
L'appel de l'ISO

Préparer une épreuve de standardisation en génomique
computationnelle

ISO calling: preparing a standardization trial in computational genomics

*Llamada ISO: preparación de un ensayo de estandarización en genómica
computacional*

ISO-Aufruf: Vorbereitung eines Normungsprozesses in der Computergenomik

Florian Jatón



Édition électronique

URL : <https://journals.openedition.org/rac/38445>

DOI : 10.4000/14k4a

ISSN : 1760-5393

Traduction(s) :

ISO calling - URL : <https://journals.openedition.org/rac/38578> [en]

Éditeur

Société d'Anthropologie des Connaissances

Ce document vous est fourni par Bibliothèque cantonale et universitaire Lausanne



Référence électronique

Florian Jatón, « L'appel de l'ISO », *Revue d'anthropologie des connaissances* [En ligne], 19-3 | 2025, mis en ligne le 28 août 2025, consulté le 01 septembre 2025. URL : <http://journals.openedition.org/rac/38445> ; DOI : <https://doi.org/10.4000/14k4a>

Ce document a été généré automatiquement le 1 septembre 2025.



Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont « Tous droits réservés », sauf mention contraire.

L'appel de l'ISO

Préparer une épreuve de standardisation en génomique computationnelle

ISO calling: preparing a standardization trial in computational genomics

Llamada ISO: preparación de un ensayo de estandarización en genómica computacional

ISO-Aufruf: Vorbereitung eines Normungsprozesses in der Computergenomik

Florian Jatón

Introduction

- 1 Depuis au moins une trentaine d'années, des chercheureuses en *Science & Technology Studies* (STS¹) ont décrit le rôle central des normes et des standards dans la construction et la diffusion d'énoncés scientifiques (Latour, 1987 ; O'Connel, 1993), d'artefacts techniques (Mallard, 1998 ; Jatón, 2017) et de procédés biomédicaux (Cambrosio *et al.*, 2006 ; Timmermans & Berg, 2003 ; Timmermans, 2015). En montrant que les mesures de fiabilité, de précision et d'efficacité dépendent de longs et coûteux réseaux référentiels stabilisés (et stabilisateurs), ces enquêtes sont parvenues à dépeindre le savoir technoscientifique et biomédical comme à la fois construit, objectif et performatif (Busch, 2013 ; Callon *et al.*, 1999 ; Timmermans & Epstein, 2010).
- 2 Cette ligne de recherche, précieuse, souffre pourtant d'un double point aveugle. Le premier concerne les sciences et techniques *bio-informatiques* – aujourd'hui centrales à la conduite de la biomédecine contemporaine (Chiapperino *et al.*, 2024) – qui reposent sur quantités de standards et bonnes pratiques, dont les ressorts restent encore peu explorés (Hà & Chow-White, 2021). Le deuxième concerne les nombreux tests, contrôles, audits et autres *épreuves* de standardisation, ainsi que leur préparation (Riom & Tanferri, 2024), qui font partie intégrante de l'écologie des pratiques technoscientifiques et biomédicales sans pour autant faire l'objet d'études sociales approfondies.

- 3 En documentant une partie de la préparation à une épreuve de standardisation d'un pipeline bio-informatique – plus précisément, une procédure d'*accréditation* nationale visant à garantir la conformité d'une suite logicielle à un standard international de contrôle qualité (qui lui-même ouvre des opportunités commerciales) –, cet article entend combler, en partie, cette double lacune. Fruit d'une enquête ethnographique de quatre mois au sein d'un centre suisse de génomique spécialisé dans l'appel de variants (*variant calling*) pour l'analyse de la lignée germinale (*germline analysis*) en contexte clinique, l'article enrichit la compréhension des processus de standardisation en biomédecine, et plus particulièrement en bio-informatique, en mettant en évidence à la fois les enjeux tactiques qui les sous-tendent et la marge d'interprétation qu'ils permettent. Mais aussi, sur un plan plus anthropologique, le texte enrichit la compréhension des phénomènes de préparation en génomique computationnelle, en soulignant la production située d'un type particulier de connaissance – à la fois épistémique, critique et stratégique – visant à conformer une plateforme bio-informatique à un standard international de qualité.
- 4 L'article est organisé comme suit. Dans un premier temps, il revient sur l'importance des études sociales des processus de standardisation, tout en soulignant l'intérêt de les associer aux études sur la préparation et sur l'informatique en action. Après un bref détour méthodologique, l'article plonge ensuite dans son terrain d'enquête pour présenter l'arrangement institutionnel, légal et technoscientifique qui façonne et influence le désir d'accréditation des acteur·rices rencontrés·es. Enfin, l'article se focalise sur une réunion de travail décisive, où des bio-informaticien·nes préparent et équipent leur demande d'accréditation en comparant et discutant deux suites logicielles pour l'appel de variants pour l'aide à la décision clinique. L'article finit par discuter les perspectives nouvelles sur l'étude sociale des processus de standardisation qu'ouvre cette analyse, en mettant en lumière les dynamiques de négociation, d'adaptation et de matérialisation des normes et standards dans les pratiques bio-informatiques.

