

# CENTRE de SÉQUENÇAGE du GÉNOME

Rapport d'étape 18 Janv. 1996

J. M. ÉGLY (Ancl. CNS)

D'après des différents entretiens que j'ai eus, il semblerait qu'il se dégage une sorte de consensus général sur l'utilité de créer un centre de grand séquençage. Il faut admettre que les esprits ont évolué à ce propos et que les opinions ne sont plus aussi tranchées que par le passé.

Les arguments des opposants au projet, dont certains peu scientifiques étaient d'ordres divers:

Tout d'abord, ils évoquaient le coût financier d'un projet qui dans la période actuelle, pourrait se faire au détriment d'autres actions, voire du fonctionnement "normal" de la recherche dans le domaine des Sciences de la Vie. Ils ajoutaient qu'il y avait déjà bien assez de "gènes à étudier" et que ceux-ci pouvaient occuper bon nombre de chercheurs de la communauté scientifique avant d'aller rechercher d'autres séquences de façon systématique. Ils suggéraient de s'interroger sur l'après génome et d'essayer de mettre en oeuvre une stratégie visant à étudier les fonctions des gènes. Ils laissaient percevoir également un scepticisme certain quant au succès d'une telle opération vu les déboires récents encore présents dans bien des esprits. Sans vouloir entrer dans un débat concernant la mission qu'aurait dû ou non remplir le GREG, il faut remarquer que les quelques résultats obtenus ainsi que ceux des programmes Généthon, ont eu le mérite de situer le problème. Il persistait également chez beaucoup, comme un sentiment de frustration, celui d'avoir en quelque sorte "loupé le coche" en ne nous intéressant pas au cDNA, dès le début des années 90 comme il l'avait été clairement suggéré dans un précédent rapport. Réflexions faites, ces mêmes opposants acceptaient l'idée d'un tel centre à condition que le projet soit bien défini (que séquençer? à quel rythme? et en combien de temps?). Je crois que nous sommes tous d'accord, sur ce point. Il est intéressant de noter que ce souhait était expressément formulé par les responsables du domaine public.

Un autre argument en défaveur de la création d'un tel centre prenait en compte notre retard réel dans le domaine du séquençage proprement dit. A cela, on peut répondre qu'en dehors des projets américain (St Louis, Missouri) et anglais (Sanger Center à Cambridge), le projet allemand entre Berlin et Heidelberg, orchestré par le Prof Lehrach et les Drs Poutska et Rosenthal d'une part, et le projet japonais (sous forme de deux à trois actions thématiques patronnées par le MITI) sont en cours d'élaboration et ne devraient démarrer que

vers la fin du premier semestre 96. De plus le séquençage systématique du génome humain vient tout juste de commencer. Je ne fais pas ici référence aux séquences de différents EST (Expressed Sequence Tags) effectués et déposés dans différentes banques de données. Par conséquent, nous sommes encore dans les temps si nous voulons mettre en oeuvre un projet de centre de séquençage.

## I. Les atouts dont nous disposons.

Après avoir brièvement énuméré les diverses réflexions et commentaires ayant entouré l'élaboration de ce document, il est bon à présent de mentionner les points forts nous permettant d'envisager favorablement la création d'un tel centre.

1) Une expérience propre et des acquis en cartographie du génome principalement issus du CEPH et du Généthon par ailleurs largement référencés.

2) Une communauté scientifique qui a fait ses preuves comme en témoigne notre participation non seulement aux grands programmes de séquençage des génomes de levure, du nématode et de l'homme, mais aussi à tous les grands congrès en tant que conférenciers invités. Il faut également noter l'excellence des équipes travaillant en génétique humaine et leur découverte de plusieurs gènes responsables de diverses maladies génétiques.

3) L'expérience des autres. Il est vrai que nous avons pris un certain retard; il est également vrai que personne en France n'a l'expérience du très grand séquençage. Notons que la situation est identique pour ceux qui comme nous décident de démarrer ce type de projets. Mais ne se pourrait-il pas que ce retard, additionné de la compétence de notre communauté scientifique, puisse être en quelque sorte un atout, dans la mesure où il nous permettrait d'éviter bien des erreurs. Il ne semble pas au travers des quelques discussions que j'ai pu avoir, et des articles que j'ai pu lire dans la presse grand public ou plus spécialisée, que les deux principaux centres que sont ceux de Cambridge et de St Louis, financés par des fonds gouvernementaux, une fondation (Wellcome Trust) et par des associations caritatives, hésitent à nous faire bénéficier de leur savoir-faire. Formellement ils sont prêts à nous accueillir pour une période de formation. Il existe une sorte de solidarité entre les chercheurs travaillant sur le séquençage du génome humain. Le travail à fournir s'avérant tellement vaste et complexe que toute bonne volonté pour se partager la tâche ne peut qu'être appréciée. La solution des problèmes rencontrés au niveau du génome humain ne pourra qu'être bénéfique pour le séquençage d'autres génomes.

## II Utilité d'un centre de Très Grand Séquençage

A ce moment de l'exposé, il nous faut mettre en exergue les avantages qu'offrirait en France la mise en place d'un centre de séquençage efficace, compétent et flexible.

### 1) *Un centre à disposition;*

Un centre capable de séquencer de grandes régions de plusieurs dizaines de kb permettrait selon les besoins dictés par des contingences scientifiques et/ou économiques, la mise en oeuvre et la réalisation de projets d'une certaine ambition dans un temps minimum.

### 2) *Un centre de référence;*

Il possédera et maintiendra un savoir faire de haut niveau. Il permettra l'amélioration et le développement de technologies plus rapides, plus efficaces et plus performantes dans la préparation des échantillons à séquencer. Il pourra, seul, en concertation, ou éventuellement sur appel d'offres, développer une informatique et/ou mener une recherche à partir des banques de données et de ses propres résultats. Ce sera un centre d'information. Ainsi, chacun, à condition qu'il fasse partie du réseau (ce sont là des modalités à définir), pourra travailler à partir et avec le centre. Il bénéficiera de ses acquis dans la recherche de séquences similaires et de ses outils.

### 3) *Un centre éducatif;*

Il permettra de former techniciens, ingénieurs et chercheurs, à la fois aux techniques de séquençage et aux problèmes d'informatique qui lui sont associés, non seulement pour lui-même mais également pour d'autres besoins. Sur ce dernier point cela peut se faire également en concertation avec d'autres structures "high tech" nécessaires à la bonne marche d'un tel établissement. Devant les enjeux économiques que le séquençage des génomes laissent entrevoir, il y aurait besoin de connaître et même de provoquer la législation et de ne pas être uniquement à l'écoute d'autrui. C'est pourquoi ce centre stimulera l'émergence d'une sorte de "culture scientifico-juridique du brevet" et par là même de cadres "scientífico-juridiques", une catégorie de personnel fort rare et cependant nécessaire dans l'hexagone.

Il faut également voir ce centre comme un pôle d'éducation d'une partie du monde industriel, concernée par certains organismes moteurs de son activité; ceux impliqués dans le recyclage ou la dépollution de l'eau par exemple. Un organisme comme l'IFREMER ou des sociétés travaillant dans le secteur de la distribution d'eau ne devraient pas être insensibles à l'activité que pourrait

développer un tel centre, comme la connaissance des génomes de ces fameuses algues qui perturbent l'équilibre écologique du littoral par exemple.

