



par Bertrand JORDAN

*L'OPA des STS
a-t-elle fait long feu ?*

La proposition faite à la fin de l'année dernière par Olson *et al.* [1] de systématiser le balisage du génome à l'aide d'une nouvelle entité, les STS, a déjà suscité beaucoup de discussions. Voyons d'abord en quoi consiste cette proposition. Il s'agit en fait de déterminer la séquence de l'ADN en un grand nombre de points du génome pour définir en chacun d'eux un STS *sequence tagged site*, (littéralement : « site étiqueté par sa séquence ») caractérisé par un couple d'oligonucléotides permettant de produire par amplification PCR un segment d'ADN long de quelques centaines de bases et spécifique de l'endroit considéré. L'ensemble des STS constituerait ainsi un maillage de balises utilisables comme repères pour la carte physique (et éventuellement pour la carte génétique); la nouveauté principale de cette proposition est la définition des balises par PCR. Du coup, en effet, le STS est un élément défini une fois pour toutes, indépendamment, par exemple, du fait que la région en question ait été clonée dans un phage, un cosmide ou un YAC : il restera valable si, un jour, la zone correspondante du génome est clonée dans un nouveau vecteur encore inconnu aujourd'hui. Le STS est, de plus, par principe aisément utilisable, puisqu'il suffit de connaître la séquence de ses deux oligonucléotides pour l'obtenir : quelques heures dans la machine à PCR avec un peu d'ADN génomique, et l'on a une sonde prête à l'emploi pour toute étude souhaitée. On remplace ainsi — et ce n'est pas le moins

de des arguments en faveur des STS — la laborieuse collecte des sondes auprès de nombreux laboratoires qui les ont isolées (ou même auprès de l'ATCC, *american type culture collection*) par la simple consultation d'une banque de données suivie d'une production locale rapide (sinon économique, voir plus bas). La définition des STS se prête enfin bien à l'intégration de données obtenues dans différents laboratoires, puisqu'une simple réaction PCR permet de savoir si tel ou tel YAC ou cosmide contient un STS donné. C'est en ce sens qu'Olson *et al.* présentaient les STS comme « un langage commun pour cartographier le génome » [1]. Où en est-on donc, une bonne année plus tard, dans l'utilisation de ce système apparemment parfait ? L'OPA des STS n'a que partiellement réussi car leur généralisation s'est heurtée à deux difficultés, l'une basement matérielle et l'autre, en quelque sorte, de principe ; et pourtant, sous une forme un peu modifiée, ils sont en train d'envahir le marché. Voyons donc ces difficultés, et cet apparent paradoxe. La difficulté matérielle se situe tout simplement au niveau du prix de revient des oligonucléotides. Certes, de nombreux laboratoires sont maintenant équipés de synthétiseurs d'oligonucléotides, et, pour les autres, il suffit de consulter les publicités de *Science* ou de *Nature* pour voir que les fournisseurs se livrent à une concurrence acharnée à coup de rabais et de promesses mirobolantes sur les délais et le contrôle de qualité de ces réactifs. Il n'en reste pas moins que

la synthèse ou l'achat d'un couple d'oligonucléotides pour reconstituer un STS revient au moins à un millier de francs, et que la mise en œuvre dans une équipe de recherche d'une centaine de STS (il faut bien cela si l'on étudie la carte d'une région un peu étendue...) entraîne par conséquent un coût non négligeable. Comme, de plus, la synthèse avec les machines actuelles fournit au moins cent, sinon mille fois plus de l'oligonucléotide qu'on n'en utilisera en plusieurs réactions PCR... les équipes qui ont préparé les oligonucléotides pour un jeu de STS en font profiter leurs collègues, et on constate que l'envoi des sondes par la poste a bien été supprimé, mais qu'il est remplacé par l'envoi des oligonucléotides ! Ce qui, évidemment, enlève beaucoup de son intérêt au système... L'universalité de ces réactifs n'est d'ailleurs pas totalement évidente dans la pratique, car elle est limitée par la reproductibilité des conditions de PCR d'un laboratoire à un autre ; or celle-ci n'est pas totale, compte tenu en particulier des performances des différentes machines à PCR : les profils thermiques produits par certaines d'entre elles n'ont parfois qu'un rapport assez lointain avec ce qui a été affiché. On lira avec intérêt, de ce point de vue, une petite étude récemment parue dans *Trends in Genetics*, qui porte sur une douzaine de machines du marché et est assez cruelle pour certaines d'entre elles [2]. Qu'on ne s'y trompe pas, je ne suis pas en train d'écrire un réquisitoire anti-STs : je mentionne seulement certaines diffi-

cultés apparues lors de la mise en œuvre de ce qui reste un concept innovant et très utile.