Quand l'étude sociale de la standardisation rencontre celle de la préparation (et de l'informatique en action)

- 5 Malgré leur apparence de 'choses ennuyeuses' (Lampland & Star, 2008, p. 11), les standards – et, plus largement, les processus de standardisation – ont fait l'objet de nombreuses publications en études sociales des sciences et des techniques. L'histoire des sciences fut sans doute la première à s'y atteler, avec notamment les travaux aujourd'hui classiques de Robert Kohler (1994) sur le façonnage collectif de la drosophile standard pour l'expérimentation génétique, Eviatar Zerubavel (1982) sur la construction du système de fuseaux horaires internationaux normalisés, ou encore – quoique de façon un peu décalée – Michel Foucault (1963 ; 1975) sur les entreprises disciplinaires de normalisation et d'uniformisation des corps. La sociologie des sciences et des techniques, dans le sillage de Bruno Latour (1987, pp. 116-160), se saisit ensuite de la thématique, avec entre autres les importants travaux de Joseph O'Connell (1993) et Alexandre Mallard (1998) sur les pratiques métrologiques dans les technosciences. Ces travaux empiriques – ainsi que certaines réflexions sociohistoriques plus larges entreprises, notamment par Alain Desrosières sur la raison statistique (2000) – ont fortement contribué à nourrir des réflexions plus fondamentales sur le rôle des

standards dans le façonnage du collectif, comme celles proposées par Geoffrey Bowker et Susan Leigh Star (2000) sur les modes de classification, Laurent Thévenot sur les régimes d'engagement (2009), ou encore Lawrence Busch (2013), pour qui les standards deviennent même ce qui permet de faire tenir le collectif.

- 6 Mais c'est dans le champ de l'étude sociale de la biomédecine que l'intérêt pour les standards a été – et demeure – le plus marqué, sans doute pour des questions de visibilité. En effet, comme le notent Stefan Timmermans et Steven Epstein (2010, p. 78), la biomédecine est un domaine dans lequel les conflits liés à la standardisation de l'être humain sont particulièrement vifs, suscitant dès lors de nombreuses controverses sociotechniques, constituant elles-mêmes des points d'entrée privilégiés pour l'analyse sociale (Rip, 1986). D'où une série d'enquêtes et de propositions importantes, notamment portées par Alberto Cambrosio et Peter Keating (1992), ainsi que Monica Casper et Adele Clarke (1998), qui ont montré que les équipements et nomenclatures qui soutiennent la recherche et les traitements biomédicaux sont pour beaucoup issus de processus institutionnels fortement chargés d'enjeux de pouvoir². Mais ce sont surtout Timmermans et ses collègues qui ont ensuite magnifié l'analyse des processus de standardisation en biomédecine, en documentant par exemple les multiples réappropriations que permettent, en pratique, les protocoles et les standards (Timmermans & Berg, 1997 ; 2003), notamment en génomique (Timmermans, 2015), ainsi que les détournements qui en font parfois des outils de contestation technopolitique (Timmermans & Almeling, 2009).
- 7 Aussi intéressantes et importantes soient-elles, les enquêtes socio-anthropologiques sur les processus de standardisation en biomédecine présentent certaines lacunes, notamment l'absence quasi totale de considération pour les disciplines bio-informatiques. Bien que les travaux de Hallam Stevens aient finement documenté la structuration de ce champ disciplinaire hybride (2013) et les modes de problématisation spécifiques, de plus en plus naturalisés, que ses praticien·nes ont contribué à diffuser dans la recherche biomédicale (Stevens, 2016), peu a été dit sur les processus de normalisation et de standardisation qui le traversent et le façonnent. Cette omission est d'autant plus regrettable que la bio-informatique, et plus largement la mobilisation de ressources computationnelles massives, joue un rôle croissant dans les représentations et les pratiques des professionnel·les de la santé (Chiapperino *et al.*, 2024) tout en reposant fortement – comme l'ensemble de l'informatique appliquée (Jaton, 2021a ; 2021b) – sur la diffusion et le partage de normes et standards plus ou moins arbitraires (Hà & Chow-White, 2021).
- 8 Une deuxième lacune réside dans les types de situations étudiées, qui se déroulent le plus souvent dans des contextes routinisés. En effet, si quelques rares enquêtes ont effectivement tenté de documenter précisément certaines pratiques d'acteur·rices confrontées à des normes et des standards (e.g., Timmermans, 1997), pratiquement aucune ne s'est risquée à décrire les pratiques situées dont dépend le *passage* d'un agencement non-standardisé à un agencement standardisé. En somme, si l'on dispose d'éléments sur les pratiques standards (pour ne pas dire standardisées), très peu a été dit sur les mécanismes propres aux *épreuves de standardisation* (tests, contrôles, audits), moments charnières dont l'issue discrétionnaire permet, ou non de faire advenir des équivalences et, en fin de compte, d'étendre son réseau³. Là réside, peut-être, un chaînon manquant de l'étude sociale de la standardisation, qui lui permettrait par