### 4) *Un plateau technique;*

Cela dépend tout d'abord du mode de fonctionnement de ce centre. L'idéal serait qu'une partie de son activité se fasse sous forme de service ou de travail en collaboration. Cependant, il faut être réaliste. Il est illusoire de songer à accueillir des équipes qui viendraient uniquement pour faire leur propre séquençage et qui travailleraient de façon plus ou moins autonome en utilisant elles-mêmes les appareils de séquençage. On peut cependant imaginer qu'il peut y avoir "aide" dans la préparation des échantillons. Le centre sera plus intéressant à utiliser en tant que prestataire de service pour la partie séquençage en raison de la sophistication de l'appareillage; quant à l'analyse et à l'exploitation des résultats, elles pourront se faire dans le centre avec l'assistance technique dont il dispose, et en collaboration avec le personnel demandeur. Ce genre de travail pourra se faire après concertation avec les responsables du centre.

### 5) *Un centre analytique et d'exploitation;*

L'aspect qui m'est apparu peut-être le plus prometteur, concerne principalement le secteur aval (**bioinformatique**) qui doit gérer l'accumulation, la classification et l'enregistrement des données issues de ce centre et de ceux qui lui sont associés au travers de divers réseaux. Cette activité pourra se faire en collaboration avec d'autres groupes selon des modalités précises à définir. Ce centre de ressources devra en outre accumuler, comparer et exploiter les résultats à partir de la masse de données qu'il possède et de celles qui lui arriveront de ses propres laboratoires ou d'autres sources. Cette activité d'exploitation et de prospective permettra d'aborder l'étude de la structure et des fonctions des gènes en question, grâce à la compétence de spécialistes en génétique, en biologie moléculaire et structurale, et en immunologie. L'étude de banques soustractives ainsi que la comparaison de plusieurs gènes issus de plusieurs espèces, permettra de mettre en lumière des régions antigéniques. Ces informations pourraient être analysées expérimentalement dans le cadre d'actions coordonnées que l'on pourrait mettre en place. On voit, là encore, l'intérêt de **situer ce centre dans un programme beaucoup plus vaste.**

### 6) *Un centre européen;*

Ce centre peut restimuler une communauté scientifique européenne qui semble quelque peu "en panne" de grands objectifs communs. Il mobilisera une grande partie de la communauté scientifique de façon plus ou moins directe en fonction des sujets grâce aux projets et aux renseignements fournis par le séquençage. L'accès à l'information, via l'ordinateur utilisé quasiment par tous ceux travaillant en génétique, biologie moléculaire et structurale, biologie du

développement ou immunologie entre autres, ne se réduira plus à chercher ou comparer des séquences; il permettra également de nouer des contacts plus facilement et plus rapidement à partir de données relativement brutes, de poser des questions, de compléter ou d'initier des collaborations entre les chercheurs. Au lieu (et en plus) de se rencontrer dans des meetings ou au travers de publications, on pourra établir des relations plus en amont, en cours "d'études". Ceci est à la portée de tous, et est, il faut en convenir, l'apanage du jeune chercheur qui se sentira encore plus concerné car il manipule l'ordinateur au quotidien.

Déjà, de nombreuses collaborations existent entre des équipes européennes et françaises; ce genre de projet ne peut que les renforcer. Divers programmes de séquençage sont mis en place dans certains états de la communauté européenne. **L'idée d'un projet de grand séquençage est accueillie avec enthousiasme par nos partenaires anglais et allemand concernés par le sujet.** Il existe un programme allemand prêt à démarrer. Il serait judicieux d'étudier comment d'une part on pourrait mener une coopération avec ces derniers, qui sont très intéressés par ce projet et d'autre part officialiser cette action au niveau des ministères ou organismes concernés. Il ne faudra pas non plus oublier les pays du nord de l'Europe comme la Hollande qui a un fort potentiel de chercheurs travaillant sur les maladies génétiques. Un tel centre pourrait être un des noyaux durs d'un vaste programme très bien coordonné au niveau européen. Il serait bon de mener dès à présent des négociations au niveau de Bruxelles.

#### 7) *Un centre dans un contexte scientifique plus large.*

Il est acquis que ce projet présente de nombreux avantages. Il ressort de tous les entretiens que j'ai pu avoir d'une part et du concept même d'un tel centre d'autre part, qu'il faut l'intégrer dès le départ dans un programme beaucoup plus vaste. En effet la motivation ne peut et ne doit pas être de faire du séquençage sans autre extrapolation. On peut imaginer par exemple qu'un des sujets propres du centre concerne l'étude d'un génome bactérien, viral ou d'une région du génome humain. Les travaux effectués aussi bien en amont pour la préparation des échantillons, qu'en aval pour l'exploitation des résultats trouveraient un début de valorisation si il y avait un retour; des travaux complémentaires pourraient ainsi se faire dans le cadre d'un programme beaucoup plus vaste concernant l'étude d'un génome bien particulier, de maladies génétiques, ou de l'après gène comme cela est également demandé par les leaders de notre communauté scientifique. Un centre comme celui que l'on voudrait mettre en place, s'intégrerait très bien dans un projet semblable à celui proposé par le Prof Lehrach au ministère allemand de la recherche.

### III Le grand séquençage dans les autres pays:

Il est assez difficile de percevoir ce qui se fait exactement ailleurs. Les déclarations intempestives des différents leaders dans le domaine, poussés par leurs autorités de tutelle et de nombreuses associations, préoccupées surtout par le séquençage du génome humain, font apparaître de nombreux projets dont beaucoup ne sont guère réalistes. Ils ont néanmoins le mérite de manifester la volonté d'une participation à l'aventure du séquençage du génome humain et de poser le problème de la connaissance du génome en général. Les arguments avancés ont pour but l'amélioration de la santé, de la production agricole, ou de l'environnement. **Tous tablent sur des retombées économiques dont personne n'est en mesure d'apprécier l'ampleur.** Il s'en suit que, réaliste ou non en ce qui concerne les retombées économiques ou scientifiques, un projet significatif dans le domaine du séquençage représente une masse de travail et un coût financier dont je ne sais si tous mesurent l'importance. Cependant, malgré les valeurs aussi disparates avancées pour le coût de telles actions (entre 0,2 et 3 dollars/base finie), tous les pays ont et développent des projets dont une partie concerne le génome humain.

