

L'analyse du chromosome Y

Christine Keyser-Tracqui & Bertrand Ludes

Institut de Médecine Légale, 11, rue Humann, 67085 Strasbourg Cedex

Reçu le 6 juin 2003

RÉSUMÉ

Malgré des difficultés initiales dans l'identification de polymorphismes génétiques de la partie non recombinante du chromosome Y humain, un grand nombre de marqueurs polymorphes sont aujourd'hui disponibles pour établir des haplotypes ou haplogroupes caractéristiques d'individus ou de populations. Parmi ces marqueurs figurent les STR (Short

Tandem Repeats) du chromosome Y, largement utilisés au cours des dernières années pour étudier l'évolution de l'Homme et identifier des individus ou des filiations en médecine légale. Après un bref rappel des particularités de ces marqueurs, nous présentons certaines des applications du typage génétique du chromosome Y.

SUMMARY Y-Chromosome analysis

After initially slow progress in identifying DNA polymorphisms on the non recombining part of the human Y chromosome, a large number of polymorphic markers are now available to define individual or population-specific haplotypes or haplogroups. Among them the Y-chromosomal STRs (Short Tan-

dem Repeats) have been increasingly used over the past few years for the study of human evolution as well as for human identification in forensic case-work and paternity testing. After a brief summary of the features of such markers, this paper deals with some applications of the Y-STR haplotyping.

INTRODUCTION

Présent en principe uniquement chez les individus de sexe masculin, le chromosome Y se caractérise par une transmission uniparentale puisque seuls les pères transmettent ce chromosome à leur descendance masculine. À l'exception de deux petites régions distales homologues du chromosome X (appelées régions pseudo-autosomales), le chromosome Y est constitué en majorité de séquences d'ADN qui lui sont spécifiques (Fig. 1). De ce fait et contrairement aux autosomes (chromosomes non sexuels), il n'échange pas de segments d'ADN avec d'autres chromosomes au cours de la division cellulaire, autrement dit il ne se recombine pas ; il est par conséquent transmis sans modification d'une génération à la suivante.

Cette caractéristique permet de retracer des lignées paternelles sur des dizaines de générations puisqu'un individu de sexe masculin aura le même chromosome Y que son père, son grand-père paternel, le père de ce dernier et ainsi de suite. Les seules modifications observées correspondent à des mutations qui surviennent au cours de l'évolution. Le chromosome Y apparaît en quelque sorte comme le pendant de l'ADN mitochondrial

(ADNmt) dont la transmission est strictement maternelle. Il permet d'accéder à des informations complémentaires à celles fournies par l'ADNmt et est actuellement largement utilisé dans les études visant à retracer l'origine et l'histoire des populations humaines (Hurles *et al.*, 1999; Underhill *et al.*, 2000; Kayser *et al.*, 2001), identifier un individu ou encore établir une filiation (Jobling *et al.*, 1997), reconstruire des généalogies (Gerstenberger *et al.*, 1999; Joblin, 2000) ou établir l'étiologie de certaines maladies (Jobling & Tyler-Smith, 2000).

Contrairement à l'ADNmt, les études portant sur les variations de séquences du chromosome Y ont progressé assez lentement car peu de marqueurs polymorphes ont été mis en évidence dans les premiers temps (Malaspina *et al.*, 1990). Depuis, de nombreuses études ont révélé un polymorphisme équivalent à celui présent sur les autosomes, et un nombre important de marqueurs moléculaires permettent aujourd'hui de définir des haplotypes ou haplogroupes (combinaisons de marqueurs) variables d'un individu ou d'un groupe ethnique à l'autre. Parmi ces marqueurs il convient de distinguer : 1) les marqueurs bialléliques à faible taux de mutation, tels les SNP (Single Nucleotide Polymorphism); 2) les mar-

