinal semblable et non affecté par la qualité de l'eau de consommation. Pour aller très loin, nous devons comparer le microbiome intestinal de la population marocaine avec une autre population. Dans ce contexte, une étude compare les microbiomes intestinaux d'une population chinoise à une population danoise [41], ils ont essayé d'identifier les différences de composition de la communauté microbienne et les caractéristiques fonctionnelles ou métaboliques, et ont trouvé un certain nombre de caractéristiques géographiques importantes du microbiote intestinal. En comparant davantage ces caractéristiques spécifiques au pays avec les habitats vivants et les facteurs environnementaux, ils ont constaté que la plupart des caractères spécifiques au pays étaient affectés par les habitats vivants et les facteurs environnementaux. Par exemple, en comparant les principales différences entre les composants de la population danoise et chinoise, ils ont constaté que les Danois abritaient plus de microbes du phylum des Firmicutes, tels que *Oenococcus* et *Lactobacillus*, tandis que plus de microbes des phylums des Proteobacteria ont été détectés dans la population chinoise.

Association entre le microbiome intestinal et les maladies :

Le microbiote intestinal joue un rôle essentiel dans la santé humaine. Cependant, les recherches sur ce sujet ont été limitées par des études basées sur la méthode de l'ARNr 16S et la taille de l'échantillon est généralement limitée dans les études anciennes [109]–[112]. Suivant le principe des études d'association à l'échelle du génome (Genome Wide Association Studies GWAS), qui ont été utilisées avec succès pour identifier de nombreuses mutations génétiques associées à des maladies, **MGWAS** devrait découvrir des différences de niveau taxonomique et des **marqueurs** génétiques en association avec des maladies. Qin et al. a proposé pour la première fois le MGWAS en 2012 et ils ont réalisé l'étude des microbiomes intestinaux sur des patients atteints de Diabète Type 2 DT2 à l'aide du séquençage métagénomique Shotgun. Généralement, deux étapes ont été appliquées dans les MGWAS.

Tout d'abord, ils ont construit un catalogue de gènes humains et ils ont recherché les gènes du microbiote intestinal distribués différemment entre les patients

DT2 et les témoins sains. Lors de la deuxième étape, 200 échantillons d'une cohorte de test ont été utilisés pour valider ces gènes marqueurs de la maladie identifiés. Il s'agissait de la première étude utilisant le MGWAS pour évaluer la corrélation entre les microbes intestinaux et les maladies humaines. Il a non seulement fourni l'association taxonomique et fonctionnelle détaillée entre les microbes intestinaux et le DT2, mais a également clairement démontré le fait attrayant que le microbiome intestinal avait le potentiel d'être utilisé dans le diagnostic du DT2 ou pour prédire le risque de maladie. Cette étude a été suivie par des MGWAS du microbiome intestinal ciblant d'autres maladies, notamment la cirrhose du foie et le Cancer Colorectal CCR [113].

En 2015, plusieurs études ont examiné l'association entre les profils taxonomiques et fonctionnels du microbiome intestinal et la séquence adénome-carcinome colorectal [114], la polyarthrite rhumatoïde (PR) et les maladies cardiovasculaires (MCV). Pour suivre le lien entre les microbes intestinaux et le développement du CRC, Feng et al. ont analysé les métagénomés d'échantillons fécaux de 44 patients atteints d'adénome colorectal, 46 patients atteints de CCR et 57 témoins sains. Une analyse de corrélation des facteurs de risque potentiels et des microbes intestinaux a indiqué qu'une consommation élevée de viande rouge semblait être associée à une excroissance de bactéries qui pourraient contribuer à un environnement intestinal plus hostile, tandis que la consommation de fruits et de légumes était négativement corrélée à cette tendance. Ce résultat suggérait la possibilité d'expliquer le lien entre régime alimentaire spécifique et le CCR par l'intermédiaire des microbes intestinaux. Cette étude a fait progresser la compréhension des liens entre les microbes intestinaux/oraux/de la peau et les maladies.

Les microbes liés aux maladies pourraient être les cibles potentielles de traitement. Sur la base de la connaissance du lien entre l'alimentation et le microbiome, des aliments fonctionnels pourraient être développés pour prévenir ou traiter les maladies.

Conclusion

Dans cette étude, nous avons analysé les microorganismes dans les eaux de puits construits à proximité des fosses septiques. Nous avons également analysé le microbiome intestinal de 3 cohortes ; une rurale utilisant ces eaux de puits comme eaux de boisson ; et deux urbaines utilisant les eaux potables. L'analyse est effectuée par séquençage moléculaire Shotgun ; 16S/ITS.

Pour l'analyse des eaux de puits ; les résultats ont confirmé la contamination par plusieurs virus et bactéries pathogènes (*Legionella* , *Acinetobacter*, *Pseudomonas_aeruginosa* , *Mycobacterium*, *Salmonella*, *Vibrio_rotiferianus* , *Stenotrophomonas*, *Sphingomonas* , *Flavobacterium*, *Caulobacter*, *Nitrosomonas*, *Nitrospira* , *Streptomyces* ...).

Les analyses de selles ont indiqué que les bactéries *Bifidobacterium* suivi de *Ruminococcus* étaient les plus abondants dans tous les échantillons.

L'abondance relative des top 10 des bactéries identifiées est presque la même entre les 3 cohortes comme a été démontré par le test statistique R^2 .

Les résultats des analyses concernant les champignons ont montré une abondance plus importantes dans les selles des ruraux par rapport à celles des urbains.

L'analyse statistique bêta diversité a montré que les ruraux partagent plus d'espèces de bactéries et champignons que les urbains.

Ce travail de recherche a démontré que la consommation d'eau contaminée n'a pas impacté directement la composition du microbiome intestinal. Néanmoins ; d'autres études ont rapporté que l'utilisation des eaux contaminées, constitue un danger sanitaire sur la population et peut engendrer diverses maladies d'origine hydrique (diarrhée, thyphoïde ...).

Le changement de la composition du microbiome intestinal, pourrait être influencé par plusieurs facteurs, en l'occurrence les modes de vie différents en termes de géographie, régime alimentaire et autres facteurs.

Bibliographie

- [1] World Health Organization, *Guidelines for drinking-water quality*. 2017.
- [2] G. Jean, R. Bayard, P. Naquin, J. Lacour, et R. Gourdon, « Les toilettes à litière biomâîtrisée : une solution de rechange à la réduction de la contamination microbiologique des eaux souterraines en Haïti », vol. 3, n° 1, p. 5, 2014.
- [3] WHO/UNICEF Joint Water Supply and Sanitation Monitoring Programme, World Health Organization, et UNICEF, Éd., *Meeting the MDG drinking water and sanitation target: the urban and rural challenge of the decade*. Geneva, Switzerland : New York: World Health Organization ; United Nations Children's Fund, 2006.
- [4] W. H. Organization, *Progress on Sanitation and Drinking Water : 2015 Update and MDG Assessment*. World Health Organization, 2015.
- [5] « How Much Water is There on Earth? » https://www.usgs.gov/special-topic/water-science-school/science/how-much-water-there-earth?qt-science_center_objects=0#qt-science_center_objects (consulté le déc. 17, 2020).
- [6] *Biogeochemistry*. Elsevier, 2013. doi: 10.1016/C2010-0-66291-2.
- [7] S. L. Murthy, « The Human Right(s) to Water and Sanitation: History, Meaning, and the Controversy over-Privatization », *Berkeley J. Int. Law*, vol. 31, n° 1, p. 89-149, 2013.
- [8] « factsheet35en.pdf ». Consulté le: déc. 17, 2020. [En ligne]. Disponible sur: <https://www.ohchr.org/documents/publications/factsheet35en.pdf>
- [9] M. Vega, R. Pardo, E. Barrado, et L. Debán, « Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis », *Water Res.*, vol. 32, n° 12, p. 3581-3592, déc. 1998, doi: 10.1016/S0043-1354(98)00138-9.
- [10] « Water, sanitation and hygiene (WASH) ». <https://www.unicef.org/eu/water-sanitation-and-hygiene-wash> (consulté le déc. 16, 2020).
- [11] G. Bolomey et J. Morel, « Solution d'avenir pour l'agriculture et l'environnement », p. 22.
- [12] UNDP, Éd., *Beyond scarcity: power, poverty and the global water crisis*. New York, NY: UNDP, 2006.
- [13] M. Haseena *et al.*, « Water pollution and human health. », *Environ. Risk Assess. Remediat.*, vol. 01, n° 03, 2017, doi: 10.4066/2529-8046.100020.
- [14] E. R. Alley, *Water quality control handbook*, 2nd ed. New York : Alexandria, Va: McGraw-Hill ; WEF Press, 2007.
- [15] A. Najah, A. El-Shafie, O. A. Karim, O. Jaafar, et A. H. El-Shafie, « An application of different artificial intelligences techniques for water quality prediction », *Int J Phys Sci*, p. 11.
- [16] « Cholera -- Worldwide, 1989 ». <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001634.htm> (consulté le déc. 18, 2020).
- [17] S. M. Shah, M. Yousafzai, N. B. Lakhani, R. A. Chotani, et G. Nowshad, « Prevalence and correlates of diarrhea », *Indian J. Pediatr.*, vol. 70, n° 3, p. 207-211, mars 2003, doi: 10.1007/BF02725583.