Il semble même que les projets Séquençage du Génome Humain masquent en partie le développement d'autres projets de séquençage qui ont une visée peut-être plus rentable, du moins à court terme. Il faut entendre par là que d'autres génomes plus simples et plus faciles à séquencer du point de vue technique, verront le jour et seront lancés grâce à cette aventure génome humain. On pourra argumenter que le séquençage des génomes de levure, de nématode, de drosophile, de poisson zèbre, de poisson fugu, ou de souris pourront faire avancer plus rapidement celui de l'homme. Là encore il faut prendre la mesure du poids des **associations caritatives qui, de façon pas toujours très objective et réaliste, insistent pour que l'on s'intéresse au séquençage du génome humain.**

Ainsi, il semble évident que **l'on doive s'intéresser au génome humain.** Il apparaît que toutes les discussions et débats menés autour de ce type de projet, par les scientifiques, les économistes et les associations caritatives, ont l'avantage de donner en quelque sorte **une légitimité à la biologie**, comme il en a été de l'énergie atomique, ou de l'espace en leur temps. De ce point de vue c'est une bonne chose car les pouvoirs publics et le grand public commencent à **percevoir** beaucoup mieux **l'intérêt de la biologie et à apprécier les perspectives qu'elle offre.** Ce projet a également quelque chose de populaire. En effet, plus concret que les projets de thérapie génique dont le succès est loin d'être acquis, il peut

7  
alimenter plus aisément l'espoir. Rendons cependant hommage à ceux qui mènent des projets de thérapie génique, très complexes, mais qui cependant ont le mérite d'exister.

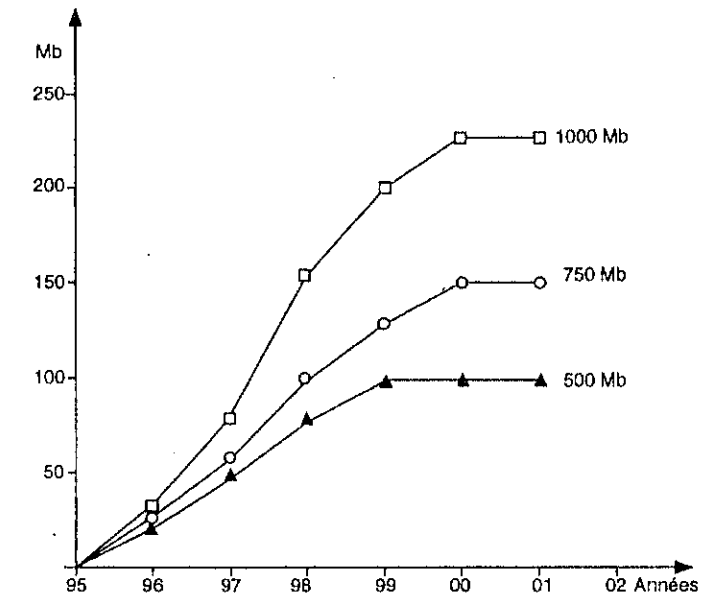
Cet engouement pour les problèmes de grand séquençage est peut-être à l'origine ou la conséquence de l'explosion de la biotechnologie que l'on perçoit aux USA. Les cabinets d'experts (source Wall Street Journal et pages économiques du Figaro) prévoient un rythme de croissance de l'ordre de 12% par an de ce secteur économique, au cours des dix prochaines années, passant d'un chiffre d'affaires de 10 milliards de dollars en 1996, à 33 milliards de dollars en 2006. Tous les secteurs en bénéficieraient: la thérapeutique humaine ainsi que le secteur diagnostique humain, auraient une croissance évaluée à plus de 13% par an. La biotechnologie agricole est beaucoup plus prometteuse avec une croissance moyenne annuelle de plus de 20%. Le chiffre d'affaires des industries concernées qui est actuellement de l'ordre de 300 millions de dollars/an, devrait être multiplié par 6 au cours des dix prochaines années. La conception de nouvelles variétés de plantes et d'animaux est envisagée grâce aux techniques de génie génétique.

Examinons brièvement la situation dans les pays concernés:

#### 1) Les USA:

Justifiant par là son titre de superpuissance, les USA, par l'intermédiaire du NIH (National Institute of Health) et du DOE (Department of Energy), consacrent d'ores et déjà pour les centres de séquençage, de cartographie et d'analyse du génome humain, plus de 100 MD/an. Le NIH finance environ une quinzaine de centres de moyenne importance pour une somme variant entre 3 et 5MD/an et par centre en fonction de la thématique étudiée, tandis que le DOE finance pour quelque 35 MD, des projets allant du séquençage à l'analyse fonctionnelle des gènes en question, en ne négligeant pas l'aspect biologie structurale. Le but affiché depuis quelques années est d'atteindre un rythme de l'ordre de 20 à 30 mégabases séquencées(Mb)/an. Ils estiment ainsi qu'en collaboration avec d'autres partenaires, le projet nématode de 100 Mb sera terminé en 1998. Devant les difficultés que laisse entrevoir le séquençage systématique du génome humain, ils ont décidé de financer par l'intermédiaire du NIH 3 "projets pilotes", capables de réaliser chacun le séquençage de 10 à 15 Mb/an; le minimum requis est 7,5 Mb/an. Des résultats de ces trois actions, dépendrait la stratégie future mise en place. Ils essaient dans le même temps de regrouper et d'identifier les équipes les plus compétentes qu'ils réuniraient en les subventionnant au travers de consortiums. Ils mettent à mettre en place des réseaux entre producteurs de séquences et détenteurs de données, optimisant au maximum l'exploitation des

7 Bis  
Fig 1 : Projets Sanger et St Louis



Après concertation, les responsables des centres de séquençage de St Louis et de Cambridge proposent de séquencer chacun 1/3 du génome humain, soit 1000 Mb. Cependant, ils n'ont obtenu un financement que pour effectuer 500 Mb. Les trois courbes représentent leurs objectifs de séquençage par an jusqu'en 2002.

résultats. Les autres génomes comme celui de la levure, du nématode et de la drosophile, font également l'objet d'études intensives.

### 2) La Grande Bretagne:

Les décisions prises en 1993 ont donné naissance au Sanger Center, le principal centre anglais qui rassemble quasiment tous les projets de séquençage. Sa mission consiste à participer à la cartographie, au séquençage et à l'analyse fonctionnelle de portions du génome humain, ainsi qu'à d'autres génomes modèles comme celui du nématode; ceci en étroite liaison avec les programmes américains. Il existe également deux autres structures au nord de l'Angleterre (dont une à Aberdeen) qui s'occupent plus spécifiquement de projets à connotation agricole (génomes du porc et d'Arabidopsis) fortement subventionnés par la CEE. Ce dernier centre effectue également une sorte de prestation de service pour quelques projets industriels de faible ampleur.

En 1995, après bilan, ce même centre établit un programme beaucoup plus ambitieux. En plus des projets levure (3 groupes de 14 personnes) et nématode (5 groupes de 12 personnes chacun), il met en route le projet génome humain.

Les projets levure et nématode étant en phase d'exploitation routinière (2,5 Mb et 8 Mb respectivement, ont été séquencés en 1995) et se terminant dans les trois ans, le centre met dès à présent ses efforts sur le génome humain (2,5 Mb séquencé sur divers chromosomes en 1995 avec un groupe de plus de 25 personnes pour le séquençage et une trentaine de personnes pour la préparation des échantillons et la cartographie des régions concernées). Le centre ambitionne de séquencer d'ici à l'an 2002, 1000 Mb de ce génome (150 Mb, pour fin 1998 selon la Figure 1), ce qui semble quelque peu irréaliste au dire même des experts. Le MRC et le Wellcome trust ont déjà alloué 70 Millions de livres au centre Sanger, pour les 7 prochaines années (10 M/an); les mêmes institutions leur ont garanti un complément de 2M par an (pour l'instant bloqué), supplémentaires si l'objectif de 500 Mb pour 2002 n'est pas atteint. (Figure 1, courbe 500Mb). Vu les résultats atteints en 1995 (2,5 Mb), il semble même difficile de réaliser pour 1997/98 le séquençage de 70 Mb. Le centre de St Louis aux USA, travaillant en très étroite collaboration avec le centre de Cambridge envisage une activité semblable, après avoir dépouillé les résultats des trois projets pilotes (voir plus haut). Le responsable du groupe séquençage du génome humain de Cambridge, D. Bentley, souligne l'effort à fournir sur la cartographie, la production de nouvelles banques à partir de chromosomes, ainsi que sur l'informatique. Une fois de plus, on constate que ce dernier secteur ne répond pas à la demande et qu'il faudrait renforcer ce pôle au niveau recherche et développement.