ailleurs de préciser les différences – fines mais cruciales pour les acteur·rices concernées – entre standardisation, certification et accréditation⁴.

- 9 Ces derniers éléments pointent à leur tour vers la question de la *préparation*, qui a récemment fait l'objet d'un regain d'attention au sein de l'anthropologie des connaissances (Riom & Tanferri, 2024). En effet, les épreuves de standardisation en biomédecine ne peuvent s'improviser, compte tenu de leurs conditions de réussite fortement encadrées et des importants enjeux financiers, commerciaux et institutionnels qui y sont associés. Ces épreuves nécessitent donc une *préparation minutieuse*, dont la nature influence inévitablement les domaines d'activité concernés. Cependant, la manière dont cette préparation s'opère et les moyens par lesquels elle s'effectue restent encore mal documentés. En somme, les compétences, expertises et savoirs mobilisés dans les *préparations aux épreuves de standardisation* demeurent méconnus, bien qu'il soit évident qu'ils ont un impact significatif sur les univers biomédicaux en quête de standardisation.
- 10 C'est précisément pour contribuer à combler ces lacunes que le présent article se propose de documenter la préparation à une épreuve de standardisation dans un domaine affilié à la bio-informatique. En s'appuyant sur le genre analytique des études de laboratoires informatiques (Peerbaye & Vinck, 2023, pp. 53-54), cet article vise à décrire minutieusement certaines des pratiques qui contribuent à ce phénomène sociotechnique, ainsi que les injonctions institutionnelles qui le mettent en branle. Et ceci implique le recours à une méthode d'enquête spécifique, qu'il convient maintenant de présenter.

Méthode

- 11 Le texte est issu d'une enquête ethnographique menée entre mars et juin 2022 au sein d'un centre suisse spécialisé en génomique computationnelle. En s'inspirant de la méthode proposée par Vinck *et al.* (2018) pour enquêter sur les objets numériques ainsi que celle proposée par Jatton (2019) et par Jatton & Vinck (2023) pour documenter la fabrication des modèles informatiques, l'ambition était de documenter une partie du travail bio-informatique ordinaire et concret dont dépend, en coulisse, le développement de la médecine dite personnalisée en Suisse⁵ (mais aussi ailleurs).
- 12 L'ambition avant tout descriptive de l'enquête fut bien accueillie par le centre en génomique – que j'appellerai dorénavant le Centre⁶ –, en commençant par le membre de son comité directeur chargé des affaires bio-informatiques. Ce professionnel respecté, actif depuis de nombreuses années dans la question du *benchmarking* et de la standardisation en génomique computationnelle, avait à cœur de montrer, et sans doute de démontrer, la nécessité des processus a priori ennuyeux de standardisation pour le développement de la médecine dite personnalisée. Ce soutien, sinon à ma recherche elle-même du moins à la thématique de *l'infra-ordinaire* bio-informatique (Lefebvre, 2013), m'a permis d'obtenir un accès privilégié au Centre. J'ai ainsi pu échanger librement avec son personnel et assister à toutes leurs séances de travail durant la durée de mon enquête. En revanche, pour des raisons évidentes de sécurité, l'accès direct aux ressources bio-informatiques m'était interdit.
- 13 Entre début mars et fin juin 2022, j'ai donc bénéficié d'un accès au Centre ainsi que d'un espace de travail dédié. À l'instar de la plupart des bioinformaticien·nes, je m'y rendais