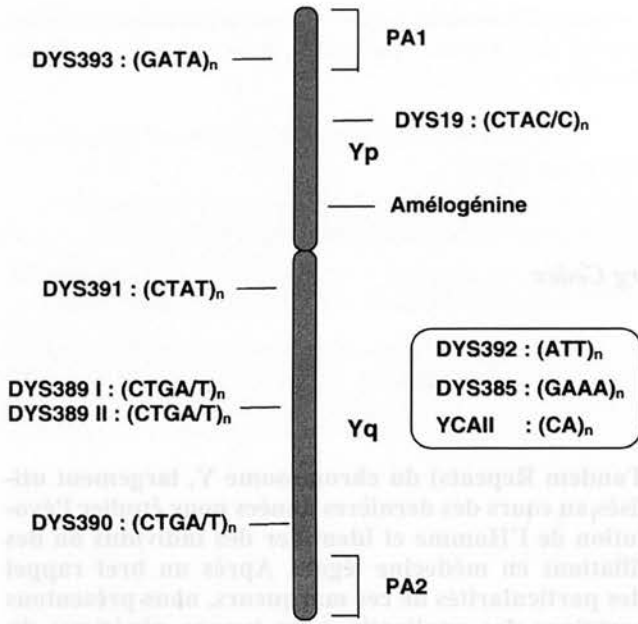


FIG. 1 – Seules les deux petites régions pseudoautosomales PA 1 et PA 2 du chromosome Y sont susceptibles d'échanges avec les chromosomes X au cours de la division méiotique. La grande majorité du chromosome Y porte des séquences d'ADN répété. La localisation de certaines d'entre elles est ici indiquée.

queurs multialléliques à taux de mutation plus élevé, tels les microsatellites ou STR (Short Tandem Repeat).

Alors que la stabilité des marqueurs bialléliques se révèle précieuse pour étudier l'histoire de l'Homme et les mouvements migratoires (Semino *et al.*, 2000; Rosser *et al.*, 2000; Zerjal *et al.*, 2002) celle, moindre, des microsatellites s'avère suffisante pour identifier des individus de même lignée paternelle dans le cadre d'investigations médico-légales ou d'études généalogiques et archéologiques. L'objet du présent article est de développer ces dernières applications en insistant sur les avantages et les limites d'une approche moléculaire basée sur l'analyse du chromosome Y.

Locus A (région particulière d'un chromosome donné)

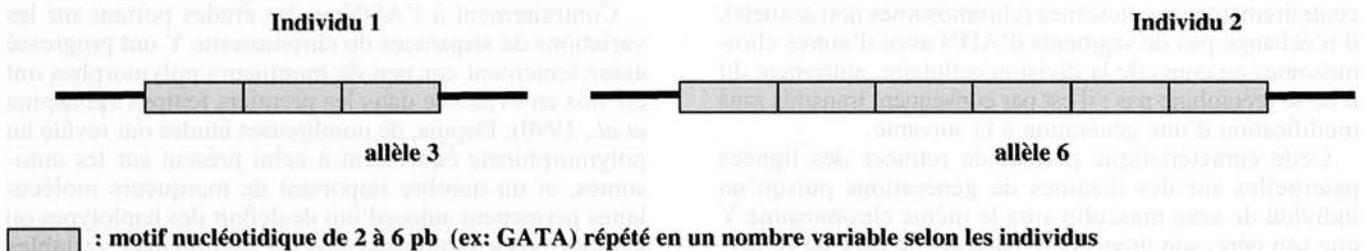


FIG. 2. – Localisés aussi bien sur les autosomes que sur les chromosomes sexuels, les microsatellites ou STR sont des séquences courtes (2 à 6 pb), non codantes, répétées les unes à la suite des autres en un nombre qui varie selon les individus. Cette variabilité du nombre de répétitions génère un polymorphisme de taille au niveau de la région étudiée. Cette différence de taille, mise en évidence au moyen de technologies de plus en plus performantes, constitue la base des techniques d'identification individuelle.

MARQUEURS MICROSATELLITES OU STR DU CHROMOSOME Y

Jusqu'à ces dernières années, l'analyse du chromosome Y en identification génétique avait pour seul objet la détermination du sexe d'un échantillon biologique donné. L'amplification d'une partie du gène de l'amélogénine, localisé sur les chromosomes sexuels (Fig. 1), permet en effet d'obtenir des fragments d'ADN dont la taille varie en fonction du sexe : un fragment de 106 pb est obtenu pour le chromosome X et un fragment de 112 pb pour le chromosome Y (Mannucci *et al.*, 1994).