L'autre difficulté, déjà mentionnée plus haut, touche au principe : les STS sont théoriquement excellents pour la cartographie... physique (comme l'indique d'ailleurs le titre de l'article d'Olson *et al.* [1]), mais ils ne sont pas, *a priori*, conçus pour effectuer la jonction entre cette approche et la cartographie génétique. Or cette jonction est primordiale, et l'histoire récente, qu'il s'agisse des difficultés du travail sur la chorée de Huntington [3] ou des progrès récents sur le syndrome de l'X fragile [4, 5], montre que, loin d'avoir détrôné la cartographie génétique, l'étude à grande échelle de l'ADN par les techniques les plus modernes a le plus besoin de balises réparties toutes les quelques mégabases, balises que seule peut fournir et ordonner à l'heure actuelle l'analyse génétique. Quitte à établir un maillage de STS tous les 100 kilobases, comme cela a été proposé, il vaudrait mieux qu'au moins une partie de ces repères fût utilisable pour une analyse génétique, c'est-à-dire permette de révéler un polymorphisme. Cet aspect ne semble pas du tout pris en compte dans la proposition des STS, du moins telle qu'elle a été initialement formulée [1].

La question, en fait, est en train de se résoudre toute seule, et d'une façon inattendue. L'apparition des microsattellites [6] dont j'ai parlé récemment en est la cause. Rappelons très rapidement qu'il s'agit de séquences particulières, du type (CA)*n* pour une des plus fréquentes, très répandues dans le génome et qui présentent un très grand polymorphisme (variation de *n*). Elles fournissent donc des repères quasi parfaits pour l'étude génétique ; leur exploitation se fait par amplification PCR à l'aide d'oligonucléotides reconnaissant les séquences en copie unique qui les encadrent suivie d'analyse de leur longueur (donc de *n*) sur un gel d'acrylamide très résolutif. La beauté de la chose, et le paradoxe dans le cadre de cette chronique, est que l'on engendre ainsi à

chaque fois un « super-STS » : c'est bien un STS, puisque la seule connaissance de la séquence des oligonucléotides (et des conditions de PCR à utiliser) permet de produire ce segment du génome ; mais c'est, bien sûr, un STS « haut de gamme » puisqu'il est en même temps un repère polymorphique pour l'analyse génétique.

La morale de cette histoire est qu'une fois de plus l'évolution des connaissances et des techniques a changé la façon dont se présente un problème *a priori* organisationnel et politique. Mais la « politique » ne perd pas ses droits puisque, même si de nombreux « super-STS » sont actuellement produits, trop peu sont mis à la disposition de la communauté, au dire des responsables des banques de données : la compétition parfois féroce sur certaines régions « chaudes » du génome humain est responsable de ce regrettable état de fait. Les responsables publics ou privés de la répartition des crédits considérables distribués pour ce type de recherche devront trouver des formules adaptées (et un peu contraignantes) pour que la collectivité profite effectivement des réactifs (les STS en l'occurrence) obtenus à l'aide de ces financements ■

Bertrand Jordan

Directeur de recherche au Cnrs, responsable du groupe génétique moléculaire humaine, CIML, Inserm/Cnrs, case 906, 13288 Marseille Cedex 9, France.

RÉFÉRENCES

1. Olson M, Hood L, Cantor C, Botstein D. A common language for physical mapping of the human genome. *Science* 1989 ; 245 : 1434-5.
2. Hoelzei R. The trouble with PCR machines. *Trends Genet* 1990 ; 6 : 237-8.
3. Roberts L. Huntington's gene : so near, yet so far. *Science* 1990 ; 247 : 624-7.
4. Brown WT. The fragile X : progress toward solving the puzzle. *Am J Hum Genet* 1990 ; 47 : 175-80.
5. Jordan BR. Fragile X-linked mental retardation and the difficulties of reverse genetics. *Bioessays* 1991 (sous presse).
6. Tautz D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 1989 ; 17 : 6463-71.