en moyenne deux fois par semaine, tandis que les autres jours, en télétravail, étaient consacrés aux analyses standards, à la maintenance des outils, à la réflexion stratégique ainsi qu'aux séances de coordination (auxquelles j'étais convié). Ce format d'enquête m'a permis d'assister à quinze séances de l'équipe bio-informatique et huit séances de coordination. J'ai également mené une trentaine d'entretiens avec l'ensemble des collaborateur·rices du Centre et ai pu observer quatre situations de programmation informatique, selon la méthode proposée par Jaton (2022).

- 14 Dès le début de mon enquête en mars 2022, la question de l'accréditation de la section bio-informatique du Centre était une préoccupation majeure pour ses membres – et par conséquent aussi pour moi. Cette accréditation, comme nous le verrons en détail, concernait la conformité d'analyses bio-informatiques complexes à la norme 15189 de l'*International Standard Organisation*⁷ (ISO), qui définit les exigences de qualité applicables aux laboratoires d'analyses médicales. Ce processus d'accréditation, à la fois rigoureux et exigeant, impliquait une préparation minutieuse en vue d'un audit officiel coûteux. Il impliquait notamment la rédaction d'un rapport détaillant avec précision le respect strict du standard ISO 15189, rapport qui servirait ensuite de support à l'audit officiel.
- 15 Ce projet d'accréditation, destiné donc à attester de la conformité à un standard reconnu par une autorité de référence, revêtait une importance stratégique pour le Centre, tout en posant de nombreux défis technoscientifiques, notamment en termes de rapidité, de lisibilité et de précision des analyses génomiques. Mais une autre difficulté, comme nous le verrons, résidait dans le fait que ce standard ISO 15189 normant la qualité des laboratoires biomédicaux restait assez évasif quant aux critères qualité des éléments computationnelles et bio-informatiques. Ce standard ne se réduisait donc pas à une simple recette à appliquer ; il imposait au Centre d'élaborer de nouvelles expériences de contrôle qualité tout en développant des interprétations adaptées de leurs résultats, une démarche à la fois exigeante, risquée et créative.
- 16 Rétrospectivement, il m'est apparu qu'une part importante de ces préoccupations et difficultés se sont retrouvées concentrées – pour ne pas dire *pliées* (Conley, 1998) – lors d'une réunion de travail capitale tenue la matinée du 28 mars 2022. C'est cette réunion de travail ainsi que les éléments institutionnels qui pèsent sur elle – qu'il m'a fallu par la suite patiemment démêler – que le texte se propose de décrire méticuleusement, afin d'y dégager des éléments nouveaux sur la pratique de standardisation en génomique computationnelle et, plus largement, sur les ressorts anthropologiques des processus de préparation en bio-informatique.