Les travaux scientifiques entrepris par plusieurs équipes internationales au début des années 1990 ont permis d'élargir cette application : en décrivant la présence sur le chromosome Y de séquences d'ADN microsatellites également appelées STR (Fig. 2), ces équipes ont mis en évidence la possibilité d'utiliser le chromosome Y pour typer les individus de sexe masculin (Roewer *et al.*, 1992). Tous les marqueurs STR présents sur le chromosome Y ne sont pas discriminants, c'est-à-dire suffisamment polymorphes pour permettre de différencier deux individus non apparentés par leurs lignées paternelles. Parmi les STR retenus par la communauté scientifique internationale chargée de coordonner ces recherches figurent sept marqueurs tétranucléotidiques (DYS19, DYS385, DYS389I, DYS389II, DYS390, DYS392, DYS393), un marqueur trinuécléotidique (DYS392) et un marqueur dinucléotidique (YCAII). Ces neuf marqueurs polymorphes ont été sélectionnés après que leur variabilité ait été testée sur les principaux groupes ethniques (américain, européen, asiatique et africain). Ce travail consiste à établir, pour chacune des régions STR étudiées, la fréquence de chacun des allèles ou variants observés et de comparer la distribution allélique de ces différents marqueurs entre les diverses populations étudiées.

Deux caractéristiques fondamentales de l'analyse du chromosome Y sont apparues au cours de ces travaux. La première est qu'il existe une spécificité ethnique substantielle dans la variabilité des marqueurs amplifiés, en

TABLEAU I. – Étude de STR du chromosome Y réalisée sur une population française de 100 individus caucasiens (Keyser-Tracqui *et al.*, 2003).

Marqueurs STR étudiés	Nombre d'haplotypes	Pouvoir discriminant
DYS19 + DYS390 + DYS391 + DYS392 + DYS393	54	54 %
DYS19 + DYS389I/II + DYS390 + DYS391 + DYS392 + DYS393	68	68 %
DYS19 + DYS389I/II + DYS390 + DYS391 + DYS392 + DYS393 + DYS385	86	86 %
DYS19 + DYS389I/II + DYS390 + DYS391 + DYS392 + DYS393 + DYS385 + YCAII	91	91 %

d'autres termes, la fréquence des allèles peut varier en fonction de la population étudiée (Keyser *et al.*, 1997; Bosch *et al.*, 2000). La seconde est que le pouvoir informatif des STR portés par le chromosome Y est moindre que celui des STR autosomiaux. Ceci s'explique aisément par le fait que la transmission de l'information portée par le chromosome Y se fait « en bloc » d'une génération à la suivante : l'haplotype du père est transmis inchangé à sa descendance masculine et ne représente finalement la contribution génétique que d'une petite partie des ancêtres de cet individu.

Le tableau I illustre le « pouvoir discriminant » des régions STR habituellement utilisées autrement dit le nombre d'haplotypes différents retrouvés au sein d'une population donnée. L'étude de cinq régions STR permet de distinguer 54 % des individus non apparentés étudiés ; l'analyse de deux marqueurs supplémentaires permet d'atteindre un « pouvoir discriminant » de 86 % ; enfin, l'amplification de neuf régions STR permet une individualisation presque complète de la population étudiée (« pouvoir discriminant » équivalent à celui des régions hypervariables HVI et HVII de l'ADNmt).

Spécificité ethnique parfois importante, pouvoir informatif réduit, au vu de ces caractéristiques on comprend combien il est important de disposer de données de référence pour interpréter les résultats d'analyses effectuées à partir du chromosome Y. De fait, une collaboration internationale étroite s'est mise en place pour tenter de collecter efficacement des données issues du plus grand nombre possible de populations. Un effort particulier a été consacré à l'étude des populations européennes puisque le site Internet dédié au regroupement de données STR du chromosome Y issues de toute l'Europe (<http://ystr.charite.de>) abrite aujourd'hui des données issues de 81 populations. Pourquoi tous ces efforts ? Autrement dit dans quels cas l'analyse du chromosome Y se révèle-t-elle intéressante en pratique ?

INVESTIGATIONS MÉDICO-LÉGALES

En criminalistique, l'intérêt de l'analyse du chromosome Y réside dans l'analyse de mélanges d'ADN. Dans les cas de viols par exemple, les traces de sperme ne sont pas toujours faciles à mettre en évidence, car elles sont souvent mélangées à des cellules féminines provenant de la victime. Si le composant majeur du mélange est représenté par de l'ADN féminin, c'est celui-ci qui sera préférentiellement amplifié au détriment de l'ADN de l'agres-

seur. L'utilisation de marqueurs spécifiques du chromosome Y permet, dans certains cas, d'amplifier de manière sélective la composante masculine du mélange et d'établir ainsi l'haplotype de l'agresseur. En ce qui concerne les viols en réunion, l'analyse du chromosome Y permet de déterminer le nombre minimum d'agresseurs avec beaucoup plus de facilité que dans le cas des STR autosomiaux ; les profils haplotypiques sont plus simples à analyser que les profils génétiques classiques du fait de la présence d'un seul allèle par région STR étudiée.

