

Le développement d'outils de généalogie génétique adaptés au contexte endogame : son utilité pour la recherche sur les migrations, par James Smith, président de l'association Atelier Généalogique, Marseille, France.

Dans mes recherches en généalogie génétique à partir de 2016, j'ai rencontré le problème de l'endogamie¹ en cherchant les parents biologiques d'une élève en généalogie et amie, Francine: en bref, l'endogamie produit une surabondance de cousins et d'homonymes, brouillant les pistes.

Six ans de recherches intensives - littéralement des milliers d'heures de travail - ont abouti à la création d'un grand arbre généalogique centré sur un lieu, Corato (ainsi que sur quelques villes environnantes), dans les Pouilles, dans le sud de l'Italie. En 2022, l'arbre généalogique compte près de 24 000 personnes, étayé par une base de données de plus de 20 000 photos d'actes authentiques, stockée et consultable sur Google Drive. Ces chiffres ont été atteints en construisant des arbres généalogiques pour plus de 100 personnes, pour la plupart des correspondances ADN de Francine. Parmi elles, plus de 60 étaient des Américains d'origine italienne. Des centaines de chercheurs sur divers sites généalogiques ont copié des extraits de notre arbre, *l'albero genealogico coratino*² (à partir des sites Ancestry, My Heritage, Geneanet, Filae).

Notre projet a une visée à la fois historique et généalogique. Aucune recherche approfondie sur l'histoire familiale des habitants de Corato ne peut ignorer le phénomène migratoire, et la recherche généalogique menée par les descendants d'émigrés est beaucoup plus riche si elle inclut le pays d'origine des personnes. Les Coratins ont commencé à émigrer vers les États-Unis à partir de 1902, mais la France a pris le relais comme principale destination à partir de 1920 environ.

Francine et moi avons traduit et fait publier en version franco-italienne un livre de l'historien Pasquale Tandoi intitulé "Quando i clandestini eravamo noi". Ce livre est toujours disponible sur demande, et une traduction anglaise, *When we were the illegal immigrants*, est disponible sur notre site web³. L'information sur nos recherches généalogiques s'est répandue et a donné naissance à un projet intitulé "Un conte de deux villes jumelées : Corato et Grenoble". Il s'agit essentiellement d'améliorer l'accès aux ressources utiles à la fois à la recherche généalogique et aux études sur la migration à partir de Corato, et plus généralement de promouvoir des échanges entre les descendants d'émigrés coratins, d'encourager la rédaction de témoignages, de faciliter l'accès aux archives *in loco*...

L'objectif principal de notre projet est de promouvoir l'écriture de l'histoire de l'émigration coratine vers la région grenobloise. Une partie non négligeable de la population de Corato⁴ a opté d'émigrer vers la France, s'installant pour la plupart à Grenoble et aux alentours, et on dit souvent que les Coratins ont beaucoup contribué au développement industriel, socio-économique et culturel de cette grande ville française. En même temps, nous sommes conscients du fait que des membres des mêmes familles qui se sont installées dans la région grenobloise ont émigré vers d'autres destinations (États-Unis, Venezuela, Brésil, Allemagne, etc), fournissant ainsi sans doute de nombreux exemples du phénomène complexe des "familles transnationales"⁵.

Aujourd'hui, notre projet est suivi par un comité scientifique sous la présidence du professeur Biagio Salvemini, professeur titulaire d'histoire moderne à l'université Aldo Moro de Bari. Un exemple de notre travail a été présenté au programme Migrations de l'Université d'Oxford en mai 2022⁶.

¹ Un calcul du niveau d'endogamie à Corato, basé sur l'étude des mariages de la période 1835-65, réalisé par l'Atelier Généalogique en collaboration avec l'association VISUGED, a montré que Corato était endogame à 97%.

² <https://www.ancestry.co.uk/family-tree/tree/167791321/recent?usePUBJs=true>

³ <https://www.emigrazione-corato.org/medias/files/when-we-were-the-illegal-immigrants2.pdf>

⁴ Tandoi a constaté que 7 450 personnes, sur un total de 13 066 émigrants ayant demandé un passeport, se sont rendues en France entre 1920 et 1959... et il n'a pas compté les immigrants illégaux. En 1921, la population de Corato était de 50 010 (cf Tandoi), en 2019, de 48 253.