- [18] J. A. Reuter, D. Spacek, et M. P. Snyder, « High-Throughput Sequencing Technologies », *Mol. Cell*, vol. 58, n° 4, p. 586-597, mai 2015, doi: 10.1016/j.molcel.2015.05.004.
- [19] V. B. Young, « The role of the microbiome in human health and disease: an introduction for clinicians », *BMJ*, vol. 356, p. j831, mars 2017, doi: 10.1136/bmj.j831.
- [20] M. J. Blaser, « The microbiome revolution », *J. Clin. Invest.*, vol. 124, n° 10, p. 4162-4165, oct. 2014, doi: 10.1172/JCI78366.
- [21] R. E. Ley, D. A. Peterson, et J. I. Gordon, « Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine », *Cell*, vol. 124, n° 4, p. 837-848, févr. 2006, doi: 10.1016/j.cell.2006.02.017.
- [22] M. A. J. Hullar et B. C. Fu, « Diet, the Gut Microbiome, and Epigenetics », *Cancer J.*, vol. 20, n° 3, p. 170-175, juin 2014, doi: 10.1097/PPO.0000000000000053.
- [23] S. V. Lynch et O. Pedersen, « The Human Intestinal Microbiome in Health and Disease », *N. Engl. J. Med.*, vol. 375, n° 24, p. 2369-2379, déc. 2016, doi: 10.1056/NEJMra1600266.
- [24] G. A. Cresci et E. Bawden, « The Gut Microbiome: What we do and don't know », *Nutr. Clin. Pract. Off. Publ. Am. Soc. Parenter. Enter. Nutr.*, vol. 30, n° 6, p. 734-746, déc. 2015, doi: 10.1177/0884533615609899.
- [25] L. Maier *et al.*, « Extensive impact of non-antibiotic drugs on human gut bacteria », *Nature*, vol. 555, n° 7698, p. 623-628, mars 2018, doi: 10.1038/nature25979.
- [26] J. Terao et R. Mukai, « Prenylation modulates the bioavailability and bioaccumulation of dietary flavonoids », *Arch. Biochem. Biophys.*, vol. 559, p. 12-16, oct. 2014, doi: 10.1016/j.abb.2014.04.002.
- [27] H. Sugahara *et al.*, « Probiotic *Bifidobacterium longum* alters gut luminal metabolism through modification of the gut microbial community », *Sci. Rep.*, vol. 5, n° 1, p. 13548, oct. 2015, doi: 10.1038/srep13548.
- [28] A. V. Tyakht *et al.*, « Human gut microbiota community structures in urban and rural populations in Russia », *Nat. Commun.*, vol. 4, n° 1, p. 2469, déc. 2013, doi: 10.1038/ncomms3469.
- [29] E. D. Sonnenburg et J. L. Sonnenburg, « Starving our Microbial Self: The Deleterious Consequences of a Diet Deficient in Microbiota-Accessible Carbohydrates », *Cell Metab.*, vol. 20, n° 5, p. 779-786, nov. 2014, doi: 10.1016/j.cmet.2014.07.003.
- [30] H. L. Simpson et B. J. Campbell, « Review article: dietary fibre-microbiota interactions », *Aliment. Pharmacol. Ther.*, vol. 42, n° 2, p. 158-179, juill. 2015, doi: 10.1111/apt.13248.
- [31] N. M. Koropatkin, E. A. Cameron, et E. C. Martens, « How glycan metabolism shapes the human gut microbiota », *Nat. Rev. Microbiol.*, vol. 10, n° 5, p. 323-335, mai 2012, doi: 10.1038/nrmicro2746.
- [32] L. G. Albenberg et G. D. Wu, « Diet and the Intestinal Microbiome: Associations, Functions, and Implications for Health and Disease », *Gastroenterology*, vol. 146, n° 6, p. 1564-1572, mai 2014, doi: 10.1053/j.gastro.2014.01.058.
- [33] B. P. Youmans *et al.*, « Characterization of the human gut microbiome during travelers' diarrhea », *Gut Microbes*, vol. 6, n° 2, p. 110-119, mars 2015, doi: 10.1080/19490976.2015.1019693.

- [34] X. Liang et G. A. FitzGerald, « Timing the Microbes: The Circadian Rhythm of the Gut Microbiome », *J. Biol. Rhythms*, vol. 32, n° 6, p. 505-515, déc. 2017, doi: 10.1177/0748730417729066.
- [35] E. A. Mayer, K. Tillisch, et A. Gupta, « Gut/brain axis and the microbiota », *J. Clin. Invest.*, vol. 125, n° 3, p. 926-938, mars 2015, doi: 10.1172/JCI176304.
- [36] J. Lu, L. Lu, Y. Yu, J. Cluette-Brown, C. R. Martin, et E. C. Claud, « Effects of Intestinal Microbiota on Brain Development in Humanized Gnotobiotic Mice », *Sci. Rep.*, vol. 8, n° 1, p. 5443, déc. 2018, doi: 10.1038/s41598-018-23692-w.
- [37] J. A. Bravo *et al.*, « Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve », *Proc. Natl. Acad. Sci.*, vol. 108, n° 38, p. 16050-16055, sept. 2011, doi: 10.1073/pnas.1102999108.
- [38] Y. Belkaid et T. W. Hand, « Role of the Microbiota in Immunity and Inflammation », *Cell*, vol. 157, n° 1, p. 121-141, mars 2014, doi: 10.1016/j.cell.2014.03.011.
- [39] M.-N. Kolopp-Sarda, « Système immunitaire muqueux et microbiote intestinal : Histoire d'une symbiose », *Rev. Francoph. Lab.*, vol. 2016, n° 484, p. 39-47, juill. 2016, doi: 10.1016/S1773-035X(16)30222-2.
- [40] « Interactions Between Commensal Fungi and the C-Type Lectin Receptor Dectin-1 Influence Colitis | Science ». <https://science.sciencemag.org/content/336/6086/1314.abstract> (consulté le mai 05, 2021).
- [41] MetaHIT Consortium *et al.*, « A human gut microbial gene catalogue established by metagenomic sequencing », *Nature*, vol. 464, n° 7285, p. 59-65, mars 2010, doi: 10.1038/nature08821.
- [42] M. D. Gershon, « The enteric nervous system: a second brain », *Hosp. Pract. 1995*, vol. 34, n° 7, p. 31-32, 35-38, 41-42 passim, juill. 1999, doi: 10.3810/hp.1999.07.153.
- [43] G. Ranucci, V. Buccigrossi, M. B. de Freitas, A. Guarino, et A. Giannattasio, « Early-Life Intestine Microbiota and Lung Health in Children », *J. Immunol. Res.*, vol. 2017, 2017, doi: 10.1155/2017/8450496.
- [44] J. H. Skalski *et al.*, « Expansion of commensal fungus *Wallemia mellicola* in the gastrointestinal mycobiota enhances the severity of allergic airway disease in mice », *PLOS Pathog.*, vol. 14, n° 9, p. e1007260, sept. 2018, doi: 10.1371/journal.ppat.1007260.
- [45] J. Tang *et al.*, « Fate Mapping of Sca1+ Cardiac Progenitor Cells in the Adult Mouse Heart », *Circulation*, vol. 138, n° 25, p. 2967-2969, déc. 2018, doi: 10.1161/CIRCULATIONAHA.118.036210.
- [46] E. Y. Hsiao *et al.*, « Microbiota Modulate Behavioral and Physiological Abnormalities Associated with Neurodevelopmental Disorders », *Cell*, vol. 155, n° 7, p. 1451-1463, déc. 2013, doi: 10.1016/j.cell.2013.11.024.
- [47] T. R. Sampson *et al.*, « Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease », *Cell*, vol. 167, n° 6, p. 1469-1480.e12, déc. 2016, doi: 10.1016/j.cell.2016.11.018.
- [48] V. C. Harris, B. W. Haak, M. Boele van Hensbroek, et W. J. Wiersinga, « The Intestinal Microbiome in Infectious Diseases: The Clinical Relevance of a Rapidly Emerging Field », *Open Forum Infect. Dis.*, vol. 4, n° 3, p. ofx144, juill. 2017, doi: 10.1093/ofid/ofx144.

- [49] A. J. Bäumler et V. Sperandio, « Interactions between the microbiota and pathogenic bacteria in the gut », *Nature*, vol. 535, n° 7610, p. 85-93, juill. 2016, doi: 10.1038/nature18849.
- [50] M. E. Hibbing, C. Fuqua, M. R. Parsek, et S. B. Peterson, « Bacterial competition: surviving and thriving in the microbial jungle », *Nat. Rev. Microbiol.*, vol. 8, n° 1, p. 15-25, janv. 2010, doi: 10.1038/nrmicro2259.
- [51] J. L. Round et S. K. Mazmanian, « The gut microbiota shapes intestinal immune responses during health and disease », *Nat. Rev. Immunol.*, vol. 9, n° 5, p. 313-323, mai 2009, doi: 10.1038/nri2515.
- [52] M. J. Ostaff, E. F. Stange, et J. Wehkamp, « Antimicrobial peptides and gut microbiota in homeostasis and pathology », *EMBO Mol. Med.*, vol. 5, n° 10, p. 1465-1483, oct. 2013, doi: 10.1002/emmm.201201773.
- [53] L. K. Chung et M. Raffatellu, « G.I. pros: Antimicrobial defense in the gastrointestinal tract », *Semin. Cell Dev. Biol.*, vol. 88, p. 129-137, avr. 2019, doi: 10.1016/j.semcdb.2018.02.001.
- [54] M. Kasubuchi, S. Hasegawa, T. Hiramatsu, A. Ichimura, et I. Kimura, « Dietary Gut Microbial Metabolites, Short-chain Fatty Acids, and Host Metabolic Regulation », *Nutrients*, vol. 7, n° 4, p. 2839-2849, avr. 2015, doi: 10.3390/nu7042839.
- [55] N. Kamada, G. Y. Chen, N. Inohara, et G. Núñez, « Control of pathogens and pathobionts by the gut microbiota », *Nat. Immunol.*, vol. 14, n° 7, p. 685-690, juill. 2013, doi: 10.1038/ni.2608.
- [56] J. A. Sorg et A. L. Sonenshein, « Bile Salts and Glycine as Cogerminants for *Clostridium difficile* Spores », *J. Bacteriol.*, vol. 190, n° 7, p. 2505-2512, avr. 2008, doi: 10.1128/JB.01765-07.
- [57] C. Lupp *et al.*, « Host-Mediated Inflammation Disrupts the Intestinal Microbiota and Promotes the Overgrowth of Enterobacteriaceae », *Cell Host Microbe*, vol. 2, n° 2, p. 119-129, août 2007, doi: 10.1016/j.chom.2007.06.010.
- [58] G. Gorkiewicz *et al.*, « Alterations in the Colonic Microbiota in Response to Osmotic Diarrhea », *PLoS ONE*, vol. 8, n° 2, p. e55817, févr. 2013, doi: 10.1371/journal.pone.0055817.
- [59] L. A. David *et al.*, « Gut Microbial Succession Follows Acute Secretory Diarrhea in Humans », *mBio*, vol. 6, n° 3, p. e00381-15, mai 2015, doi: 10.1128/mBio.00381-15.
- [60] M. C. Mattioli, A. J. Pickering, R. J. Gilsdorf, J. Davis, et A. B. Boehm, « Hands and Water as Vectors of Diarrheal Pathogens in Bagamoyo, Tanzania », *Environ. Sci. Technol.*, vol. 47, n° 1, p. 355-363, janv. 2013, doi: 10.1021/es303878d.
- [61] F. Sanger, S. Nicklen, et A. R. Coulson, « DNA sequencing with chain-terminating inhibitors », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 74, n° 12, p. 5463-5467, déc. 1977, doi: 10.1073/pnas.74.12.5463.
- [62] « Sequencing technologies — the next generation | Nature Reviews Genetics ». <https://www.nature.com/articles/nrg2626> (consulté le juin 03, 2021).
- [63] D. A. Wheeler *et al.*, « The complete genome of an individual by massively parallel DNA sequencing », *Nature*, vol. 452, n° 7189, p. 872-876, avr. 2008, doi: 10.1038/nature06884.
- [64] « Accurate whole human genome sequencing using reversible terminator chemistry | Nature ». <https://www.nature.com/articles/nature07517> (consulté le juin 03, 2021).

