

En science forensique, les traces d'ADN sont des indices pouvant être déterminants pour la résolution d'une enquête, notamment lorsque le profil ADN d'un suspect concorde avec celui retrouvé sur une scène de crime. Toutefois, il est nécessaire pour l'enquête d'évaluer le poids de cette concordance, c'est-à-dire l'hypothèse qu'elle soit purement fortuite. Une concordance fortuite survient lorsque la trace est laissée par un inconnu qui possède le même profil génétique que le suspect. La probabilité de concordance fortuite dépend des fréquences des diverses variantes d'un gène existant dans la population d'intérêt pour l'enquête. Un échantillon génétique de référence, collecté dans cette population d'intérêt, sert de banque d'ADN pour évaluer la rareté d'un profil génétique. Les laboratoires d'expertise disposent d'échantillons d'ADN séparés pour des groupes ethniques différents: Caucasiens, Asiatiques, Noirs, etc. Cette façon de constituer une banque pose deux problèmes : elle est valable seulement lorsqu'on a affaire à des populations ethniquement homogènes et ces banques ne reflètent pas la composition génétique des traces laissées dans l'environnement par les activités humaines au quotidien.

Mon hypothèse est que les fréquences d'allèles dans les traces d'ADN d'une population d'intérêt cosmopolite diffèrent des fréquences d'allèles fournies par les banques de référence typiques. L'objectif général de ce projet est de tester cette hypothèse en utilisant une nouvelle approche pour 1) mieux définir la population d'intérêt à l'enquête et son impact sur le poids statistique de la preuve et 2) établir une banque d'ADN de référence à partir de traces prélevées aléatoirement dans cette population.

Il a d'abord été nécessaire d'échantillonner l'ADN humain présent naturellement dans l'environnement. Montréal, une cosmopolite, a été ciblée par cette collecte. Préalablement, il importait de bien définir les zones d'échantillonnage afin de délimiter la population d'intérêt en fonction de l'échelle de la criminalité, et ce, à l'aide de méthodes d'analyse criminelle. Des zones d'échantillonnage ont été ainsi identifiées à partir de la variable de densité des différents types d'infractions au travers des nombreux postes de quartiers. Un total de 700 coordonnées géographiques déterminées aléatoirement ont été générées sur les territoires sélectionnés. À chaque point, un seul objet délaissé (verres jetables, gommes, mégots de cigarettes, etc.) a été récolté comme source d'ADN dans un rayon de 100 mètres. L'ADN était ensuite extrait, amplifié puis analysé en laboratoire. Ensuite, les données ont été comparées à celles de la banque de référence du Laboratoire de sciences judiciaires et de médecine légale pour mesurer la différence quantitative dans la rareté des profils génétiques (probabilités de concordance fortuite). Des analyses graphiques de l'information génétique ont permis d'établir la distribution des fréquences d'allèles, et des analyses des données génotypiques multi-locus ont permis d'étudier la structure d'une population d'échantillonnage. Enfin, ces données pourront être comparées à de futurs échantillons pour suivre l'évolution génétique des populations.

Les laboratoires judiciaires utilisent des banques d'ADN de référence de populations qui ne sont pas représentatives de la population d'intérêt pour un cas criminel particulier. Mes travaux permettront de vérifier la validité de cette approche et d'évaluer l'alternative qui est d'échantillonner des traces d'ADN plutôt que des individus volontaires, pour mieux cibler la population d'intérêt. Ce projet s'avère particulièrement important dans le contexte de la mobilité à grande échelle des individus dans les sociétés modernes. Ce projet aura certainement de grandes retombées, car elle sera applicable dans les nombreuses grandes villes multiethniques, ou dans de nombreuses villes où les variations génétiques ne sont pas en lien avec la séparation des territoires.

Septembre à décembre 2019

- Revue de littérature sur la génétique forensique et l'impact des bases de données sur l'évaluation de la preuve ADN
- Collecte d'informations sur la criminalité dans la ville de Montréal et analyse terrain en fonction des postes de quartiers desservis par le Service de Police de la Ville de Montréal (SPVM)
- Cours de maîtrise de biologie cellulaire avancée (3 crédits)

Janvier à avril 2020

- Apprentissage du logiciel R (langage de programmation)
- Développement de cartes de données dans le but de déterminer les zones d'échantillonnage ciblées par la criminalité (point chaud de la criminalité)
- Apprentissage du logiciel QGIS afin de créer, éditer, visualiser et analyser des données géospatiales
- Cours de maîtrise de biologie moléculaire avancée (3 crédits)

Mai à août 2020

- Apprentissage du système géodésique mondial sur l'utilisation du Système de positionnement global (Global positioning system, GPS)
- Préparation de la période de terrain intensif (collecte de données)
- Collecte de données (traces ADN) dans les postes de quartiers de la ville de Montréal préalablement sélectionnés (durée de 2 mois)

Septembre à décembre 2020

- Analyses de laboratoire :
- Extractions des échantillons d'ADN avec le kit DNA IQ
 - Tests de dosage des profils génétiques à partir des traces ADN par PCR en temps réel, QuantStudio™ 5

Plusieurs problématiques sont survenues au niveau des protocoles et beaucoup de délais engendrés par la COVID (approvisionnement de matériel et attente de résultats par le LSJML) ont nécessité de nouvelles analyses allant des mois de janvier à août 2021.

Janvier à avril 2021

- Nouvelles analyses de laboratoire sur de nouveaux échantillons d'ADN
- Envoie des échantillons au Laboratoire de sciences judiciaires et de médecine légale (LSJML)
- Cours de rédaction et communication scientifique en biologie cellulaire et moléculaire (3 crédits)

Mai à août 2021

- Nouvelles analyses de laboratoire sur de nouveaux échantillons d'ADN
- Envoie des échantillons au Laboratoire de sciences judiciaires et de médecine légale (LSJML)

Septembre à décembre 2021

- Analyse statistique de données compilées à partir de l'échantillonnage via le logiciel R : Test de différenciation, test de déficit d'hétérozygotie, analyse de fréquences alléliques
- Analyses comparatives pour mesurer la différence quantitative dans la rareté des profils génétiques : valeur de probabilité de concordance fortuite (banque de données du LSJML vs banque de données générée à partir de l'échantillonnage)
- Apprentissage du Logiciel Structure, programme permettant d'utiliser des données génotypiques multi-locus pour étudier la structure d'une population d'échantillonnage
- Présentation du séminaire de maîtrise, présentation d'une communication orale sur mon sujet de recherche (3 crédits)

Janvier à avril 2022

- Rédaction du mémoire ainsi que d'un article scientifique
- Soumission et dépôt du mémoire final (prévue pour avril 2022)