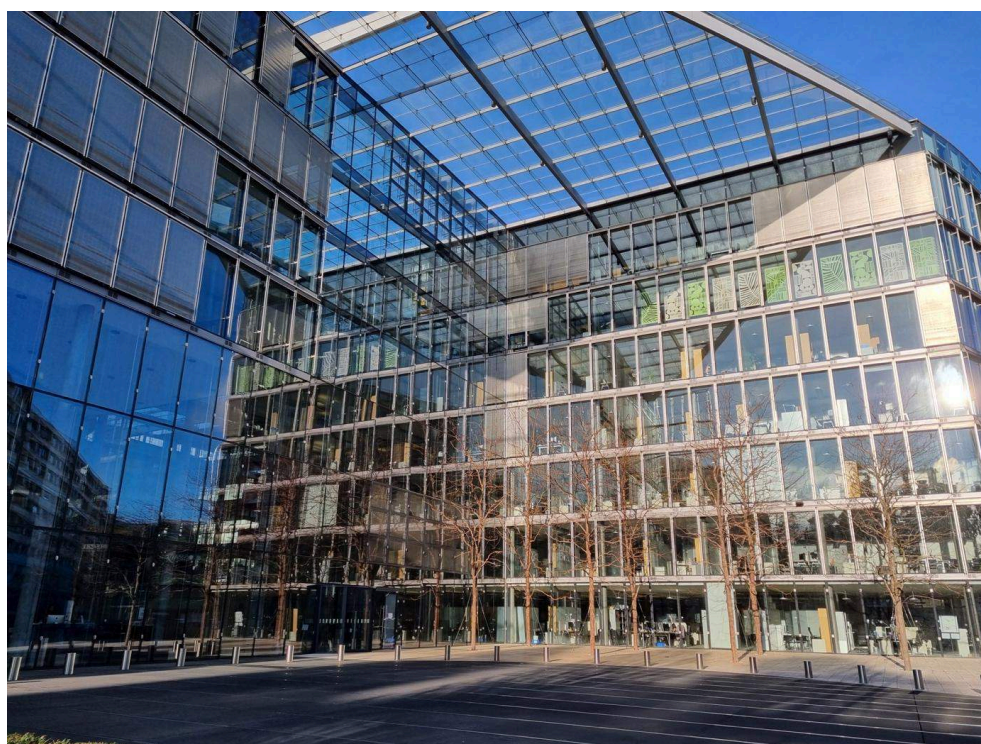
ISO 15189 : Une norme de laboratoire, pour les gouverner tous

Un agencement institutionnel complexe

- 17 De nombreux facteurs institutionnels nourrissent la stratégie du Centre dans sa quête d'accréditation. Commençons donc par tirer ce fil, en précisant que le Centre est constitué de deux plateformes. Une plateforme *computation*, dont les membres sont responsables de ce qui est communément appelé la partie « sèche » (*dry part*) de l'analyse génomique, c'est-à-dire tout ce qui concerne le traitement informatique des données séquencées. La partie dite « humide » (*wet part*), qui englobe la préparation des

échantillons biologiques, l'amplification et le séquençage, est prise en charge par la plateforme *séquençage*, situé trois étages au-dessus de la plateforme computation, au sein du même vaste immeuble dédié à l'innovation biomédicale, appelé ici le parc Biomed (voir fig. 1). Si les deux plateformes entretiennent des liens étroits, les sept membres de la plateforme séquençage sont, pour la plupart, des technicien·nes et scientifiques de laboratoire biomédical, tandis que les sept membres de la plateforme computation sont, pour la plupart, des analystes de données, des ingénieures logiciel et des spécialistes en sécurité informatique, regroupé·es ici sous la dénomination large de bio-informaticien·nes.

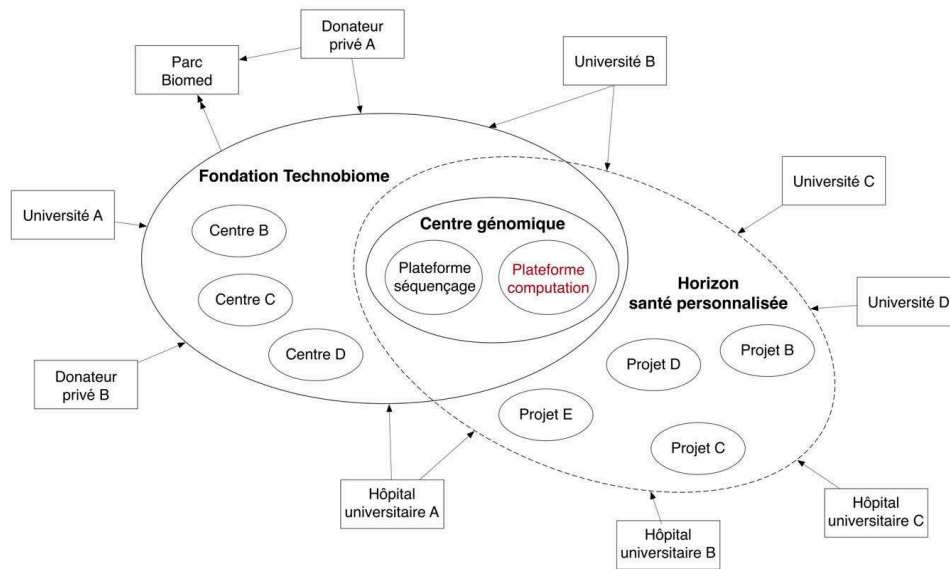
Figure 1 : Cour d'entrée du Parc Biomed



Source : auteur.