- [65] K. J. McKernan *et al.*, « Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding », *Genome Res.*, vol. 19, n° 9, p. 1527-1541, sept. 2009, doi: 10.1101/gr.091868.109.
- [66] J. Shendure et H. Ji, « Next-generation DNA sequencing », *Nat. Biotechnol.*, vol. 26, n° 10, p. 1135-1145, oct. 2008, doi: 10.1038/nbt1486.
- [67] G. E. Fox, L. J. Magrum, W. E. Balch, R. S. Wolfe, et C. R. Woese, « Classification of methanogenic bacteria by 16S ribosomal RNA characterization », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 74, n° 10, p. 4537-4541, oct. 1977.
- [68] R. P. Dickson *et al.*, « Analysis of Culture-Dependent versus Culture-Independent Techniques for Identification of Bacteria in Clinically Obtained Bronchoalveolar Lavage Fluid », *J. Clin. Microbiol.*, vol. 52, n° 10, p. 3605-3613, oct. 2014, doi: 10.1128/JCM.01028-14.
- [69] R. Ranjan, A. Rani, A. Metwally, H. S. McGee, et D. L. Perkins, « Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing », *Biochem. Biophys. Res. Commun.*, vol. 469, n° 4, p. 967-977, janv. 2016, doi: 10.1016/j.bbrc.2015.12.083.
- [70] « Site de la Direction Régionale de Rabat-Salé-Kénitra », *Site de la Direction Régionale de Rabat-Salé-Kénitra*. <https://www.hcp.ma/region-rabat> (consulté le nov. 20, 2020).
- [71] « Monographie de la région Rabat-Salé-Kénitra | Centre d'Echange d'Information sur la Biodiversité de la Région Rabat-Salé-Kénitra ». <http://ma.chm-cbd.net/rabat-sale-kenitra/region/monographie-de-la-region-rabat-sale-kenitra> (consulté le nov. 20, 2020).
- [72] « Eaux souterraines », *ABHS*, avr. 19, 2016. <http://www.abhsebou.ma/presentation-du-bassin/eaux-souterraines/> (consulté le nov. 20, 2020).
- [73] « analyse-de-l_eau-Rodier.pdf ».
- [74] Z. R. Corporation, « Quick-DNA Miniprep Plus », p. 19.
- [75] « Quant-iT™ PicoGreen ® dsDNA Reagent and Kits », p. 7.
- [76] « BioTek: Gen5 Software Features for Imaging & Microscopy - Overview ». <https://www.biotek.com/products/software-robotics-software/gen5-software-features-for-imaging-microscopy/> (consulté le déc. 24, 2020).
- [77] « 5989-7725FR.pdf ». Consulté le: déc. 24, 2020. [En ligne]. Disponible sur: <https://www.agilent.com/cs/library/brochures/5989-7725FR.pdf>
- [78] « Chemical Analysis, Life Sciences, and Diagnostics | Agilent ». <https://www.agilent.com/> (consulté le déc. 24, 2020).
- [79] « US4683195A - Process for amplifying, detecting, and/or-cloning nucleic acid sequences - Google Patents ». <https://patents.google.com/patent/US4683195A/en> (consulté le déc. 25, 2020).
- [80] « Nextera XT DNA Library Prep Kit Reference Guide (15031942) », p. 30.
- [81] M. Baym, S. Kryazhimskiy, T. D. Lieberman, H. Chung, M. M. Desai, et R. Kishony, « Inexpensive Multiplexed Library Preparation for Megabase-Sized Genomes », *PLOS ONE*, vol. 10, n° 5, p. e0128036, mai 2015, doi: 10.1371/journal.pone.0128036.
- [82] « Integrated DNA Technologies | IDT », *Integrated DNA Technologies*. <https://eu.idtdna.com/pages> (consulté le janv. 04, 2021).

- [83] « Agencourt AMPure XP.pdf ». Consulté le: déc. 28, 2020. [En ligne]. Disponible sur:
https://research.fredhutch.org/content/dam/stripe/hahn/methods/mol_biol/Agencourt%20AMPure%20XP.pdf
- [84] « 16s-metagenomic-library-prep-guide-15044223-b.pdf ». Consulté le: déc. 28, 2020. [En ligne]. Disponible sur:
https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf
- [85] J. J. Kozich, S. L. Westcott, N. T. Baxter, S. K. Highlander, et P. D. Schloss, « Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform », *Appl. Environ. Microbiol.*, vol. 79, n° 17, p. 5112-5120, sept. 2013, doi: 10.1128/AEM.01043-13.
- [86] A. M. Bolger, M. Lohse, et B. Usadel, « Trimmomatic: a flexible trimmer for Illumina sequence data », *Bioinformatics*, vol. 30, n° 15, p. 2114-2120, août 2014, doi: 10.1093/bioinformatics/btu170.
- [87] A. P. Masella, A. K. Bartram, J. M. Truszkowski, D. G. Brown, et J. D. Neufeld, « PANDAsseq: paired-end assembler for illumina sequences », *BMC Bioinformatics*, vol. 13, n° 1, p. 31, févr. 2012, doi: 10.1186/1471-2105-13-31.
- [88] P. D. Schloss *et al.*, « Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities », *Appl. Environ. Microbiol.*, vol. 75, n° 23, p. 7537-7541, déc. 2009, doi: 10.1128/AEM.01541-09.
- [89] C. Quast *et al.*, « The SILVA ribosomal RNA gene database project: improved data processing and web-based tools », *Nucleic Acids Res.*, vol. 41, n° Database issue, p. D590-D596, janv. 2013, doi: 10.1093/nar/gks1219.
- [90] « Why do I have such a large distance matrix », <https://mothur.org>.
<https://mothur.org> (consulté le déc. 28, 2020).
- [91] R. C. Edgar, « Search and clustering orders of magnitude faster than BLAST », *Bioinformatics*, vol. 26, n° 19, p. 2460-2461, oct. 2010, doi: 10.1093/bioinformatics/btq461.
- [92] « R: a language and environment for statistical computing ». <https://www.gbif.org/fr/tool/81287/r-a-language-and-environment-for-statistical-computing> (consulté le déc. 28, 2020).
- [93] P. J. McMurdie et S. Holmes, « phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data », *PLoS ONE*, vol. 8, n° 4, p. e61217, avr. 2013, doi: 10.1371/journal.pone.0061217.
- [94] P. R. Chowdhury, Y. Boucher, K. A. Hassan, I. T. Paulsen, H. W. Stokes, et M. Labbate, « Genome Sequence of *Vibrio rotiferianus* Strain DAT722[▽] », *J. Bacteriol.*, vol. 193, n° 13, p. 3381-3382, juill. 2011, doi: 10.1128/JB.05089-11.
- [95] « Salmonella in surface and drinking water: Occurrence and water-mediated transmission | Elsevier Enhanced Reader ». <https://reader.elsevier.com/reader/sd/pii/S0963996911004169?token=D6AED096E1B5A14B73CFB91A526F7FE6EC43AE9540FDA9FE8B88E9FEF6B8571E1DEE498A5B0A541B063FDFC4CC542A8B> (consulté le févr. 11, 2021).
- [96] S. Allegra, F. Berger, P. Berthelot, F. Grattard, B. Pozzetto, et S. Riffard, « Use of Flow Cytometry To Monitor *Legionella* Viability », *Appl. Environ. Microbiol.*, vol. 74, n° 24, p. 7813-7816, déc. 2008, doi: 10.1128/AEM.01364-08.