### 3) Le Japon:

Ce pays a été jusqu'à présent un observateur attentif de l'avancée de la technologie de séquençage et de la nature des projets qui se déroulaient au niveau international. Le projet "Frontière humaine" qu'il parraine financièrement en grande partie, a fortement contribué à sa prise de décision. De façon directe ou indirecte, il participe financièrement à de nombreux projets industriels qui se développent aux USA. Il est en train de bâtir divers projets au niveau académique, laissant à son ministère de l'industrie le MITI le soin du financement, sur au moins 4 sites, avec une vingtaine de personnes et une subvention globale de 10 M de dollars. Cette dispersion de sites ne soulève pas l'enthousiasme des scientifiques. Il est intéressant d'analyser la composition de leur groupe travaillant sur un projet archaebactérie: ce projet dont la partie séquençage démarre (3 Mb) en mars 1996, devrait durer 1 an maximum, demande dès le départ la participation de généticiens, informaticiens et biologistes structuraux, une façon de se projeter au plus vite sur "l'après gène".

### 4) Le Canada:

Des projets sur les génomes humains (10), ainsi que sur ceux de levure et drosophile (7 pour les deux sujets) sont en cours de réalisation dans le cadre de collaborations internationales. Ils sont, ainsi que 3 centres de données et 4 projets informatiques, subventionnés par le Conseil de la Recherche et le Ministère de l'Industrie, responsable de la recherche. Le Conseil de Recherches Médicales assure le développement scientifique; pour cela, il a reçu 22 M de dollars canadiens en 1995. Très vigilant sur ce qui touche à la propriété industrielle, il prévoit de mettre l'effort sur la cartographie à haute définition d'une partie du génome humain, ainsi que de celui de la souris. Ils ont des projets sérieux sur le séquençage d'organismes bactériens simples (*Sulfolobus*, *Solfata*...).

### 5) L'Allemagne:

En dehors d'une participation au même titre que de nombreux états de la communauté européenne, aux projets levure, nématode et drosophile en collaboration avec diverses équipes internationales, l'Allemagne met en place avec l'appui soutenu du ministre concerné, Mr Rutgers un vaste projet dont la thématique est la suivante "Identification et Explication de la structure, de la fonction et de la régulation des gènes humains et en particulier ceux revêtant une importance médicale". Une attention spéciale sera portée sur l'origine des maladies génétiques, le recensement des facteurs de risque, la découverte de substances actives, la mise au point de tests diagnostiques et d'essais de thérapie". Cette thématique n'exclut pas des travaux sur le poisson zèbre et la souris qui ne peuvent qu'être bénéfiques au projet génome humain. Il y aura des appels

d'offres qui solliciteront des équipes préalablement remarquées dans d'autres pays dont la France.

Dès le départ, ce pays marque sa volonté de s'intégrer dans un contexte international et surtout européen:

- désir de prolonger les consortiums opérationnels sur la levure et le nématode,
- désir de cadrer leur centre de ressource ou de séquençage dans le contexte européen,
- désir de s'ancrer dans le programme HUGO ,
- désir de stimuler une action au niveau de Bruxelles (CEE).

On ne parle pas de grand centre de séquençage, en dépit du fait que des groupes de faible ampleur sont en cours de formation. Il semblerait même que l'on privilégie une recherche plus verticale incluant l'après gène. Le séquençage proprement dit se ferait sur des zones restreintes et serait soumis à appel d'offres. Le montage d'un centre en France, serait plus que fortement apprécié par notre voisin allemand.

Ce centre sans murs, soutenu par les organismes de recherche, le secteur santé et des fonds industriels, s'articulera entre le Deutsche Krebs Forschung Zentrum (Berlin) et le Max Plank Institute (Heidelberg); son activité sera suivie par un comité propre au centre, par un comité scientifique international et à un niveau politique.

- Le comité scientifique aura pour mission, l'évaluation de la qualité, du choix et de l'orientation des programmes de recherche. Il devra de plus, assurer une veille technologique.

- Ce centre se dote d'un comité politique ou des représentants du ministère de la recherche cotoyeront des responsables scientifiques, des représentants du secteur industriel, médical et agricole ainsi que des représentants du monde associatif. Il tiendra compte de priorités financières (leur nature n'est pas précisée), pourra modifier certaines orientations et se préoccupera des problèmes éthiques particulièrement sensibles Outre-Rhin.

En 1995, la dotation était de 50 M de DM. Son fonctionnement est prévu pour une période de 8 ans. Le centre de ressource sera accessible aux chercheurs du secteur publique et privé, aux médecins soignants ainsi qu'à tous les laboratoires d'analyse. Il enregistrera toute idée et suggestion en provenance de ces utilisateurs.

Des conversations que nous avons pu avoir avec des responsables, il ressort qu'ils seraient prêts à tisser des liens étroits avec la France.

#### 6) La Suisse:

Consacrant 2,7% de son PIB à la recherche, le séquençage du génome ne fait pas partie de ses préoccupations. Ce secteur dont se décharge l'Etat, est laissé aux mains des gros groupes industriels qui y participent de façon plus ou moins directe, par l'intermédiaire de leurs filiales américaines.

Beaucoup de pays de la CEE participent aux divers projets "grand séquençage". Cela se fait dans le cadre de collaborations ou de divers appels d'offres. Une étude plus détaillée mériterait de passer en revue les divers pays. Il n'y a pas d'autres gros projets qui se dessinent en URSS ou en Chine. Il y a peut-être des projets à Taiwan.

#### 7) L'Inde:

Considéré comme un grand réservoir de maladies génétiques, ce pays a mis en place un groupe de travail en 1993, avec comme objectif la création d'un réseau de recherche clinique. Ce pays s'intéresserait dans un premier temps, aux hémoglobinopathies, aux thalassémies, à la dystrophie myotonique, à la maladie de Huntington et à tout gène associé aux problèmes de stérilité. De plus l'Inde désire créer une banque de données sur les maladies génétiques, s'intéresser aux projets concernant le génome du vibron cholérique, de l'amibe dysentérique et étudier les mutations de p53 dans la population. De plus amples informations manquent. Il faut noter qu'ils basent leurs projets sur une collaboration avec le CEPH suite à des contacts avec le Prof. Dausset et le Dr Beckmann (Genethon)

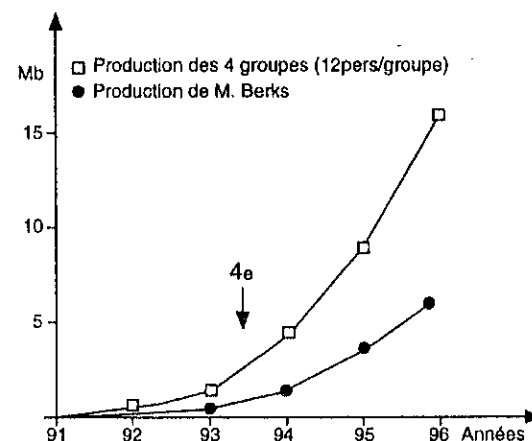
#### IV) Un exemple de grand séquençage: Le groupe de M. Berks.