L'analyse du chromosome Y présente également un intérêt dans les recherches en paternité, lorsqu'il s'agit d'établir un lien de filiation alors que le père présumé est décédé. Dans ce type de situation, il est possible d'avoir accès à l'information portée par le chromosome Y du père disparu en utilisant l'ADN de tout autre parent de même lignée paternelle (il peut s'agir du grand-père paternel, de l'oncle paternel, de son fils, etc.). Si l'haplotype de l'enfant soumis à l'analyse ne correspond pas à celui du parent testé, la paternité est exclue. Soulignons cependant que, dans cet exemple, l'exclusion de paternité peut provenir d'une génération antérieure. Il se peut très bien que le père décédé n'ait pas été le fils légitime du grand-père sollicité pour l'analyse tout en étant le père biologique de l'enfant testé.

L'analyse du chromosome Y présente donc des limites liées essentiellement aux difficultés d'interprétation des résultats en médecine légale. À l'instar des autres systèmes utilisés en identification génétique, le chromosome Y représente un outil d'investigation permettant d'exclure un suspect : si l'haplotype obtenu à partir d'une trace biologique diffère de celui obtenu à partir du prélèvement sanguin d'un suspect, il peut être affirmé avec certitude que la trace ne provient pas du suspect. Aucun risque d'erreur n'est affecté au résultat. Le problème est différent lorsqu'il s'agit d'inclure un suspect. Dans le cas des profils génétiques autosomiaux, le seul individu susceptible de présenter la même empreinte génétique qu'un suspect est son vrai jumeau (le cas est peu fréquent). Pour le Y en revanche, un haplotype identique peut être retrouvé chez toute personne de même lignée paternelle mais également par hasard dans la population dont est issu le suspect. Les inclusions sont par conséquent difficilement supportées par un calcul statistique. Quoi qu'il en soit, l'information selon laquelle le suspect ne peut être exclu par l'analyse peut constituer un élément utile à l'ins-truction. De la même façon, dans les tests de paternité, l'exclusion est affirmée si les haplotypes obtenus ne sont pas concordants. S'ils le sont, on ne peut exclure comme père tout autre personne de même lignée paternelle.

ÉTUDES GÉNÉALOGIQUES OU ARCHÉOLOGIQUES

Le potentiel des analyses du chromosome Y est également exploité dans des domaines historiques. L'un des exemples les plus retentissants concerne Thomas Jefferson, élu président des États-Unis en 1801. En 1998, la célèbre revue *Nature* publie un article relançant le débat sur les relations entre le troisième président américain et l'une de ses esclaves, Sally Hemings (Foster *et al.*, 1998). Les auteurs de cet article ont comparé l'haplotype Y de l'un des descendants de Field Jefferson, un oncle paternel de Thomas Jefferson, avec celui d'un descendant par voie paternelle de l'un des fils de Sally Hemings, Eston. Les résultats de ces recherches montrent clairement une identité entre les deux haplotypes. Les auteurs de l'article concluent que l'explication la plus probable est que l'ancien président Thomas Jefferson était bien le père de Eston Hemings, l'un des fils de Sally.

Dans les études explorant l'ADN extrait de tissus anciens, l'analyse du chromosome Y, bien que difficile à mettre en place du fait de la quantité et de la qualité souvent médiocre des molécules extraites, représente néanmoins une source d'informations précieuses pour appréhender les rites culturels ou familiaux des populations du passé. Ainsi, dans une étude consacrée à l'analyse moléculaire de restes anciens, exhumés d'une nécropole Xiongnu datant de plus de 2000 ans (site d'Egyin Gol, nord de la Mongolie), notre équipe a mis en évidence, sur la base d'haplotypes Y communs, des regroupements d'individus de même lignée paternelle au sein du cimetière. Ces analyses du chromosome Y ont également permis de préciser et confirmer des liens de proche parenté mis en évidence par les STR autosomaux (Crubézy *et al.*, 2002 ; Keyser-Tracqui *et al.*, 2003).

CONCLUSION

En médecine légale le typage du chromosome Y trouve un intérêt dans des cas particuliers où les STR autosomaux et l'ADNmt se révèlent inappropriés. Il s'agit d'une technique d'investigation complémentaire des systèmes classiquement utilisés, relativement peu développée dans les laboratoires d'identification génétique.