- [97] S. T. Tokajian, F. A. Hashwa, I. C. Hancock, et P. A. Zalloua, « Phylogenetic assessment of heterotrophic bacteria from a water distribution system using 16S rDNA sequencing », *Can. J. Microbiol.*, févr. 2011, doi: 10.1139/w05-007.
- [98] A. Carvalheira, J. Silva, et P. Teixeira, « Acinetobacter spp. in food and drinking water – A review », *Food Microbiol.*, vol. 95, p. 103675, mai 2021, doi: 10.1016/j.fm.2020.103675.
- [99] M. Gajdács et E. Urbán, « Prevalence and Antibiotic Resistance of *Stenotrophomonas maltophilia* in Respiratory Tract Samples: A 10-Year Epidemiological Snapshot », *Health Serv. Res. Manag. Epidemiol.*, vol. 6, août 2019, doi: 10.1177/2333392819870774.
- [100] « Flavobacterium - an overview | ScienceDirect Topics ». <https://www.sciencedirect.com/topics/immunology-and-microbiology/flavobacterium> (consulté le févr. 11, 2021).
- [101] P. W. J. J. van der Wielen, S. Voost, et D. van der Kooij, « Ammonia-Oxidizing Bacteria and Archaea in Groundwater Treatment and Drinking Water Distribution Systems », *Appl. Environ. Microbiol.*, vol. 75, n° 14, p. 4687-4695, juill. 2009, doi: 10.1128/AEM.00387-09.
- [102] J. M. Regan, G. W. Harrington, H. Baribeau, R. D. Leon, et D. R. Noguera, « Diversity of nitrifying bacteria in full-scale chloraminated distribution systems », *Water Res.*, vol. 37, n° 1, p. 197-205, janv. 2003, doi: 10.1016/S0043-1354(02)00237-3.
- [103] B. Zaitlin et S. B. Watson, « Actinomycetes in relation to taste and odour in drinking water: Myths, tenets and truths », *Water Res.*, vol. 40, n° 9, p. 1741-1753, mai 2006, doi: 10.1016/j.watres.2006.02.024.
- [104] J.-H. Shin, M. Sim, J.-Y. Lee, et D.-M. Shin, « Lifestyle and geographic insights into the distinct gut microbiota in elderly women from two different geographic locations », *J. Physiol. Anthropol.*, vol. 35, déc. 2016, doi: 10.1186/s40101-016-0121-7.
- [105] H. Safouane et K. Yassine, « Hydrogéologie et Qualité des eaux de la nappe du Gharb », p. 54.
- [106] S. Cairncross et R. Feachem, *Environmental Health Engineering in the Tropics: Water, Sanitation and Disease Control*. Routledge, 2018.
- [107] P. R. Hunter, J. M. Colford, M. W. LeChevallier, S. Binder, et P. S. Berger, « Waterborne diseases. », *Emerg. Infect. Dis.*, vol. 7, n° 3 Suppl, p. 544, 2001.
- [108] « The Human Intestinal Microbiome in Health and Disease | NEJM ». <https://www.nejm.org/doi/full/10.1056/NEJMra1600266> (consulté le déc. 18, 2020).
- [109] I. Adlerberth *et al.*, « Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? », *Pediatr. Res.*, vol. 59, n° 1, p. 96-101, janv. 2006, doi: 10.1203/01.pdr.0000191137.12774.b2.
- [110] M. G. Dominguez-Bello *et al.*, « Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, n° 26, p. 11971-11975, juin 2010, doi: 10.1073/pnas.1002601107.
- [111] « Development of gut microbiota in infants not exposed to medical interventions - PubMed ». <https://pubmed.ncbi.nlm.nih.gov/21143523/> (consulté le juill. 10, 2021).

- [112] « Human gut microbiome viewed across age and geography | Nature ». <https://www.nature.com/articles/nature11053> (consulté le juill. 10, 2021).
- [113] G. Zeller *et al.*, « Potential of fecal microbiota for early-stage detection of colorectal cancer », *Mol. Syst. Biol.*, vol. 10, n° 11, nov. 2014, doi: 10.15252/msb.20145645.
- [114] « Feng Y, et al. (2015) | SGD ». <https://www.yeastgenome.org/reference/S000179552> (consulté le juill. 10, 2021).

Annexes

Annexe 1: Protocole du PicoGreen

- "Stock A" : Préparer 1 mL du stock de 2 ug/ml :
20 µL du stock d'ADN Lambda + 0,98 mL de 1x TE⁴
- "Stock B" : Préparer 1 mL du stock de 0,2 ug/ml :
100 µL du stock de 2 µg/ml + 900 µL de TE
- Préparer 1 mL du stock de 50 ng/ml :
25 µL du stock de 2 µg/ml + 975 µL de 1x TE
- Préparer 1 mL du stock de 5 ng/ml :
100 µL du stock de 50 ng/ml + 900 µL de TE
- Calculer la quantité de TE et de Pico-GREEN, TE nécessaire pour tous les échantillons (en remplissant les cases grises sous le nom #Samples)
Vol (mL) nécessaire du Pico-Green dilué = $2 + 0,125 * (\#échantillons + 1)$
- **Vol (mL) nécessaire 1x TE** = $1 + 0,125 * (\#échantillons + 1)$ + **Vol (mL) nécessaire du Pico-Green dilué**
- Préparer 1x TE et le stock du Pico-Green dilué :
Vol (mL) de 20 x stock TE = **Vol (mL) nécessaire 1x TE** / 20
- **Volume (mL) d'EAU UPW**
Vol (mL) nécessaire 1x TE - **Vol (mL) de 20x stock TE**
- **Vol (µL) du stock Pico-Green 200x stock TAMPON** = **Vol (mL) nécessaire Pico-Green dilué** / 200 * 1000
- **Volume (mL) de TE** = **Vol (mL) nécessaire du Pico-Green dilué** - (**Vol (µL) de Pico-Green 200x stock** / 1000)

⁴ TE tampon (ou tampon Tris EDTA) est l'un des tampons largement utilisés dans la préparation et le stockage des protocoles d'ADN et d'ARN. Le tampon TE est également connu sous le nom de tampon T 10 E 1 en raison de leur rapport dans la concentration de travail, qui est respectivement de 10 mM et 1 mM. Comme son nom l'indique, le tampon TE est composé de deux réactifs, c'est-à-dire Tris et EDTA.

Annexe 2: Protocole de Beckman

Sélection de la taille: 600ul (pcr + son amplifié) divisé en 12 puits (50ul chacun)

1. Plusieurs heures à l'avance, sortez les billes magnétiques Ampure XP du réfrigérateur afin qu'elles soient à température ambiante. Vortexer soigneusement les billes avant utilisation. Prenez 780ul (65ul dans chaque puits)
2. Diluer de l'éthanol frais à 80%. (4800: 3840ul 100% d'éthanol + 960ul d'eau)
3. Ajouter 30 ul de billes magnétiques à chaque échantillon. Bien mélanger en pipetant de haut en bas. Incuber 8 minutes
4. Placez la plaque sur le support magnétique. Attends 5 minutes
5. En gardant la plaque sur le support magnétique, aspirez et jetez le liquide de chaque puits en laissant les billes magnétiques
6. En gardant la plaque sur le support magnétique, ajoutez 200 ul d'éthanol à 80% dans chaque puits. Incuber pendant 30 secondes et aspirer. Ne dérangez pas les billes magnétiques
7. Répétez l'étape 4.
8. Retirer soigneusement toute trace d'éthanol des puits et laisser sécher pendant 5 minutes (toujours sur support magnétique)
9. Retirer la plaque PCR du support magnétique et suspendre les billes magnétiques dans 20 ul d'eau. Mélanger en pipetant de haut en bas
10. Placer la plaque sur un support magnétique pendant 2 minutes
11. Transférer de l'eau (qui contient maintenant de l'ADN) sur une nouvelle plaque PCR (nous avons utilisé un eppendorf). Ces bibliothèques d'ADN sont prêtes à être séquencées

Annexe 3: Metadata

#NAME	Sex	Age	Water	Cohort	Family
SE1	M	56	Well	Rural	Morocco
SE2	M	3	Well	Rural	Morocco
SE3	F	26	Well	Rural	Morocco
SE4	F	50	Well	Rural	Morocco
SE5	F	35	Well	Rural	Morocco
SE6	F	40	Well	Rural	Morocco
SE7	F	68	Well	Rural	Morocco
SE8	F	59	Well	Rural	Morocco
SE9	M	41	Well	Rural	Morocco
SE10	M	14	Well	Rural	Morocco
SE11	F	3	Well	Rural	Morocco
SE12	M	16	Well	Rural	Morocco
SE13	F	8	Well	Rural	Morocco
SE14	M	58	Well	Rural	Morocco
SE15	M	46	Well	Rural	Morocco
SE16	M	2	Well	Rural	Morocco
SE17	F	6	Well	Rural	Morocco
SE18	F	60	Well	Rural	Morocco
SE19	F	31	Well	Rural	Morocco
SE20	M	34	Well	Rural	Morocco
SE21	M	63	Well	Rural	Morocco
SE22	M	7	Well	Rural	Morocco
SE23	F	34	Well	Rural	Morocco
SE24	F	30	Well	Rural	Morocco
SE25	M	14	Well	Rural	Morocco
SE26	F	8	Well	Rural	Morocco
SE27	F	56	Well	Rural	Morocco
SE28	F	26	Tap	Urban	Morocco
SE29	M	23	Tap	Urban	Morocco
SE30	F	25	Tap	Urban	Morocco
SE31	F	48	Tap	Urban	Morocco
SE32	M	13	Tap	Urban	Morocco
SE33	M	26	Tap	Urban	Morocco
SE34	F	14	Tap	Urban	Morocco
SE35	M	58	Tap	Urban	Morocco
SE36	F	49	Tap	Urban	Morocco
SE37	M	56	Well	Rural	Morocco
SE38	M	56	Well	Rural	Morocco
SE39	M	37	Tap	Urban	Morocco
SE40	M	54	Tap	Urban	Morocco