Afin de bien fixer les idées, il m'a paru intéressant de décrire l'évolution du projet de Mary Berks un chef de projet au Sanger Center (Figure 2).

M. Berks a intégré le centre en 1991 en même temps que John Sulston, l'actuel directeur, ce qui signifie qu'elle a une expérience de 4 ans derrière elle; elle travaille principalement sur le séquençage du nématode (*C. elegans*). Post-Doc, elle anime une équipe de 12 personnes, dont elle-même. Compte tenu du contexte et de l'expérience acquise, son objectif pour 1995, était de 2 Mb séquencés; elle séquencera un peu plus de 3Mb (3,3 Mb) dont 0,2 de participation au projet génome humain. L'objectif total pour le secteur nématode était de 8 Mb; deux autres groupes ont pu séquencer 2Mb chacun et le quatrième (formé courant 1994) a produit 1,5Mb.



Fig 2 : Montée en Production  
du séquençage de C. Elegans



Ce graphique donné à titre d'exemple indique la montée en productivité du séquençage du nématode (*C. elegans*) au centre Sanger à Cambridge entre 1991 et 1996. Il est à noter la performance du groupe de M. Berks (le plus ancien des 4 groupes concernés par le projet) qui à lui seul projette de remplir plus du 1/3 de l'objectif fixé de 16 Mb pour 1996. On notera également que le 4e groupe formé n'est pas encore en pleine puissance de productivité.

L'objectif fixé pour 1995 a donc été atteint et dépassé. Chacun des groupes comprend 12 personnes, dont seulement 1/3 possède une expérience en biologie (niveau BTS). Quatre appareils de séquençage, effectuant 12-14 séquences par jour, sont en activité.

Mary Berks est confiante de pouvoir réaliser avec ses collègues, en 1996, le séquençage de 16 Mb (Figure 2), sans augmentation significative de personnel et de moyens (moins de 20% de coût supplémentaire).

Il est surtout intéressant de noter l'évolution de la productivité des différentes équipes (3,1 Mb, groupe formé fin 1992) et 1,5 Mb, groupe formé 1994), dont la seule différence est due à l'expérience acquise. M. Berks insiste sur ce point, et nous dit que débarrassée des problèmes de cartographie, grâce à une excellente préparation d'échantillons, et grâce à des "finishers" très expérimentés (personnes qui analysent la qualité des résultats de séquençage et qui redemandent et reprécisent les zones délicates), ce qui est le cas, le séquençage peut aller vite. Divers contrôles attestent quotidiennement de la qualité des résultats jugés "acceptable" si 70% des séquences sont bonnes et "very good" si on dépasse les 80%, ce qui est son cas. Elle va même jusqu'à dire que son personnel n'est pas surchargé (groupe de 12 personnes). La nouvelle génération de séquenceurs (ABI 377) devrait permettre une augmentation de production de données de plus de 25%. Son groupe a une taille optimale permettant efficacité et flexibilité. Il s'agit là d'un "module" type, constituant élémentaire d'un centre de grand séquençage.

Il faut noter que **18 mois** représentent une durée minimale difficilement compressible pour atteindre ce type de performance.

#### V) Le Centre de Très Grand Séquençage:

De l'avis des différentes personnes ayant une expérience dans le domaine, il apparaît que la décision de monter un centre de séquençage, semble être attendue par la communauté scientifique nationale et internationale, de même que par le public. Il en résulte que cette opération sera très observée et que nous nous devons de réussir.

##### 1) Un objectif réaliste et modeste de 30 Mb/an:

L'expérience dans le domaine grand séquençage n'existe pas en France. Par conséquent il faut monter le centre de toutes pièces en essayant autant que possible, de bénéficier de l'expérience existant ailleurs.

Pour être des partenaires crédibles et considérés, on se doit d'avoir une production minimale comprise entre 25 et 30 Mb séquencées/an (Figure 3). C'est

un objectif raisonnable que l'on se fixerait dès le départ et qui tient compte des réalités techniques et informatiques du moment. (Tableau 1, cas A et B; coefficient d'efficacité de 12%). On peut prétendre favorablement à un rendement supérieur grâce aux nouveaux séquenceurs et aux nouveaux logiciels qui seront prochainement accessibles (cas B; 20%); ainsi avec une structure de centre quasiment identique, il sera possible de nous rapprocher de l'**objectif haut de 50Mb**. Il n'est guère réaliste de prétendre séquencer plus, avant d'avoir évalué notre capacité à remplir l'objectif de production de 30 Mb/an dans un temps aussi court. Il sera toujours temps d'augmenter par la suite la capacité de production du centre et de fixer d'autres objectifs.

### 2) *Un seul centre, dans un premier temps:*

Le rapport de Mr F. Galibert, et le sentiment de plusieurs personnalités scientifiques, favorisaient la création de plusieurs centres de petite taille. Or, selon l'avis des scientifiques présents dans des centres de séquençage, il en résulterait une perte d'efficacité: on ne pourrait assurer une veille technologique correcte, en dehors du fait que l'on augmenterait les coûts de fonctionnement. Par exemple, le centre de St Louis tire sa force d'une confrontation permanente des différentes techniques mises au point par chaque groupe et de leur transfert dans tous les groupes; il en résulte dans ce cas, un gain de productivité. D'autre part, il n'y a pas suffisamment de compétences en France pour justifier une dispersion de centres.

Je pense que dans un premier temps, un seul centre serait suffisant. Il pourrait cependant prendre de l'ampleur, dans le cadre de son contrat ou d'une décision ad hoc (prise au bout de deux ans par exemple). Il s'en suivrait une augmentation de taille, ou la création d'une deuxième unité par une excroissance sur un site distinct. Le second centre pourrait être créé avec du personnel issu de, ou formé dans le premier centre. Il faudrait dans ce cas garder une direction commune.

### 3) *La localisation du centre :*

La décision du lieu d'implantation du centre, doit être prise avec grand sérieux et en tenant compte de plusieurs paramètres. D'abord, l'installation de ce centre ne doit pas se faire dans le cadre d'une politique d'aménagement du territoire. Il est certain que dans la période présente les acteurs de la vie locale feront un effort financier et/ou politique pour influencer toutes décisions concernant son implantation. C'est certainement une bonne chose que d'aider et de manifester de l'intérêt pour ce projet mais, il faut savoir que, **la bonne technologie se fait dans de bons endroits.**

Le centre se voudrait d'avoir une dimension nationale, européenne voire internationale. Donc il doit être facile d'accès.