⁵ Cf *Famiglie transnazionali*, in P.Corti, M.Sanfilippo, a cura di, *Storia d'Italia. Annali 24, Migrazioni*, Torino, Einaudi, 2009, pp.303-316.

⁶ <https://www.emigrazione-corato.org/medias/files/summaries-submitted-for-publication-on-the-compass-blog.pdf>

Une expérience visant à mieux identifier les cousins pertinents dans un contexte endogame

Wikitree⁷, un site web généalogique américain, est actuellement engagé dans le développement d'une application appelée "X-friends" (concepteur : Greg Clarke) qui pourrait bien être une contribution significative, non seulement à la recherche généalogique génétique mais aussi aux études de migration basées sur la "mémoire" de la migration, en d'autres termes, sur l'histoire familiale. Cette approche bi-disciplinaire visant à recueillir les témoignages des personnes ayant émigré serait d'actualité d'un point de vue historiographique. Plus de 100 ans après le principal flux migratoire, de nombreux tabous concernant le récit de la migration et de ses multiples facettes sont en train d'être levés.

L'expérience généalogique que nous menons en collaboration avec Wikitree concerne la poursuite de la recherche des parents biologiques de Francine en adaptant les méthodes de recherche en généalogie génétique pour tenir compte du contexte endogame de Corato. Notre hypothèse est que la méthodologie, si elle s'avère fructueuse, pourrait être pertinente pour la recherche et l'identification de toute personne habitant un lieu endogame.

En principe, étant donné que le schéma de transmission du chromosome X est très spécifique (le père ne transmet qu'à ses filles, la mère à tous ses enfants), la recherche de porteurs potentiels de l'X - à la fois 'vers le haut', dans l'ascendance, jusqu'à plusieurs "ancêtres terminaux" et 'vers le bas' jusqu'à toutes les personnes qui ont pu hériter de l'X de ces derniers - réduira le nombre de cousins pertinents à prendre en considération lors de la recherche, en éliminant certaines lignées masculines généalogiques.

Les concordances potentielles sur le chromosome X peuvent être utilement recherchées pour n'importe qui dans un arbre généalogique si ce dernier est suffisamment grand et construit avec précision. La précision est, bien sûr, essentielle : si l'arbre généalogique est défectueux, les concordances X potentielles seront imaginaires. Les normes strictes d'enregistrement des sources adoptées par Wikitree devraient normalement empêcher cela, et un outil de confirmation d'ADN développé conjointement par Wikitree et Gedmatch, ainsi que d'autres outils de généalogie génétique, devraient contribuer à éviter les erreurs.

La procédure technique de l'expérience d'X-friends peut être décrite comme suit :

1. A partir de l'*albero genealogico coratino*, pour chacune des principales 'correspondances ADN' (*DNA matches*) autosomales (80cM⁸ ou plus) du sujet (c'est-à-dire Francine), le logiciel Family Tree Maker exporte un Gedcom⁹ de ses ancêtres directs et de leurs descendants sur 9 générations.
2. Avec son application X-friends¹⁰, Wikitree extrait de chacun des gedcoms un fichier .csv contenant les noms des concordances X potentielles identifiées par une seule opération de bas vers le haut, puis du haut vers le bas du "widget" conçu pour suivre les trajectoires de transmission de ces chromosomes X. Ce fichier .csv est créé d'abord à partir d'une personne ayant fait un test ADN sélectionnée dans l'arbre généalogique produit par le gedcom.
3. La fusion de plusieurs de ces fichiers .csv – convertis en fichiers Excel .xlsx - devrait permettre d'identifier les personnes potentiellement porteuses des mêmes chromosomes X que Francine à partir de plusieurs sources, et ainsi indiquer quelles sont les lignées généalogiques les plus prometteuses.

La procédure décrite ci-dessus a été répétée 16 fois, produisant 16 fichiers .csv. Le tableau d'ascendance ci-dessous, créé après fusion des fichiers, illustre comment des cousins pertinents

⁷ Wikitree est une société américaine décrite comme un " site web de généalogie libre, partagé et social qui permet aux utilisateurs de faire des recherches et de contribuer à leurs propres arbres généalogiques personnels tout en construisant et en collaborant à un arbre généalogique mondial unique au sein du même système " (Wikipedia).