SE41	M	41	Tap	Urban	Morocco
SE42	F	36	Tap	Urban	Morocco
SE43	M	35	Tap	Urban	Morocco
SE44	M	30	Tap	Urban	Morocco
SE45	F	19	Tap	Migrant	Guinea
SE46	F	19	Tap	Migrant	Zombia
SE47	F	19	Tap	Migrant	Zombia
SE48	F	22	Tap	Migrant	Guinea
SE49	F	22	Tap	Migrant	Guinea
SE50	F	22	Tap	Migrant	Guinea
SE51	F	19	Tap	Migrant	Guinea
SE52	F	20	Tap	Migrant	Guinea
SE53	F	22	Tap	Migrant	Guinea
SE54	M	24	Tap	Migrant	Guinea
SE55	F	21	Tap	Migrant	Guinea
SE56	M	20	Tap	Migrant	Guinea
SE57	F	19	Tap	Migrant	Zombia
SE58	F	19	Tap	Migrant	Zombia
SE59	M	56	Well	Rural	Morocco
SE60	F	22	Tap	Migrant	Guinea
SE61	F	26	Tap	Migrant	Guinea
SE62	F	25	Tap	Migrant	Guinea
SE63	M	25	Tap	Migrant	Guinea
SE64	F	22	Tap	Migrant	Guinea
SE65	F	21	Tap	Migrant	Sudan
SE66	M	25	Tap	Migrant	Sudan
SE67	F	20	Tap	Migrant	Guinea
SE68	F	21	Tap	Migrant	Sudan
SE69	F	24	Tap	Migrant	SouthSudan
SE70	F	19	Tap	Migrant	Liberia
SE71	F	21	Tap	Migrant	Liberia
SE72	F	19	Tap	Migrant	Nigeria
SE73	F	17	Tap	Migrant	Nigeria
SE74	M	20	Tap	Migrant	Nigeria
SE75	F	18	Tap	Migrant	Nigeria
SE76	M	21	Tap	Migrant	SouthSudan
SE77	F	20	Tap	Migrant	SouthSudan
SE78	M	18	Tap	Migrant	SouthSudan
SE79	M	20	Tap	Migrant	Gombia
SE80	M	23	Tap	Migrant	Tanzania

Annexe 4: Barcodes 16S

SampleID	Plate	Well	Golay Barcode
1	1	A1	AGCCTTCGTCGC
2	1	A2	TCCATACCGGAA
3	1	A3	AGCCCTGCTACA
4	1	A4	CCTAACGGTCCA
5	1	A5	CGCGCCTTAAAC
6	1	A6	TATGGTACCCAG
7	1	A7	TACAATATCTGT
8	1	A8	AATTTAGGTAGG
9	1	A9	GACTCAACCAGT
10	1	A10	GCCTCTACGTCG
11	1	A11	ACTACTGAGGAT
12	1	A12	AATTCACCTCCT
13	1	B1	CGTATAAATGCG
14	1	B2	ATGCTGCAACAC
15	1	B3	ACTCGCTCGCTG
16	1	B4	TTCCTTAGTAGT
17	1	B5	CGTCCGTATGAA
18	1	B6	ACGTGAGGAACG
19	1	B7	GGTTGCCCTGTA
20	1	B8	CATATAGCCCGA
21	1	B9	GCCTATGAGATC
22	1	B10	CAAGTGAAGGGA
23	1	B11	CACGTTTATTCC
24	1	B12	TAATCGGTGCCA
25	1	C1	TGACTAATGGCC
26	1	C2	CGGGACACCCGA
27	1	C3	CTGTCTATACTA
28	1	C4	TATGCCAGAGAT
29	1	C5	CGTTTGGAAATGA
30	1	C6	AAGAACTCATGA
31	1	C7	TGATATCGTCTT
32	1	C8	CGGTGACCTACT
33	1	C9	AATGCGCGTATA
34	1	C10	CTTGATTCTTGA
35	1	C11	GAAATCTTGAAG
36	1	C12	GAGATACAGTTC
37	1	D1	GTGGAGTCTCAT
38	1	D2	ACCTTACACCTT
39	1	D3	TAATCTCGCCGG

40	1	D4	ATCTAGTGGCAA
41	1	D5	ACGCTTAACGAC
42	1	D6	TACGGATTATGG
43	1	D7	ATACATGCAAGA
44	1	D8	CTTAGTGCAGAA
45	1	D9	AATCTTGCGCCG
46	1	D10	AGGATCAGGGAA
47	1	D11	AATAACTAGGGT
48	1	D12	TATTGCAGCAGC
49	1	E1	TGATGTGCTAAG
50	1	E2	GTAGTAGACCAT
51	1	E3	AGTAAAGATCGT
52	1	E4	CTCGCCCTCGCC
53	1	E5	TCTCTTTCGACA
54	1	E6	ACATACTGAGCA
55	1	E7	GTTGATACGATG
56	1	E8	GTCAACGCTGTC
57	1	E9	TGAGACCCTACA
58	1	E10	ACTTGGTGTAAG
59	1	E11	ATTACGTATCAT
60	1	E12	CACGCAGTCTAC
61	1	F1	TGTGCACGCCAT
62	1	F2	CCGGACAAGAAG
63	1	F3	TTGCTGGACGCT
64	1	F4	TACTAACGCGGT
65	1	F5	GCGATCACACCT
66	1	F6	CAAACGCACTAA
67	1	F7	GAAGAGGGTTGA
68	1	F8	TGAGTGGTCTGT
69	1	F9	TTACACAAAGGC
70	1	F10	ACGACGCATTTG
71	1	F11	TATCCAAGCGCA
72	1	F12	AGAGCCAAGAGC
73	1	G1	GGTGAGCAAGCA
74	1	G2	TAAATATACCCT
75	1	G3	TTGCGGACCCTA
76	1	G4	GTCGTCCAAATG
77	1	G5	TGCACAGTCGCT
78	1	G6	TTACTGTGGCCG
79	1	G7	GGTTCATGAACA
80	1	G8	TAACAATAATTC

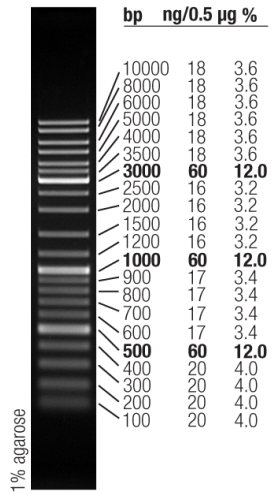
Annexe 5: Barcodes ITS

SampleID	Plate	Well	Barcode
1	A	A1	TTACCGACGAGT
2	A	A2	ATTGGACACGCT
3	A	A3	TCGCATGGATAC
4	A	A4	AGCGAACCTGTT
5	A	A5	GCGATATATCGC
6	A	A6	GTCAGCCGTAA
7	A	A7	TCCAGATAGCGT
8	A	A8	AAGGCGCTCCTT
9	A	A9	GCTCACAATGTG
10	A	A10	TTGACGACATCG
11	A	A11	CTTAGAACGTGC
12	A	A12	CGGTTACATAG
13	A	B1	CGATAGGCCTTA
14	A	B2	TGCGCTGAATGT
15	A	B3	TGGAGTAGGTGG
16	A	B4	TAGACACCGTGT
17	A	B5	TCAGCTGACTAG
18	A	B6	TAAGTCGGCCTA
19	A	B7	GCTCCTTAGAAG
20	A	B8	ATGGCCTGACTA
21	A	B9	TTGCAAGTACCG
22	A	B10	CCTAGTAAGCTG
23	A	B11	CTAGGATCACTG
24	A	B12	TATGAACGTCCG
25	A	C1	CTTGTGCGACAA
26	A	C2	AGCAGAACATCT
27	A	C3	ACGTGCCTTAGA
28	A	C4	TGAACTAGCGTC
29	A	C5	TATTCAGCGGAC
30	A	C6	TAATCGGTGCCA
31	A	C7	GCGTCCATGAAT
32	A	C8	CGTAAGATGCCT
33	A	C9	CTGTTACAGCGA
34	A	C10	ACGATCATCTGG
35	A	C11	GTAACGGCTCTA
36	A	C12	CCATGCTTAGAG
37	A	D1	GTACGCACAGTT
38	A	D2	TTAGAGCCATGC
39	A	D3	ATAAGGTCGCCT

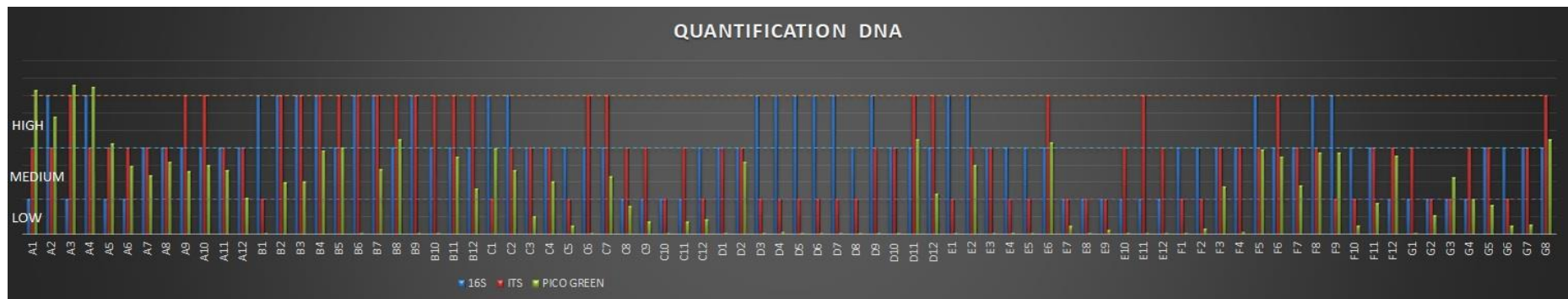
40	A	D4	AGTGGCACTATC
41	A	D5	CCAGAAGTG TTC
42	A	D6	CTACTAGCGGTA
43	A	D7	TAGCGTTCAGA
44	A	D8	GTGAGTCATA CC
45	A	D9	TGGTCCTACAAG
46	A	D10	TACGCGTACAGT
47	A	D11	GAGCCATCTGTA
48	A	D12	CGTCCGTATGAA
49	A	E1	GATACGTT CGCA
50	A	E2	CAGCTGGTTCAA
51	A	E3	TTAAGCGCCTGA
52	A	E4	CCTGCGAAGTAT
53	A	E5	TTGTAGCCGACA
54	A	E6	TCTGTAGAGCCA
55	A	E7	GTTCTCTTCTCG
56	A	E8	CTCTGAGGTAAC
57	A	E9	CAGGATTCGTAC
58	A	E10	TCACTGCTAGGA
59	A	E11	ACATGTCACGTG
60	A	E12	ATTCTGCCGAAG
61	A	F1	GGACTTCCAGCT
62	A	F2	TGCATACACTGG
63	A	F3	ACGCAATGTCTG
64	A	F4	GCTCGAAGATTC
65	A	F5	AGACGTTGCTAC
66	A	F6	TAGAGCTGCCAT
67	A	F7	TTGGCTCTATTC
68	A	F8	GACTTCATGCGA
69	A	F9	CTGCATACTGAG
70	A	F10	AATTGTGTCGGA
71	A	F11	AGTATTCGCGCA
72	A	F12	TTCGCAGATACG
73	A	G1	GTGGTGGTTTCC
74	A	G2	CTCATGGTAGCA
75	A	G3	CTCACAACCGTG
76	A	G4	GCGGACTATTCA
77	A	G5	ATCGCTTAAGGC
78	A	G6	TCAGGACGTATC
79	A	G7	GCATTACTGGAC
80	A	G8	GCTATGGA ACTC