Le centre voudrait développer des sujets propres et travailler en collaboration. Donc, il doit être implanté **près d'autres centres** ou la recherche publique et/ou privée est dense.

Le centre peut être attracteur pour d'autres projets publics ou privés. Donc sa localisation est importante.

### 4) *Le centre conçu pour fonctionner sur une période donnée:*

En fonction du volume et de la nature du projet qu'il aura mission de mener à terme, le centre devra **disposer d'un budget**. Si le centre veut être efficace **dès le départ**, ce budget devra lui être alloué et ce sur une période de **5 ans**. **C'est une condition nécessaire à sa création**. Ainsi les responsables pourront travailler et ne pas se préoccuper de trouver de l'argent. Il est à noter que les grands centres bénéficient tous d'un financement dès le départ sur une période allant de 5 à 7ans. **Ce centre doit fonctionner de manière autonome.**

Cela n'exclue pas que le centre puisse être financé par ailleurs: CEE, contrats publics ou industriels, associations caritatives, fondations diverses. Ce point devra être discuté au niveau des gestionnaires du centre.

On peut imaginer par exemple, qu'après 2-3 ans d'activité, une *analyse critique* de son fonctionnement soit faite, avec à la suite, décision de poursuite, de fermeture ou d'extension. Au bout des cinq ans, le maintien du centre devra être décidé.

Dès le départ, il faut recruter des techniciens BT et BTS et quelques ingénieurs. Il peut y avoir du personnel provenant d'établissements publics mais dans une faible proportion. Il faut être conscient que cette activité est sujette à un renouvellement rapide du personnel, soit parce que le travail, s'il n'est pas bien organisé, peut vite devenir monotone, soit parce que le prestige de cette activité et les compétences de son personnel puisse intéresser d'autres activités. Il faut être prêt à pallier à toute déficience de personnel car il en va du bon rendement du centre.

### 5) *Le centre dans le tissu français:*

La création de ce centre nécessitera un énorme investissement tant financier qu'intellectuel. En France, il existe des structures ayant une certaine compétence. Ces structures relèvent du secteur public, privé ou en relation avec des associations caritatives. Aussi serait-il bon dans un premier temps de rencontrer le centre de gestion informatique "Infogène" de Villejuif, les équipes du Genethon, les associations comme l'AFM ou luttant contre le cancer, les fondations comme le CEPH ainsi que des sociétés privées travaillant dans ce domaine comme GENSET.



### 6) La gestion du centre:

Afin de responsabiliser les pouvoirs publics, et les personnes extérieures ou intérieures à, et concernées par ce centre, il faudrait installer deux types de comités.

a) Un comité à connotation politique et économique: Ce comité regrouperait des personnalités représentant les ministères de la recherche, de la santé, de l'industrie et du monde agricole, des organismes de recherche, des associations caritatives et des fondations. Il devra y avoir également des industriels, quelques scientifiques. Ce comité se **préoccupera des problèmes juridiques, éthiques et financiers**. Il est difficile de lui dénier le droit d'intervention dans les orientations que prendra le centre.

b) Un comité scientifique composé de personnalités scientifiques compétentes nationales et internationales, du monde public ou privé ainsi que des représentants des organismes de recherche. Il aura pour mission d'évaluer les projets, d'analyser les travaux et d'apprécier la qualité et la compétence du centre. Il aidera à la veille technologique et proposera des choix.

c) Un comité scientifique interne au centre devra être mis en place comme dans toute structure de ce type.

### 7) Le responsable du centre:

Le succès de l'opération grand séquençage dépend en partie du ou des leaders que l'on aura mis en place. Le **responsable** devra avoir, outre des compétences en séquençage, analyse et informatique, des qualités d'organisateur et un esprit quelque peu "industrie" (être le plus rentable, et avoir en tête le côté contrat à honorer). Ce responsable qui doit être un scientifique devra travailler **avec un gestionnaire** qui s'occupera des problèmes de logistique et d'infrastructure afin que lui même n'ait à s'occuper que de la partie technologique et scientifique. Il est bon et même souhaité que le responsable du centre ait un projet propre dont l'intérêt sera apprécié par les comités mentionnés ci-dessus.

Il est important que le responsable soit reconnu au niveau international de par son action dans le domaine du séquençage proprement dit ou de par l'impact de ses travaux en biologie ou génétique. Ainsi il aura plus facilement accès à l'information et pourra surtout beaucoup plus vite s'intégrer dans les différents réseaux internationaux. Il est également absolument indispensable que le responsable de ce centre n'ait que (et uniquement) la responsabilité du centre et n'exerce aucune autre fonction annexe, afin de pouvoir se consacrer entièrement à ce projet. N'en faisons pas non plus comme on en a trop souvent l'habitude en France, le Mr Séquençage du Génome que l'on exhibera dans diverses manifestations et comités quelconques. Laissons-le et donnons-lui tous les

moyens de s'occuper entièrement de son projet. Il y a un énorme travail et il y a tout à monter.

Le responsable devra également réunir autour de lui des compétences très diverses: génétique humaine, biologie moléculaire et structurale, informatique et mathématiques. Ces compétences seront acquises grâce à un personnel propre ou par contrat pour résoudre un problème bien particulier. On ne peut également exclure que certaines questions comme par exemple celle concernant la mise au point d'algorithme soit sous-traitée non seulement par des équipes provenant des organismes publics mais également par des sociétés privées. Il est également possible et même recommandé qu'une partie de ce travail s'intègre dans des réseaux comme le centre de ressource de Heidelberg.

Ce centre devra avoir à court ou moyen terme une dimension européenne, par le biais de collaborations et de contrats avec différents partenaires. Il pourrait prétendre être le centre ou l'un des centres européens de référence avec des projets ou programmes issus de Bruxelles ou en , donc il faut dès le départ avoir un recrutement large.

### 8) Une place clé à l'informatique:

L'informatique est un des paramètres important pour le succès de ce centre. Elle conditionne le stockage, l'analyse et la validation des données issues du séquençage. L'efficacité et la rapidité dans l'analyse (polishing) de ces données permettent d'augmenter de façon considérable non seulement leur qualité, mais également le rendement. L'excellence des outils et des informaticiens français est universellement reconnue, notamment par les spécialistes des banques de données de séquences. L'excellence d'un bon outil informatique serait un atout supplémentaire qui pourrait marquer notre spécificité par rapport aux autres centres. **Un bon informaticien ayant des compétences en séquençage est donc indispensable. Le souhait et la logique voulant que l'on s'intéresse à l'après gène,** il nous faut également une compétence en Structure/Fonction.

## VI) Projets du Centre de Grand Séquençage :

Il faut que dès le départ soit précisé, cahier des charges à l'appui, le(s) projet(s) du centre: objectif et temps. En d'autres termes, **quoi séquençer, combien de Mb et en combien de temps?**

En dehors du débat sur l'opportunité d'un centre de grand séquençage, il en est un autre qui tient à la **définition** même du ou **des projets**.