⁸ cM : centimorgans, unité de mesure de l'ADN partagé avec une autre personne

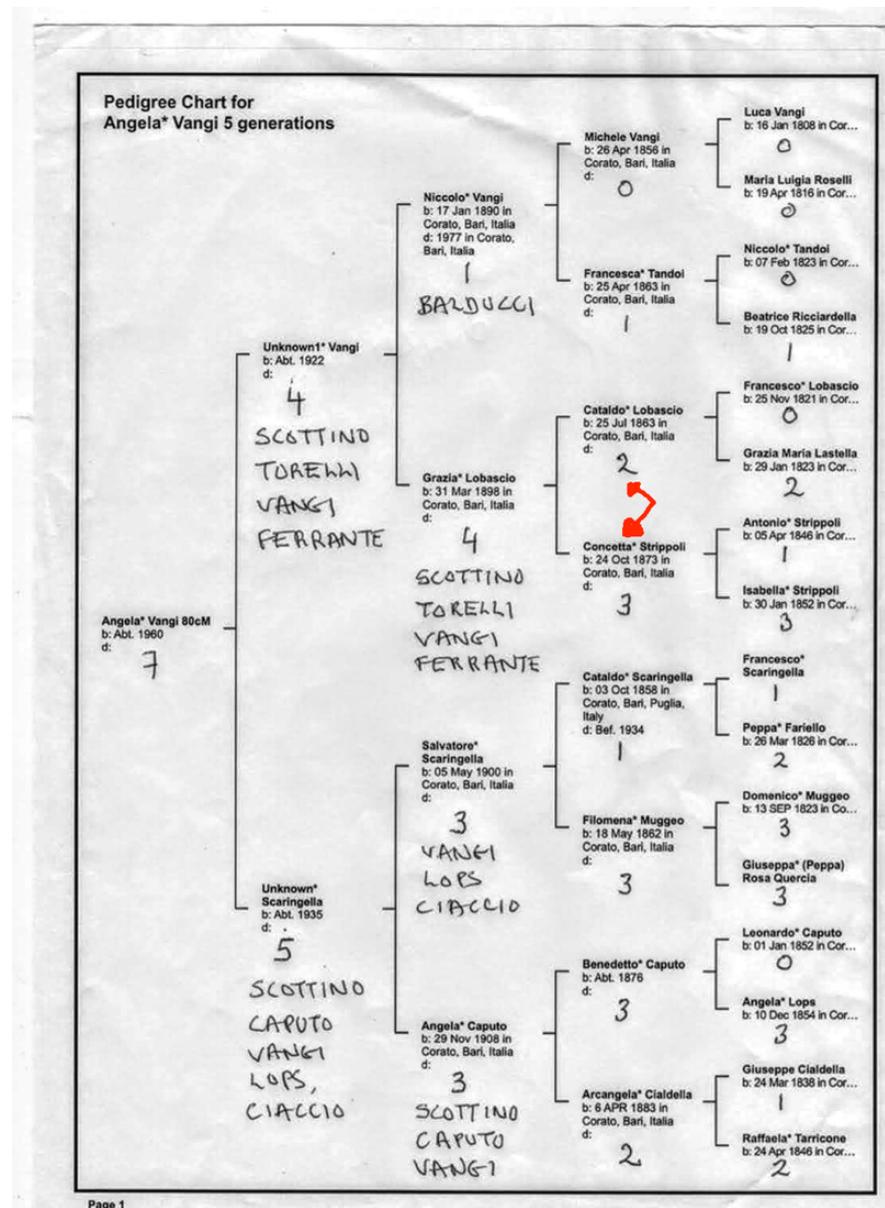
⁹ Les fichiers GEDCOM sont des fichiers séquentiels en format texte, qui permettent les échanges de données informatisées entre les différents logiciels de généalogie. Un Gedcom est créé soit à partir d'un arbre généalogique, permettant de transférer cet arbre d'un logiciel de généalogie à un autre ; soit à partir d'un fichier .csv approprié (voir plus loin).