En anthropologie moléculaire, l'analyse des polymorphismes du chromosome Y permet de préciser les relations qui existent entre différents individus ou populations et accompagnent bien souvent des données issues de l'ADNmt. Elle permet dans certains cas de mettre en relief des différences liées au sexe notamment dans les schémas migratoires. Enfin, elle valide bien souvent les données obtenues à l'aide de marqueurs autosomaux ou mitochondriaux, dont l'authenticité est souvent mise en cause du fait de la nature dégradée de molécules d'acides nucléiques extraites de substrats anciens.

BIBLIOGRAPHIE

- Bosch E., Calafell F., Perez-Lezaun A., Comas D. *et al.*, Y chromosome STR haplotypes in four populations from northwest Africa. *Int. J. Legal Med.*, 2000, 114, 36-40.
- Crubézy E., Keyser C. & Ludes B., Les surprises de l'ADN ancien. *La Recherche*, 2002, 353, 44-47.
- Foster E. A., Jobling M. A., Taylor P. G., Donnelly P. *et al.*, Jefferson fathered slave's last child. *Nature*, 1998, 396, 27-28.
- Gerstenberger J., Hummel S., Schultes T., Häck B. & Herrmann B., Reconstruction of a historical genealogy by means of STR analysis and Y-haplotyping of ancient DNA. *Eur. J. Hum. Genet.*, 1999, 7, 469-477.
- Hurles M. E., Veitia R., Arroyo E., Armenteros M. *et al.*, Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. *Am. J. Hum. Genet.*, 2001, 65, 1437-1448.
- Jobling M. A., Pandya A. & Tyler-Smith C., The Y chromosome in forensic analysis and paternity testing. *Int. J. Leg. Med.*, 1997, 110, 118-124.
- Jobling M. A., In the name of the father: surnames and genetics. *Trends Genet.*, 2001, 17, 353-357.
- Jobling M. A. & Tyler-Smith C., New uses for new haplotypes. *TIG*, 2000, 16, 356-362.
- Kayser M., Krawczak M., Excoffier L., Dieltjes P. *et al.*, An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. *Am. J. Hum. Genet.*, 2001, 68, 990-1018.
- Kayser M., Caglia A., Corach D., Fretwell N. *et al.*, Evaluation of Y-chromosomal STRs: a multicenter study. *Int. J. Leg. Med.*, 1997, 110, 125-133.
- Keyser-Tracqui C., Crubézy E., & Ludes B., Nuclear and mitochondrial DNA analysis of a 2,000-year-old necropolis in the Egyin Gold valley of Mongolia. *Am. J. Hum. Genet.*, 2003, 73, 247-260.
- Keyser-Tracqui C., Ricaut F., Blandin P. & Ludes B., French allele frequencies and haplotypes of nine Y-specific STRs. *J. F. S.*, 48, 1, 2003.
- Malaspina P., Persichetti F., Novelletto A., Iodice C. *et al.*, The human Y chromosome shows a low level of DNA polymorphism. *Ann. Hum. Genet.*, 1990, 54, 297-305.
- Mannucci A., Sullivan K. M., Ivanov P. L. & Gill P., Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. *Int. J. Leg. Med.*, 1994, 106, 190-193.
- Roewer L., Arnemann J., Spurr N. K., Grzeschik K. H. & Epplen J. T., Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum. Genet.*, 1992, 89, 389-394.
- Rosser Z. H., Zerjal T., Hurles M. E., Adojaan M. *et al.*, Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am. J. Hum. Genet.*, 2000, 67, 1526-1543.
- Semino O., Passarino G., Oefner P. J., Lin A. A. *et al.*, The genetic legacy of paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science*, 2000, 290, 1155-1159.
- Spurdle A. B. & Jenkins T., The Y chromosome as a tool for studying human evolution. *Curr. Opin. Genet. Dev.*, 1992, 2, 487-491.
- Underhill P. A., Passarino G., Lin A. A., Shen P. *et al.*, The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann. Hum. Genet.*, 2001, 65, 43-62.
- Zerjal T., Wells R. S., Yuldashova N., Ruzibakiev R. & Tyler-Smith C., A genetic landscape reshaped by recent events: Y-chromosomal insights into Central Asia. *Am. J. Hum. Genet.*, 2002, 71, 466-482.