Il y a des besoins en séquençage pour les cinq ans à venir comme l'indique le Tableau 2. Il a été difficile d'avoir des valeurs et des projets plus précis en dehors de l'INRA. Il se peut que la création du centre amènera et générera des projets

spécifiques supplémentaires. La participation éventuelle au projet séquençage du génome humain n'est pas comptabilisée; elle devra faire l'objet de discussions approfondies au niveau international. Le projet propre que pourrait développer le responsable du centre n'est pas indiqué également. On peut facilement imaginer qu'il y aurait pour les cinq prochaines années au moins 200 Mb à séquencer.

### 1) Faut-il participer au séquençage du génome humain?

Le séquençage du génome humain est pour les deux tiers, aux mains des anglo-saxons; les centres de St Louis et de Cambridge projettent à eux seuls, de séquencer 1000 Mb chacun. C'est un projet. A noter à ce propos qu'ils n'excluent pas et cela est scientifiquement compréhensible d'intervenir dans les 1000 Mb restantes par le biais de programmes ou de collaborations diverses.

Tout d'abord, il n'est pas vrai que nous n'aurons pas accès aux séquences. **Toute séquence, au dire même des responsables des centres, sera accessible.**

- Par contre, seront vraisemblablement brevetées, les séquences dont on aura démontré la fonction. Il y a des cDNA dont la séquence est soit inaccessible soit brevetée. Les deux centres de grand séquençage ne sont pas concernés par ce problème: le centre Sanger n'a déposé aucun brevet sur les séquences. Les discussions et débats à ce sujet se poursuivent encore à l'heure actuelle aux USA.

- D'autre part, Il semble **certain** que le séquençage du génome humain ne **soit pas terminé pour 2002**. Il y a beaucoup de **problèmes techniques**:

Afin d'évaluer le degré de faisabilité de leur projet génome humain, les USA lancent des opérations pilotes tests dont les résultats seront longuement analysés avant d'aller plus avant. De plus, ils ont d'énormes restrictions financières et les prévisions de financement de ces projets sont révisées à la baisse. Ils sont même à la recherche de partenaires pour poursuivre leurs actions. C'est dire que l'accès à leur technologie et à leurs banques de DNA sont possibles.

Le centre Sanger dont l'objectif était de séquencer l'équivalent de 1000Mb pour l'an 2002, soit 1/3 du génome, n'a reçu un financement (10 M livres/an pour 7 ans), que pour séquencer 500Mb. Ils envisagent de séquencer pour 1997/98, c'est à dire pour dans deux ans et demi, 70 Mb alors qu'ils ont séquencé en gros tous chromosomes confondus, environ 3-4 Mb en 1995. Il semble que cet objectif soit également très optimiste.

En conséquence, **il est possible de participer à l'aventure génome humain, car il y a encore de quoi faire.**

Par contre, il semble difficile de créer un centre de séquençage dont une partie de l'activité ne soit concernée par le génome humain.

### 2) Les autres génomes:

En ce qui concerne le secteur santé publique, les choix sont plus délicats. Tout d'abord, en dehors du génome humain, les génomes des organismes pathogènes les plus importants sont en train d'être séquencés ou en voie de l'être. En ce qui concerne l'aspect vétérinaire, le vaccin étant très peu onéreux, cette industrie hésite à s'offrir de la high tech.

Les autres génomes intéressants (Tableau 2), concernent l'agriculture: certains génomes animaux comme celui du porc, ceux des bactéries lactiques ou des plantes. Deux organismes séduisants en raison de leur taille et d'intérêts divers seraient Arabidopsis et le poisson Fugu; des projets ponctuels sur ces deux espèces sont déjà lancés. Les archaebactéries présentent également un intérêt. Là aussi des projets sont en cours.

Quoiqu'il en soit, le choix du sujet devra être fait après une étude sérieuse tenant compte du contexte international.

### 3) Faut-il breveter?

Il n'est pas question de produire de l'information qui comme trop souvent dans ce domaine sera exploitée par autrui. Le débat sur la brevetabilité ou non de ce que l'on aimerait appeler surtout pour le génome humain, "élément du patrimoine de l'humanité" est un débat auquel il faut certes participer, mais que cela ne se fasse pas au détriment d'un futur par trop diffus. En attendant que des décisions soient prises à ce niveau, faisons comme si et même comme beaucoup d'autres, brevetons ce qui peut l'être, nous serons mieux à même de négocier et de faire part de notre générosité le moment venu.

Il est clair également que les séquences auront beaucoup plus d'intérêt s'il a été démontré au préalable qu'elles permettaient l'expression d'une fonction. Cela signifie qu'il serait bon que, comme nous l'avons déjà souligné, la création de ce centre se fasse dans un cadre plus large concernant l'après gène.

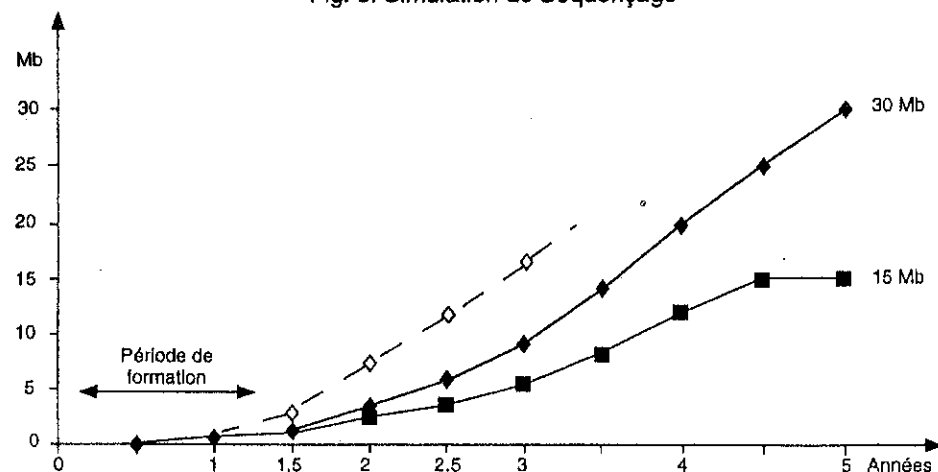
## VII) Structure du centre de grand séquençage.

Faut-il considérer ce centre comme un instrument prestataire de service? Et si oui comment? Ce type de service doit-il être accessible à tout un chacun, doit-il être gratuit ou doit-il faire l'objet d'agréments entre les différentes parties concernées? Cela devra faire l'objet de réflexions.

Un tel centre devrait développer environ quatre types d'activités:

- 1) Préparation des banques et des échantillons.
- 2) Séquençage proprement dit.

Fig. 3: Simulation de Séquençage



Cette figure simule les capacités de séquençage d'un centre français selon que l'on projette un financement pour 15 Mb (3 groupes) ou 30 Mb (6 groupes ; se référer au Tableau 2 également). La courbe (-----), montre que pour un même projet, on puisse avoir des rendements supérieurs consécutifs à une évolution technologique attendue et à une compétence bien établie.

3) Accumulation et validation des données.

4) Exploitation des résultats.