Annexe 6: Résultats du Gel

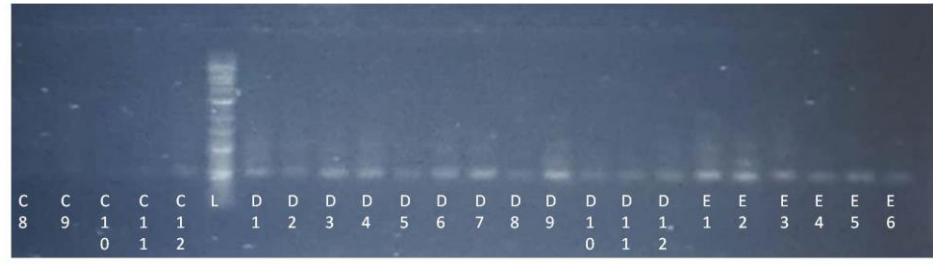
Ladder :



0,5 µg/lane, 8 cm length gel,
1X TAE, 7 V/cm, 45 min



Gel results 16S:



Line1:

Line2:

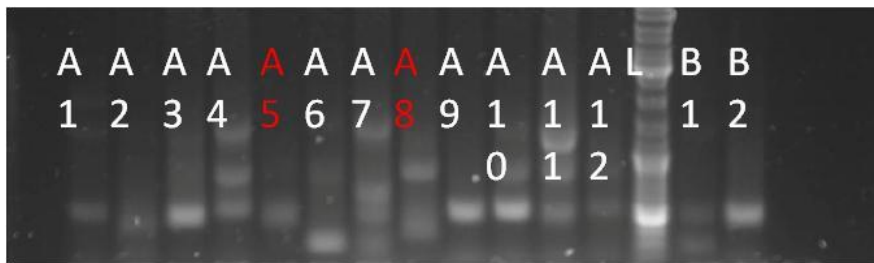


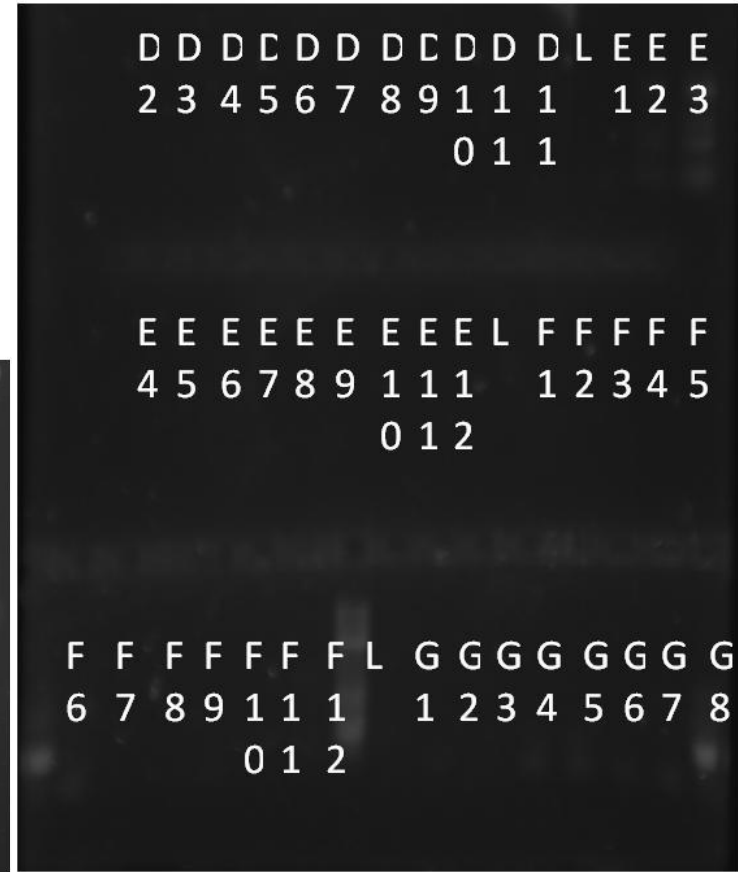
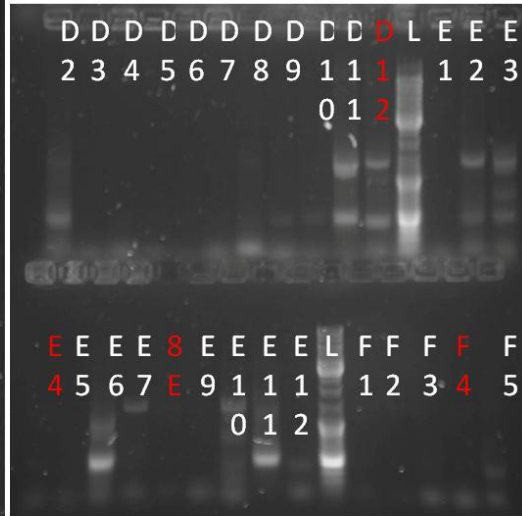
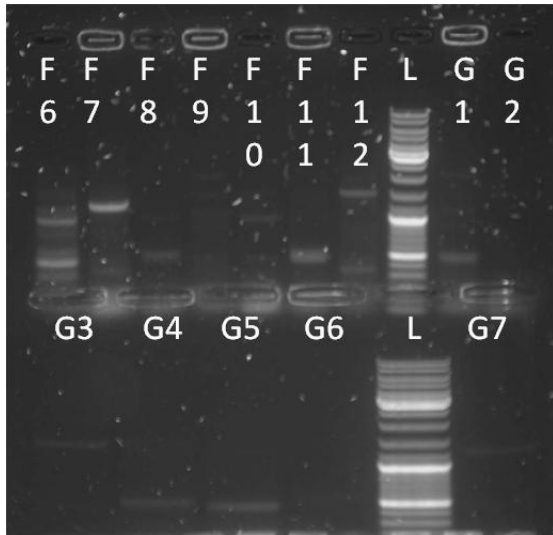
Line3:

Line4:



Gel results ITS :





Annexe 7: Résultats de Shotgun

ID	ID	Urban	Migrant
k__Archaea	1,44302	2,228	4,37286
k__Archaea p__Euryarchaeota	1,44302	2,228	4,37286
k__Archaea p__Euryarchaeota c__Methanobacteria	1,44302	2,228	4,37286
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales	1,44302	2,228	4,37286
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae	1,44302	2,228	4,37286
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae g__Methanobrevibacter	1,44302	2,13867	3,87423
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae g__Methanobrevibacter s__Methanobrevibacter_smithii	1,0109	1,8149	3,08197
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae g__Methanobrevibacter s__Methanobrevibacter_smithii t__Methanobrevibacter_smithii_unclassified	1,0109	1,8149	3,08197
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae g__Methanobrevibacter s__Methanobrevibacter_unclassified	0,43212	0,32377	0,79226
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae g__Methanosphaera	0	0,08933	0,49863
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae g__Methanosphaera s__Methanosphaera_stadtmanae	0	0,08933	0,49863
k__Archaea p__Euryarchaeota c__Methanobacteria o__Methanobacteriales f__Methanobacteriaceae g__Methanosphaera s__Methanosphaera_stadtmanae t__GCF_000012545	0	0,08933	0,49863

k__Bacteria	98,55698	97,74594	95,62714
k__Bacteria p__Actinobacteria	40,7099	34,90746	50,72333
k__Bacteria p__Actinobacteria c__Actinobacteria	40,7099	34,90746	50,72333
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales	0,13703	0,11865	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Actinomycetaceae	0,13703	0	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Actinomycetaceae g__Actinomyces	0,13703	0	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Actinomycetaceae g__Actinomyces s__Actinomyces_turicensis	0,13703	0	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Actinomycetaceae g__Actinomyces s__Actinomyces_turicensis t__GCF_000296505	0,13703	0	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Micrococcaceae	0	0,11865	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Micrococcaceae g__Rothia	0	0,11865	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Micrococcaceae g__Rothia s__Rothia_mu cilaginosa	0	0,11865	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Actinomycetales f__Micrococcaceae g__Rothia s__Rothia_mu cilaginosa t__Rothia_mucilaginosa_unclassified	0	0,11865	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales	34,02599	30,44752	44,75794
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae	34,02599	30,44752	44,75794
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium	34,02599	30,44752	44,75794
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__ _Bifidobacterium_adolescentis	17,97544	18,20063	10,46911
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__ _Bifidobacterium_adolescentis t__Bifidobacterium_adolescentis_unclassified	17,97544	18,20063	10,46911

