

## Les métiers du séquençage à grande échelle

PAR ÉLISABETH DE LAVERGNE



Sur la plateforme du CNG, Doris Lechner et Yannis Duffourd vérifient les paramètres et spécifications de la nouvelle machine de séquençage

Sur des équipements exceptionnels, les spécialistes du séquençage à grande échelle de l'Institut de génomique décryptent les génomes d'organismes vivants. Au Centre national de séquençage, techniciens, managers et bio-informaticiens séquent les génomes encore inconnus de plantes ou d'animaux. Au Centre national de génotypage, ils séquent et comparent les génomes de personnes saines et de personnes malades.

Deux bâtiments datant des années 1980 : Genopole, Genoscope situés à Évry, dans l'Essonne, une ville qui accueille de nombreux centres de recherche en biologie génétique. Le Centre national de séquençage (CNS) et le Centre national de génotypage (CNG) constituent l'Institut de génomique du CEA. Chacun a sa spécialité, tous les deux utilisent les technologies de séquençage à haut débit, ou encore à grande échelle. Des spécialistes du séquençage se relaient autour de deux plateformes de machines uniques en ►►

►► France : au CNS, ils décryptent les génomes encore inconnus de microbes, animaux ou plantes pour la recherche fondamentale et d'éventuelles valorisations en médecine, agriculture ou dépollution de l'environnement ; au CNG, ils comparent de grands nombres de génomes humains sains et malades, afin d'identifier les zones mutées susceptibles d'induire des maladies.

« Au CNS, nous disposons de quatre technologies pour un parc de 25 machines qui permettent de répondre aux différents besoins de séquençage de la biodiversité, explique Patrick Wincker, responsable du laboratoire de séquençage. Des machines qui, sans cesse, sont remises à niveau, améliorées ou remplacées. Les moyens techniques conditionnent notre positionnement à l'international. » Une évolution à grande vitesse des technologies qui exige beaucoup des équipes : « Tous les trois mois environ, nous mettons en place un protocole plus performant, plus rapide ou plus sûr. Ces changements impactent tous les spécialistes qui travaillent autour du séquençage : les cinquante techniciens supérieurs et les deux cadres de la plateforme, les deux cadres qui assurent le développement des machines, les douze informaticiens système et les vingt bio-informaticiens. » Un effort qui permet au CNS de se classer parmi les dix

premiers centres mondiaux, même si les centres américains ou chinois disposent de près d'une centaine de machines et que de nouveaux centres se créent partout dans le monde. Aujourd'hui, le centre travaille en partenariat, pour plus de la moitié de ses projets, avec des laboratoires des universités ou du CNRS qui ont besoin de moyens de séquençage. Les autres projets de recherche fondamentale sont propres au centre. Le plus médiatisé : celui du séquençage des génomes de micro-organismes marins récoltés autour du monde par l'expédition Tara Océans. Le CNS travaille aussi au décryptage des génomes de la truffe noire du Périgord – un résultat tout récent – des clémentines, des bananes, du café, du colza, de la truite... ou encore du virus H1N1, en lien avec le programme transverse du CEA Technologies pour la santé. Dernier exemple, le projet Cloaca maxima de séquençage du génome des bactéries des stations d'épuration, afin de comprendre et d'optimiser les mécanismes de dégradation et de proposer au marché des procédés innovants de "chimie verte" de dépollution.

À quelques mètres du CNS, le CNG est équipé différemment. Ici, le parc de machines est uniforme, car elles ne séquentent qu'un seul et même génome : celui de l'homme. « Notre mission est de comparer

**S**ous hotte, elle manipule flacons et pipettes. Dans le séquenceur, elle dépose une lame recouverte de minuscules billes portant de l'ADN. Avec son ingénieure, elle discute du nouveau protocole qu'elle évalue et rédige. Bientôt, elle formera les équipes. Caroline Menguy, technicienne supérieure du développement, référente technique de la plateforme de séquençage, semble savoir tout faire : « C'est parce que je travaille ici depuis dix ans. J'ai occupé tous les postes de travail et je suis capable de remplacer quelqu'un au pied levé. » Caroline Menguy est l'une des plus anciennes, dans une équipe jeune et plutôt féminine. Autour d'elle, les techniciennes évoluent au sein de salles vitrées. Posés en ligne sur des paillasses, les séquenceurs n'impressionnent pas beaucoup : ils ressemblent à de grosses boîtes en plastique. Pourtant, ici, l'organisation est industrielle et les machines fonctionnent 24 heures sur 24, week-end compris. « Le travail sur la plateforme n'est pas facile, soumis à deux impératifs contradictoires : d'une part, rester concentré sur des gestes répétitifs, très contraints, et, d'autre part, modifier souvent nos habitudes pour nous adapter à de nouveaux protocoles. Mes missions principales sont aujourd'hui de tester



**TECHNICIENNE SUPÉRIEURE AU CNS**

“ Il s'agit toujours d'apprendre ”