L'idée serait dans un premier temps de démarrer le centre vers mars/avril 1996 en se donnant comme objectif pour la fin de 96 d'un séquençage de l'ordre de un à deux Mb. Ainsi, le personnel pourrait immédiatement se familiariser avec des techniques simples. Démarrer en créant un seul centre avec un objectif modeste et raisonnable qui consisterait à atteindre un rythme de 25-30 Mb/an. Ce rythme sera fonction du temps nécessaire à former le personnel. (Figure 3). Il prendra également en compte le type de génome choisi comme l'indique les valeurs Tableau 1, ainsi que l'évolution des technologies. Ainsi, il n'est pas exclu que l'on puisse atteindre des objectifs nettement supérieurs comme indiqué sur la courbe en pointillé. Ces valeurs sont issues de l'expérience menée sur le projet nématode comme mentionné Figure 2.

Différentes stratégies pourraient être étudiées pour atteindre de tels objectifs. Il serait possible d'acquérir dans un temps relativement bref, la technologie de séquençage en utilisant directement des échantillons provenant d'autres centres ou groupes. On apprendrait ainsi pendant les neuf premiers mois à faire marcher les machines. Dans le même temps, on apprendrait à préparer le matériel, à faire des banques, affiner des cartes sur un projet bien défini et qui serait le sujet propre du centre. Les résultats issus du séquençage ainsi que ceux existant par ailleurs pourront permettre de se familiariser avec les technologies informatiques.

**C onclusion:**

Il n'est certainement pas inutile de créer un centre de grand séquençage véritablement opérationnel en France; ce voeu répond à un besoin de la communauté scientifique. Il pourra alors intéresser le secteur économique, conforter et générer d'autres projets. Le succès de cette opération est cependant conditionné par les points suivant:

1) La création et l'activité du centre devra s'insérer dans un large contexte tenant compte de l'activité scientifique et économique liée à l'après gène.

2) Une telle création devra avoir une dimension européenne. Une coordination de ce projet avec des développements analogues en particulier en Allemagne et dans la CEE sera indispensable.

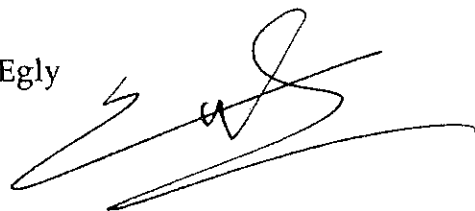
3) Ce centre devra être autonome et avoir dès sa création un budget propre pour une période minimale de 5 ans. Les premières estimations permettant de séquencer 30 +/- 8 Mb/an font état d'un investissement de 75 MF la première année, avec un fonctionnement moyen de 50-55 MF/an.

4) Ce centre dont on aura choisi après mure réflexion, le lieu et le(s) responsable(s), devra avoir un projet précis avec devoir d'objectif. Une partie de son activité devra concerner le génome humain.

PS. Beaucoup de personnes ont pris sur leur temps pour m'éclairer sur les problèmes du séquençage, ne voulant les citer tous par crainte d'ommissions, qu'ils sachent que je les en remercie.

Strasbourg le 18 Janvier 1996

J.M. Egly



20 21  
Tableau 1  
Rendement et Coût de Séquençage

		A		B	
		2 électrophorèses/jour 32 échantillons <sup>1a</sup> 400 bases lues <sup>2</sup>		2 électrophorèses/jour 60 échantillons <sup>1b</sup> 400 bases lues	
1 groupe 4 séquenceurs		4 x 2 x 32 x 400		4 x 2 x 60 x 400	
1 jour		0,1 Mb		0,19 Mb	
1 semaine 9 électrophorèses		0,56 Mb		1,05 Mb	
1 an (40 semaines)		22,4 Mb		42 Mb	
Coefficient d'efficacité	12 % <sup>3</sup>	2,8 Mb		5 Mb	
	20 % <sup>3</sup>		4,5 Mb		8,4 Mb
3 groupes 12 séquenceurs		8,4 Mb	13,5 Mb	15 Mb	25,2 Mb
Coût de fonctionnement	1,8F/b <sup>4</sup>	15 MF	19,5MF <sup>5</sup>	19,5MF <sup>6</sup>	26MF <sup>5</sup>
	5F/b <sup>4</sup>	45 MF	60MF <sup>5</sup>	60MF <sup>6</sup>	78MF <sup>5</sup>
6 groupes 24 séquenceurs		17 Mb	27 Mb	30 Mb	45 Mb
Coût de fonctionnement	1,8/b <sup>4</sup>	30 MF	40 MF	40 MF	70 MF

1. Selon le système mis en place, il peut y avoir 32 (a) ou 60 (b) échantillons déposés.
2. Rendements en lecture entre 70% (acceptable) et 85 % (très bon) de bases effectivement lues. Lors d'une électrophorèse, il peut être lu 500 bases dans le cas optimum. En fait en réalité, 400 le sont, ce qui est correct.
3. Ce coefficient rend compte de nombre de bases qu'il faut lire, afin d'obtenir une séquence prête à être déposée dans une banque de données. Il y a en général 5 redondances pour le cas A.
4. Le coût de fonctionnement est calculé à partir de données du Sanger Center. Les paramètres pris en compte sont les suivants: nombre d'échantillons; nombre de bases lues; nombre de personnes; nombre d'appareillage; infrastructure du laboratoire et salaires. Dans ces chiffres, ne sont pas inclus le service informatique qui doit compter 20 personnes dont 5 ingénieurs d'entretiens pour toute l'informatique du centre.  
Selon ces critères, les équipes les plus performantes (best team) et travaillant sur des sujets relativement simples (Nematode) permettent d'atteindre des prix de revient de 1,5-1,8 F/base séquencée (17-20 pences/base) alors que les équipes qui sont en période de croissance, et avec des sujets plus complexes (Genome humain), ont des coûts de 4,5-5,4F/base
5. Il est évident que selon l'évolution de la technologie, on pourra atteindre un rendement proche des 20 % (13,5 Mb), donc aux augmentations du nombre de bases réellement séquencées de plus de 60 % pour une augmentation de l'ordre de 25-30 % (19,5MF au lieu de 15 MF). De même, avec des appareils plus sophistiqués (cas B), arrivant très prochainement sur le marché, il y aura également une augmentation de bases lues, pour des coûts de revient plus bas.  
Il faut noter également que le coût des appareillages pour 3 groupes de 12 personnes serait environ de 13 MF (séquençage); 4 MF (préparation des échantillons) et 2MF (informatique). Pour fixer les idées le coût des produits pour le séquençage revient (cas A) à 3 centimes par base lue.

Tableau 2  
Besoins en Séquençage<sup>1</sup>  
au niveau français

<i>Organisme</i>	<i>Domaine</i>	<i>Quantité Mb</i>
CEA	Projets en collaborations avec d'autres organismes et en cours de définition	5
CNRS	Besoins recouvrent des projets avec d'autres organismes et projets propres	10-15
IFREMER	Archaeobacteries <sup>2</sup> Plasmides et Virus	2 <sup>2</sup>
INRA	Genome Porcin Genomes Bacteriens	20-25 30
INSERM	Concerne surtout le Genome Humain ; et les maladies génétiques	5-10
Milieu Industriel	Groupes comme Rhone-Poulenc SANOFI, L'OREAL ont été contactés	5-10

1. Estimations Prévisionnelles pour les 5 prochaines années, tous organismes confondus.
2. Equipes jeunes en cours de formation; elles sont en train de définir ce type de projets.