k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_angulatum	5,80642	0,03243	0,94276
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_angulatum t__GCF_000156635	5,80642	0,03243	0,94276
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_bifidum	1,23175	0,07707	0,55038
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_bifidum t__Bifidobacterium_bifidum_unclassified	1,23175	0,07707	0,55038
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_breve	0	0	0,12389
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_breve t__Bifidobacterium_breve_unclassified	0	0	0,12389
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_catenulatum	3,30449	1,26917	4,20913
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_catenulatum t__GCF_000173455	3,30449	1,26917	4,20913
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_dentium	0	0,44292	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_dentium t__Bifidobacterium_dentium_unclassified	0	0,44292	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_longum	5,12624	8,85892	23,78054
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_longum t__Bifidobacterium_longum_unclassified	5,12624	8,85892	23,78054

k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_pseudocatenulatum	0,58164	1,56639	4,68212
k__Bacteria p__Actinobacteria c__Actinobacteria o__Bifidobacteriales f__Bifidobacteriaceae g__Bifidobacterium s__Bifidobacterium_pseudocatenulatum t__GCF_000173435	0,58164	1,56639	4,68212
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales	6,54688	4,3413	5,96539
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae	6,54688	4,3413	5,96539
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Adlercreutzia	0	0,02225	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Adlercreutzia s__Adlercreutzia_equolifaciens	0	0,02225	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Adlercreutzia s__Adlercreutzia_equolifaciens t__GCF_000478885	0	0,02225	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Collinsella	6,01193	4,2536	5,96539
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Collinsella s__Collinsella_aerofaciens	6,01193	4,2536	5,96539
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Collinsella s__Collinsella_aerofaciens t__GCF_000169035	6,01193	4,2536	5,96539
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Eggerthella	0,30824	0,06544	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Eggerthella s__Eggerthella_unclassified	0,30824	0,06544	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Olsenella	0,22671	0	0
k__Bacteria p__Actinobacteria c__Actinobacteria o__Coriobacteriales f__Coriobacteriaceae g__Olsenella s__Olsenella_unclassified	0,22671	0	0
k__Bacteria p__Bacteroidetes	0,60427	0,05027	0,65402

k__Bacteria p__Bacteroidetes c__Bacteroidia	0,60427	0,05027	0,65402
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales	0,60427	0,05027	0,65402
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae	0	0,01271	0,14628
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides	0	0,01271	0,14628
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_caccae	0	0	0,048
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_caccae t__Bacteroides_caccae_unclassified	0	0	0,048
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_fragilis	0	0	0,04716
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_fragilis t__Bacteroides_fragilis_unclassified	0	0	0,04716
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_pliebeius	0	0,01271	0
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_pliebeius t__GCF_000187895	0	0,01271	0
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_vulgatus	0	0	0,05112
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Bacteroidaceae g__Bacteroides s__Bacteroides_vulgatus t__Bacteroides_vulgatus_unclassified	0	0	0,05112
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Prevotellaceae	0,60427	0,03757	0,50775
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Prevotellaceae g__Prevotella	0,60427	0,03757	0,50775
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Prevotellaceae g__Prevotella s__Prevotella_copri	0,60427	0,03757	0,50775

k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Prevotellaceae g__Prevotella s__Prevotella_copri t__GCF_000157935	0,60427	0,03757	0,50775
k__Bacteria p__Firmicutes	54,59249	58,60946	42,09459
k__Bacteria p__Firmicutes c__Bacilli	5,19519	9,93823	2,52059
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales	5,19519	9,93823	2,52059
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Enterococcaceae	0,27747	0,06809	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Enterococcaceae g__Enterococcus	0,27747	0,06809	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Enterococcaceae g__Enterococcus s__Enterococcus_faecium	0,27747	0,06809	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Enterococcaceae g__Enterococcus s__Enterococcus_faecium t__Enterococcus_faecium_unclassified	0,27747	0,06809	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae	3,4408	0,59807	1,81964
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae g__Lactobacillus	3,4408	0,59807	1,81964
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae g__Lactobacillus s__Lactobacillus_gasseri	0	0,25152	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae g__Lactobacillus s__Lactobacillus_gasseri t__Lactobacillus_gasseri_unclassified	0	0,25152	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae g__Lactobacillus s__Lactobacillus_mucosae	0	0	0,09352
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae g__Lactobacillus s__Lactobacillus_mucosae t__GCF_000248095	0	0	0,09352
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae g__Lactobacillus s__Lactobacillus_ruminis	3,4408	0,34654	1,72613
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Lactobacillaceae g__Lactobacillus s__Lactobacillus_ruminis t__Lactobacillus_ruminis_unclassified	3,4408	0,34654	1,72613

k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae	1,47692	9,27207	0,70094
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus	1,47692	9,27207	0,70094
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_angularinosus	0	0,06877	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_angularinosus t__Streptococcus_angularinosus_unclassified	0	0,06877	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_mutans	0	0,09494	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_mutans t__Streptococcus_mutans_unclassified	0	0,09494	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_parasanguinis	0	2,7141	0,07465
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_parasanguinis t__Streptococcus_parasanguinis_unclassified	0	2,7141	0,07465
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_salivarius	1,47692	5,97833	0,6263
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_salivarius t__Streptococcus_salivarius_unclassified	1,47692	5,97833	0,6263
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_vestibularis	0	0,41593	0
k__Bacteria p__Firmicutes c__Bacilli o__Lactobacillales f__Streptococcaceae g__Streptococcus s__Streptococcus_vestibularis t__Streptococcus_vestibularis_unclassified	0	0,41593	0
k__Bacteria p__Firmicutes c__Clostridia	40,42319	39,82244	33,67003
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales	40,42319	39,82244	33,67003
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Clostridiaceae	0	0	0,25022

k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Clostridiaceae g__Clostridium	0	0	0,25022
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Clostridiaceae g__Clostridium s__Clostridium_leptum	0	0	0,07061
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Clostridiaceae g__Clostridium s__Clostridium_leptum t__GCF_000154345	0	0	0,07061
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Clostridiaceae g__Clostridium s__Clostridium_nexile	0	0	0,17961
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Clostridiaceae g__Clostridium s__Clostridium_nexile t__GCF_000156035	0	0	0,17961
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae	6,10073	5,44462	8,24479
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium	6,10073	5,44462	8,24479
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_eligen s	0	0,26962	0,02384
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_eligen s t__GCF_000146185	0	0,26962	0,02384
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_hallii	1,83097	1,28158	1,54484
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_hallii t__GCF_000173975	1,83097	1,28158	1,54484
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_ramul us	0	0,34192	0,01985
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_ramul us t__GCF_000469345	0	0,34192	0,01985
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_rectal e	4,26976	3,25113	6,50675

k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_rectale t__Eubacterium_rectale_unclassified	4,26976	3,25113	6,50675
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_siraeum	0	0	0,14951
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_siraeum t__Eubacterium_siraeum_unclassified	0	0	0,14951
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_sp_3_1_31	0	0,30036	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Eubacteriaceae g__Eubacterium s__Eubacterium_sp_3_1_31 t__Eubacterium_sp_3_1_31_unclassified	0	0,30036	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae	16,8201	15,27956	11,65453
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Anaerostipes	0	0,19306	0,08445
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Anaerostipes s__Anaerostipes_hadrus	0	0,19306	0,08445
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Anaerostipes s__Anaerostipes_hadrus t__GCF_000332875	0	0,19306	0,08445
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Blautia	8,37871	6,88063	5,04468
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Blautia s__Ruminococcus_gnavus	0	0,95618	0,0205
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Blautia s__Ruminococcus_gnavus t__Ruminococcus_gnavus_unclassified	0	0,95618	0,0205
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Blautia s__Ruminococcus_obeum	5,52052	1,81038	1,56392
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Blautia s__Ruminococcus_obeum t__Ruminococcus_obeum_unclassified	5,52052	1,81038	1,56392

k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Blautia s__Ruminococcus_torques	2,85819	4,11408	3,46025
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Blautia s__Ruminococcus_torques t__Ruminococcus_torques_unclassified	2,85819	4,11408	3,46025
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Butyrivibrio	0,05207	0	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Butyrivibrio s__Butyrivibrio_crossotus	0,05207	0	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Butyrivibrio s__Butyrivibrio_crossotus t__GCF_000156015	0,05207	0	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Coprococcus	1,07538	1,10019	0,53527
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Coprococcus s__Coprococcus_catu	0,69245	0,38694	0,22752
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Coprococcus s__Coprococcus_catu t__GCF_000210555	0,69245	0,38694	0,22752
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Coprococcus s__Coprococcus_comes	0,38293	0,71325	0,30775
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Coprococcus s__Coprococcus_comes t__GCF_000155875	0,38293	0,71325	0,30775
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Dorea	6,26022	3,64801	3,81755
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Dorea s__Dorea_formicigenerans	1,98616	1,04612	1,37814
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Dorea s__Dorea_formicigenerans t__Dorea_formicigenerans_unclassified	1,98616	1,04612	1,37814
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Dorea s__Dorea_longicatena	4,27406	2,60189	2,43941

k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Dorea s__Dorea_longicatena t__GCF_000154065	4,27406	2,60189	2,43941
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname	0,85346	0,48676	1,8476
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_2_1_58FAA	0	0,20498	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_2_1_58FAA t__GCF_000218465	0	0,20498	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_3_1_46FAA	0	0,12385	0,74381
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_3_1_46FAA t__GCF_000209405	0	0,12385	0,74381
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_5_1_57FAA	0,01076	0	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_5_1_57FAA t__GCF_000218425	0,01076	0	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_5_1_63FAA	0,03273	0,15793	0,18088
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_5_1_63FAA t__GCF_000185525	0,03273	0,15793	0,18088
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_8_1_57FAA	0,80998	0	0,92291
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Lachnospiraceae_noname s__Lachnospiraceae_bacterium_8_1_57FAA t__GCF_000185545	0,80998	0	0,92291
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia	0,20026	2,9709	0,32498