Dans la hotte de nébulisation, **CAROLINE MENGUY** prépare l'échantillon d'ADN pour la fragmentation

**Caroline Menguy,**

technicienne supérieure du développement

Bac S, parcours universitaire du DEUG SVT jusqu'à la maîtrise en biochimie à l'université Paris VII Denis Diderot

et d'adapter les nouveaux protocoles que nous fournissent les fabricants, de les rédiger et de former les équipes, de suivre leur mise en place. À l'automne, nous allons recevoir de toutes nouvelles machines. Maintenir la production tout en mettant en

place l'organisation et le nouveau fonctionnement de la plateforme sera un gros challenge. Pas possible de capitaliser sur mes acquis, il s'agit toujours d'apprendre ! Mais ce changement permanent est aussi la source de ma forte motivation. »

**J**ulie Poulain est l'une des deux responsables de production de la plateforme de séquençage. Chef des équipes, ingénieure, elle recueille les projets des chercheurs du CNS et des chercheurs partenaires, une fois qu'ils ont été évalués et acceptés par le conseil scientifique. Avec ses collaborateurs, elle définit la faisabilité des projets puis traduit leurs besoins en opérations techniques de séquençage. « Nous travaillons sur une centaine de projets en parallèle, qui durent de quelques mois à plusieurs années : il s'agit de gérer les priorités, d'organiser les travaux sur les différents types de séquenceurs et de planifier, avec l'équipe technique, l'occupation des machines. J'ai aussi la responsabilité de la mise en œuvre des protocoles et de la traçabilité des manip', en lien avec la responsable de l'assurance qualité. » Autre interface : la collaboration avec la trentaine d'informaticiens du CNS s'est renforcée depuis 2006, avec la mise en place des appareils de troisième génération : « Ils ont fait exploser les débits, les données, les besoins d'analyse, en même temps que les projets, prenant moins de temps, se sont multipliés. » Entre problèmes techniques à régler et préparation de l'avenir, Julie Poulain est aussi responsable des achats de la plateforme, en lien avec les acheteurs du centre CEA de Fontenay-aux-Roses, et manager de l'équipe : « Même si j'évolue



**CHEF DES ÉQUIPES AU CNS**

## “ Un métier d'interactions ”

De gauche à droite : Dominique Robert, JULIE POULAIN, Odette Beluche et Sandrine Lebled suivent l'enrichissement en billes portant des fragments d'ADN, dernière étape avant dépôt sur le séquenceur

dans un environnement technique et scientifique hyperpointu, je crois que ce que je trouve le plus difficile à réussir, c'est le management, les relations humaines. C'est exiger beaucoup des équipes, tout en maintenant leur cohésion et leur motivation qui sont la force du CNS. »

### Julie Poulain, ingénieure

BAC Science et technique de laboratoire – BTS analyse biologique – parcours universitaire de la licence de biologie à la maîtrise de biologie cellulaire et physiologie, mention génétique moléculaire et cellulaire, option microbiologie à l'Université Paris XI, Orsay – DEA d'écologie microbienne appliquée aux sols et milieux aquatiques à l'Université Claude Bernard de Lyon

les génomes de milliers de personnes atteintes d'une maladie à ceux de milliers de personnes saines, raconte Jorg Hager, responsable de projet de génétique humaine. Chaque génome s'écarte plus ou moins de la normale, mais quand ceux des malades montrent des variations similaires, nous suspectons des mutations qui jouent un rôle dans la maladie. Il nous arrive parfois de compléter ces travaux par le séquençage de modèles animaux de maladies humaines. Valider ces mutations comme favorisant cette maladie permet d'en comprendre les causes, de développer des outils diagnostiques et des approches thérapeutiques. » Ainsi, le CNG a travaillé sur l'asthme, l'hypertension ou le diabète. Il est aujourd'hui très actif sur le cancer. En 2010, sur les machines de deuxième génération, le séquençage du génome humain entier a duré trois semaines. Le premier séquençage dans les années 1990 a demandé six ans ! « Nous avons besoin d'une informatique extrêmement puissante. Les quantités de données et les besoins de stockage sont énormes. Dans chaque génome, 600 000 à 1 million de variations sont analysées. » Comme au CNS, les profils des douze techniciens, cadres et bio-informaticiens de la plateforme de séquençage du CNG, se révèlent très pointus. Le besoin de formation est permanent. « Les personnels recrutés n'ont jamais approché de telles machines, explique Doris Lechner. Ils mettent

plusieurs mois à être opérationnels, puis sont formés au fil des évolutions des machines et des protocoles. Pour ceux qui aiment apprendre et remettre en question leurs acquis, le CNG est passionnant ! ■

### Une histoire récente

C'est le programme télévisé Téléthon, né en 1987 autour de la génétique des maladies rares, qui a financé le Généthon, premier laboratoire en génétique implanté à Évry. Vingt ans plus tard, le bioparc d'Évry compte 20 laboratoires de recherche publique et 69 entreprises. Le Centre national de séquençage (CNS) et le Centre national de génotypage (CNG) ont été créés en 1997 pour dix ans, par le secrétaire d'État à la Recherche François d'Aubert. Au bout de cette période, en mai 2007, ils ont intégré le CEA au sein de l'Institut de génomique.