¹⁰ <https://apps.wikitree.com/apps/clarke11007/Xfriends.php>

peuvent être identifiés en trouvant des personnes potentiellement porteuses d'une concordance X avec Francine à partir de différentes sources. Angela Vangi, par exemple, qui est une correspondance autosomale de 80cM avec Francine, partage potentiellement avec elle des chromosomes X provenant de 7 sources différentes.

Cependant, l'application X-friends est capable d'identifier tous les porteurs potentiels des mêmes chromosomes X parmi les descendants des ancêtres de n'importe quelle personne, et pas seulement les personnes qui ont passé un test ADN. Le tableau ci-dessous indique, par exemple, que Cataldo Lobascio et Concetta Strippoli sont des porteurs potentiels du X à partir de 2 et 3 sources respectivement. En construisant l'arbre généalogique de ce couple particulier, et en développant sa descendance, il est possible d'identifier d'autres personnes qui étaient (ou sont) des porteurs potentiels de l'X à partir de plusieurs sources, mais qui n'ont pas fait de test ADN.

Notre hypothèse est qu'en reconvertissant les fichiers .csv en gedcoms et en créant ainsi un arbre généalogique spécifique dans lequel les personnes ayant effectué un test ADN sont incluses, un modèle peut émerger indiquant les lignées généalogiques les plus prometteuses des ancêtres directs de Francine. Notre hypothèse est aussi que l'application X-friends pourrait aider à vérifier l'identité des émigrés d'une manière similaire.



Les enjeux du développement de l'application X-friends par Wikitree

L'expérience avec Wikitree est en cours ; en d'autres termes, le processus expérimental décrit ci-dessus n'est pas encore terminé. La reconversion des fichiers .csv des porteurs potentiels de concordances X en gedcoms¹¹ permettra, après la fusion des arbres généalogiques ainsi produits, de visualiser un arbre généalogique spécialement ciblé. Si l'expérience est concluante, notre *albero genealogico coratino* pourrait être utilisé par toute personne recherchant l'histoire de sa famille à Corato ou dans les pays de destination. La même méthodologie pourrait être appliquée avec n'importe quel grand arbre basé sur l'étude d'un lieu unique endogame. Elle pourrait s'inscrire dans une approche bi-disciplinaire de l'histoire de l'émigration, en reliant les perspectives généalogiques et historiques. Notre hypothèse est donc que la généalogie peut trouver sa place dans une conception de l'histoire basée sur la "mémoire", en identifiant les migrants et leurs descendants capables et désireux de témoigner.

Dans le contexte italien, le processus d'identification est rendu difficile non seulement par l'endogamie mais aussi par le manque d'accès aux registres paroissiaux et les restrictions imposées par le Stato Civile - où les registres des naissances ne sont librement accessibles que jusqu'en 1910 - ainsi que par la législation sur ce que les Italiens appellent la *privacy* (dont une partie semble être de plus en plus imposée par l'UE). Les registres du Stato Civile comportent également de nombreuses imprécisions et des cas d'erreurs d'identité, aggravés par l'analphabétisme. La capacité d'identifier les personnes au cours de la période 1910-1930 est particulièrement importante, car c'est une époque pendant laquelle de nombreux migrants sont nés.

Fondamentalement, pour que l'histoire de la migration puisse être écrite sur la base de la "mémoire" de celle-ci, c'est-à-dire à partir d'histoires familiales, la recherche généalogique doit trouver des moyens de surmonter ces restrictions. Bien sûr, il existe d'autres moyens d'accéder à l'information généalogique : les manifestes des navires, les permis de travail et de séjour périmés, les recensements, les mémoires et les thèses universitaires, les articles de journaux, les données de naturalisation et les données militaires, les archives de famille, etc. En tant que moyen de vérifier l'identité, l'application "X-friends" de Wikitree pourrait bien être un complément très utile aux outils de généalogie génétique existants.

Marseille, octobre 2022

¹¹ L'Atelier Généalogique y travaille avec Jean Chabaud, le concepteur du logiciel VISUGED (<https://www.visuged.org/index.htm>) qui a été utilisé en mars 2022 pour calculer le niveau d'endogamie à Corato.