k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia s__Roseburia_hominis	0	0,00916	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia s__Roseburia_hominis t__GCF_000225345	0	0,00916	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia s__Roseburia_intestinalis	0	1,9004	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia s__Roseburia_intestinalis t__Roseburia_intestinalis_unclassified	0	1,9004	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia s__Roseburia_inulinivorans	0,20026	1,06134	0,0811
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia s__Roseburia_inulinivorans t__GCF_000174195	0,20026	1,06134	0,0811
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Lachnospiraceae g__Roseburia s__Roseburia_unclassified	0	0	0,24388
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Peptostreptococcaceae	0,44368	0,66639	1,51661
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Peptostreptococcaceae g__Peptostreptococcaceae_noname	0,44368	0,66639	1,51661
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Peptostreptococcaceae g__Peptostreptococcaceae_noname s__Clostridium_bartlettii	0,13137	0,21712	0,07297
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Peptostreptococcaceae g__Peptostreptococcaceae_noname s__Clostridium_bartlettii t__GCF_000154445	0,13137	0,21712	0,07297
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Peptostreptococcaceae g__Peptostreptococcaceae_noname s__Peptostreptococcaceae_noname_unclassified	0,3123	0,44927	1,44364
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae	17,05868	18,43187	12,00388
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Faecalibacterium	4,01568	5,21571	1,93749

k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Faecalibacterium s__Faecalibacterium_prausnitzii	4,01568	5,21571	1,93749
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Faecalibacterium s__Faecalibacterium_prausnitzii t__Faecalibacterium_prausnitzii_unclassified	4,01568	5,21571	1,93749
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus	9,62825	12,22348	8,87403
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_bromii	5,76443	5,10812	8,42335
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_bromii t__GCF_000209875	5,76443	5,10812	8,42335
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_callidus	0	0	0,02868
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_callidus t__GCF_000468015	0	0	0,02868
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_lactaris	0,24136	0,39415	0,422
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_lactaris t__Ruminococcus_lactaris_unclassified	0,24136	0,39415	0,422
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_sp_5_1_39BFAA	3,62246	6,7212	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Ruminococcus s__Ruminococcus_sp_5_1_39BFAA t__GCF_000159975	3,62246	6,7212	0
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Subdoligranulum	3,41475	0,99269	1,19236
k__Bacteria p__Firmicutes c__Clostridia o__Clostridiales f__Ruminococcaceae g__Subdoligranulum s__Subdoligranulum_unclassified	3,41475	0,99269	1,19236
k__Bacteria p__Firmicutes c__Erysipelotrichia	8,41685	4,73809	5,64574

k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales	8,41685	4,73809	5,64574
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae	8,41685	4,73809	5,64574
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Catenibacterium	4,215	1,70956	3,95226
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Catenibacterium s__Catenibacterium_mitsuokai	4,215	1,70956	3,95226
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Catenibacterium s__Catenibacterium_mitsuokai t__GCF_000173795	4,215	1,70956	3,95226
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Coprobacillus	0	0,25385	0,01381
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Coprobacillus s__Coprobacillus_unclassified	0	0,25385	0,01381
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Erysipelotrichaceae__noname	4,20186	2,77468	1,67967
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Erysipelotrichaceae__noname s__Clostridium_amosum	0	0,01301	0
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Erysipelotrichaceae__noname s__Clostridium_amosum t__GCF_000154485	0	0,01301	0
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Erysipelotrichaceae__noname s__Eubacterium_biforme	4,20186	2,76167	1,67967
k__Bacteria p__Firmicutes c__Erysipelotrichia o__Erysipelotrichales f__Erysipelotrichaceae g__Erysipelotrichaceae__noname s__Eubacterium_biforme t__GCF_000156655	4,20186	2,76167	1,67967
k__Bacteria p__Firmicutes c__Negativicutes	0,55725	4,1107	0,25823
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales	0,55725	4,1107	0,25823
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae	0,55725	4,1107	0,25823

k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Dialister	0,28905	0,06571	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Dialister s__Dialister_succinatiphilus	0,28905	0,06571	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Dialister s__Dialister_succinatiphilus t__GCF_000242435	0,28905	0,06571	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas	0	1,63329	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas s__Megamonas_funiformis	0	0,1385	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas s__Megamonas_funiformis t__GCF_000245775	0	0,1385	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas s__Megamonas_hypermegale	0	0,43642	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas s__Megamonas_hypermegale t__GCF_000209975	0	0,43642	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas s__Megamonas_rupellensis	0	0,05465	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas s__Megamonas_rupellensis t__GCF_000378365	0	0,05465	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megamonas s__Megamonas_unclassified	0	1,00372	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megasphaera	0	0,94692	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Megasphaera s__Megasphaera_unclassified	0	0,94692	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Mitsuokella	0,22663	0,11347	0,25823

k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Mitsuokella s__Mitsuokella_multacida	0	0	0,08023
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Mitsuokella s__Mitsuokella_multacida t__GCF_000155955	0	0	0,08023
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Mitsuokella s__Mitsuokella_unclassified	0,22663	0,11347	0,17801
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Veillonella	0,04157	1,35132	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Veillonella s__Veillonella_dispar	0	0,35958	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Veillonella s__Veillonella_dispar t__GCF_000160015	0	0,35958	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Veillonella s__Veillonella_parvula	0	0,01021	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Veillonella s__Veillonella_parvula t__Veillonella_parvula_unclassified	0	0,01021	0
k__Bacteria p__Firmicutes c__Negativicutes o__Selenomonadales f__Veillonellaceae g__Veillonella s__Veillonella_unclassified	0,04157	0,98153	0
k__Bacteria p__Proteobacteria	0,58288	2,32501	0,16604
k__Bacteria p__Proteobacteria c__Gammaproteobacteria	0,58288	2,32501	0,16604
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Enterobacteriales	0,58288	0,72956	0,16604
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Enterobacteriales f__Enterobacteriaceae	0,58288	0,72956	0,16604
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Enterobacteriales f__Enterobacteriaceae g__Escherichia	0,58288	0,72956	0,16604
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Enterobacteriales f__Enterobacteriaceae g__Escherichia s__Escherichia_coli	0,40142	0,71902	0

k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Enterobacteriales f__Enterobacteriaceae g__Escherichia s__Escherichia_coli t__Escherichia_coli_unclassified	0,40142	0,71902	0
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Enterobacteriales f__Enterobacteriaceae g__Escherichia s__Escherichia_unclassified	0,18146	0,01054	0,16604
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Pasteurellales	0	1,59545	0
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Pasteurellales f__Pasteurellaceae	0	1,59545	0
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Pasteurellales f__Pasteurellaceae g__Actinobacillus	0	0,06121	0
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Pasteurellales f__Pasteurellaceae g__Actinobacillus s__Actinobacillus_unclassified	0	0,06121	0
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Pasteurellales f__Pasteurellaceae g__Haemophilus	0	1,53424	0
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Pasteurellales f__Pasteurellaceae g__Haemophilus s__Haemophilus_parainfluenzae	0	1,53424	0
k__Bacteria p__Proteobacteria c__Gammaproteobacteria o__Pasteurellales f__Pasteurellaceae g__Haemophilus s__Haemophilus_parainfluenzae t__Haemophilus_parainfluenzae_unclassified	0	1,53424	0
k__Bacteria p__Spirochaetes	0	0,01566	0
k__Bacteria p__Spirochaetes c__Spirochaetia	0	0,01566	0
k__Bacteria p__Spirochaetes c__Spirochaetia o__Spirochaetales	0	0,01566	0
k__Bacteria p__Spirochaetes c__Spirochaetia o__Spirochaetales f__Spirochaetaceae	0	0,01566	0
k__Bacteria p__Spirochaetes c__Spirochaetia o__Spirochaetales f__Spirochaetaceae g__Treponema	0	0,01566	0
k__Bacteria p__Spirochaetes c__Spirochaetia o__Spirochaetales f__Spirochaetaceae g__Treponema s__Treponema_succinifaciens	0	0,01566	0
k__Bacteria p__Spirochaetes c__Spirochaetia o__Spirochaetales f__Spirochaetaceae g__Treponema s__Treponema_succinifaciens t__GCF_000195275	0	0,01566	0

k__Bacteria p__Verrucomicrobia	2,06744	1,83807	1,98916
k__Bacteria p__Verrucomicrobia c__Verrucomicrobiae	2,06744	1,83807	1,98916
k__Bacteria p__Verrucomicrobia c__Verrucomicrobiae o__Verrucomicrobiales	2,06744	1,83807	1,98916
k__Bacteria p__Verrucomicrobia c__Verrucomicrobiae o__Verrucomicrobiales f__Verrucomicrobiaceae	2,06744	1,83807	1,98916
k__Bacteria p__Verrucomicrobia c__Verrucomicrobiae o__Verrucomicrobiales f__Verrucomicrobiaceae g__Akkermansia	2,06744	1,83807	1,98916
k__Bacteria p__Verrucomicrobia c__Verrucomicrobiae o__Verrucomicrobiales f__Verrucomicrobiaceae g__Akkermansia s__Akkermansia_muciniphila	2,06744	1,83807	1,98916
k__Bacteria p__Verrucomicrobia c__Verrucomicrobiae o__Verrucomicrobiales f__Verrucomicrobiaceae g__Akkermansia s__Akkermansia_muciniphila t__GCF_000020225	2,06744	1,83807	1,98916
k__Viruses	0	0,02606	0
k__Viruses p__Viruses_noname	0	0,02606	0
k__Viruses p__Viruses_noname c__Viruses_noname	0	0,02606	0
k__Viruses p__Viruses_noname c__Viruses_noname o__Caudovirales	0	0,02606	0
k__Viruses p__Viruses_noname c__Viruses_noname o__Caudovirales f__Siphoviridae	0	0,02606	0
k__Viruses p__Viruses_noname c__Viruses_noname o__Caudovirales f__Siphoviridae g__C2likevirus	0	0,02606	0
k__Viruses p__Viruses_noname c__Viruses_noname o__Caudovirales f__Siphoviridae g__C2likevirus s__C2likevirus_unclassified	0	0,02606	0