### À savoir

Le génome, les molécules d'ADN ou chromosomes renferment, sous forme chimique, l'ensemble des instructions ou gènes nécessaires au fonctionnement d'une cellule ou d'un organisme. Les gènes sont eux-mêmes constitués de nucléotides. Dans les nucléotides : des bases, qui se succèdent sous quatre formes (A, C, T et G). Le séquençage est la détermination de l'ordre de succession de ces bases. Les chromosomes humains sont longs de plusieurs dizaines à plusieurs centaines de millions de nucléotides. Chaque manipulation de séquençage ne livre qu'une séquence de 100 à 1000 nucléotides. Il faut donc en réaliser un grand nombre, en reprenant le séquençage des bordures pour raccorder les séquences les unes aux autres. La séquence complète d'un chromosome humain demande ainsi des dizaines de millions d'analyses.

**A**u milieu de son équipe de techniciens supérieurs, au coin d'une rangée de machines et son cahier à la main, Doris Lechner fait le point hebdomadaire sur l'avancement des travaux de séquençage de la plateforme. Responsable de la production, elle est un peu l'équivalente de Julie Poulain, mais côté CNG. Ici, les séquençages concernent presque tous le génome humain et se font sur des machines identiques, mais ils nécessitent le même talent de chef d'orchestre. Doris Lechner réagit de même au plus vite devant les pannes, sait se réorganiser face aux changements de programme, et voit remettre en question ses savoir-faire avec l'arrivée des nouveaux séquenceurs de deuxième génération. « Maintenant, nous avons un nouvel atout pour attirer les meilleurs. C'est aussi très gratifiant de travailler dans un centre qui garde sa position à la pointe de la technologie et de la recherche. En Europe, seuls le Royaume-Uni et la France sont capables de



### RESPONSABLE CHAÎNE DE PRODUCTION

“ Une production stable et fiable ”

**DORIS LECHNER** (assise) fait le point sur l'avancement des opérations de séquençage avec Florence Jobard et Damien Delafoy

**Doris Lechner,**

responsable chaîne de production

BAC BBZ à Miesbach/Bavière – Étude de professorat à Technische Universität München, Munich – étude de biotechnologie à BEUTH University of Applied Science, Berlin

rester techniquement dans la course. Un tel équipement nous donne des responsabilités fortes : j'ai calculé qu'un technicien manipule chaque jour dans ses mains l'équivalent du prix d'une voiture. » Doris Lechner n'est pas chercheuse, mais côtoie les scientifiques et les médecins partenaires. « Notre objectif final, ce sont leurs publications et les débouchés pour la santé. Les enjeux sociaux et

économiques sont énormes. Ainsi notre programme sur les tumeurs cancéreuses, qui vise à comparer l'ADN de la même personne, prélevé dans les tissus sains et dans les tumeurs. Nous cherchons à trouver comment le cancer modifie le patrimoine génétique, afin de se développer sans être éliminé par le système immunitaire. À la clef, peut-être, un nouveau traitement ? »

**I**l reçoit toutes les données issues des séquençages. Yannis Duffourd est l'un des trois bio-informaticiens du CNG. Tout le séquençage de la plateforme passe par lui : des téraoctets de données chaque mois. « Je reçois des millions de séquences à aligner, c'est-à-dire à positionner au bon endroit sur le génome. Cela est réalisé automatiquement par un logiciel que nous a fourni le fabricant, mais je filtre au préalable les données aberrantes des données exploitables. Parfois, il est nécessaire d'utiliser des logiciels différents pour optimiser l'alignement. Au final, en général, j'arrive à aligner 80 % des séquences. Pour un génome entier, cela me prend plusieurs jours. » Une fois cette première étape terminée, il s'agit de comparer les génomes sains et malades entre eux : « Plusieurs jours de calcul à nouveau, cette fois grâce à des logiciels que nous avons développés en interne. Au final, nous obtenons les positions de mutations ou de délétions, insertions, duplications, toutes sortes d'anomalies génétiques. Certaines n'ont aucun effet, il s'agit d'identifier celles qui pourraient être responsables de la maladie. » Yannis Duffourd pilote en parallèle quinze supercalculateurs, pour cinq projets. « Je suis biologiste de formation, avec un diplôme de master en informatique. Le rôle



### BIO-INFORMATICIEN

“ Le rôle de l'informatique ne cesse de se renforcer ”

**YANNIS DUFFOURD** dans son bureau : les supercalculateurs n'empêchent pas les raisonnements sur tableau blanc

des bio-informaticiens ne cesse de se renforcer, avec des problématiques, des outils de travail exceptionnels et des enjeux vraiment intéressants. C'est un poste

d'envergure ! Ce qui m'intéresserait particulièrement, ce serait de pouvoir bénéficier des moyens de calcul haute performance du CEA. »