- 18 Le Centre, qui regroupe donc la plateforme séquençage et la plateforme computation au sein du parc Biomed, constitue le bras médical d'une initiative plus vaste de recherche et d'innovation : l'Horizon santé personnalisée (appelée ici Horizon), lancée en 2016. Rattachée administrativement à une école polytechnique, cette initiative Horizon vise à coordonner les multiples projets locaux en santé personnalisée et médecine de précision, sous la direction d'un comité de pilotage composé de professeur·es, de vice-recteur·rices et de cadres hospitaliers. Bien que l'initiative Horizon joue un rôle central dans les orientations stratégiques du Centre, elle ne dispose pas des moyens juridiques pour engager directement du personnel salarié. C'est pourquoi, dès sa création, le Centre a été administrativement rattaché à la Fondation Technobiome (ici appelée Fondation), une organisation à but non lucratif soutenue par des donateur·rices privé·es et des institutions académiques, et responsable de la gestion des groupes hébergés au sein du parc Biomed (lui-même une structure autonome sur le plan organisationnel, voir fig. 2).

Figure 2 : Schéma de l'arrangement institutionnel dans lequel s'inscrit le Centre génomique. Les flèches pleines signifient « soutien financier », la double flèche signifie « hébergé par ». Le cercle en pointillé autour de l'initiative Horizon indique sa situation administrative lâche



Source : auteur.

- 19 Bien que l'organisation institutionnelle qui soutient le Centre puisse sembler complexe – et le soit effectivement –, cet enchevêtrement n'implique pas nécessairement de grandes lourdeurs administratives. En effet, la plupart des dirigeant·es des différentes entités concernées – Centre, Horizon, Fondation, etc. – sont les mêmes personnes, qui cumulent plusieurs mandats. Néanmoins, certains obstacles administratifs et juridiques demeurent difficiles à surmonter. Par exemple, bien qu'Horizon et la Fondation, les deux structures de tutelle du Centre, soient soutenues par des universités et des hôpitaux universitaires, cela ne confère pas pour autant au Centre le statut d'institution académique. Ainsi, la plateforme computation, tout comme sa « sœur », la plateforme séquençage, ne sont pas directement affiliées à une université ou à un hôpital universitaire.
- 20 Pour le Centre, être une institution sans affiliation académique directe présente certains avantages, notamment la flexibilité d'embaucher ou de restructurer en fonction des ressources financières disponibles. Mais cela présente aussi des inconvénients notables, notamment l'obligation de se présenter en tant que *prestataire de services* auprès des clients potentiels dont il dépend fortement. En effet, même si le Centre a bénéficié d'un financement important pour son lancement – notamment pour acquérir du matériel de séquençage haut de gamme et attirer du personnel qualifié – la poursuite et l'expansion de ses activités de séquençage et d'analyse dépendent en grande partie des revenus qu'il est en mesure de générer. En ce sens, même s'il fait partie d'une organisation à but non lucratif, le Centre – et ses deux plateformes – peut être considéré comme une *entité semi-commerciale* (Callon, 2017) cherchant à générer des revenus à réinvestir dans son organisation mère, la Fondation. Et c'est précisément ce besoin de générer des revenus pour assurer la pérennité de ses activités – et, idéalement, lancer de nouveaux projets et engager de nouvelles collaborateur·rices – qui nourrit, du moins en partie, la volonté de Centre d'obtenir une accréditation nationale attestant de la fiabilité de ses services de séquençage et d'analyse.

- 21 Mais à quoi renvoie l'accréditation d'un laboratoire d'analyses médicales – actif ou non dans la génomique – pour le cas de la Suisse ? Et quelle est l'institution habilitée à l'accorder ? Et en vertu de quels critères et épreuves ? À ce stade, il est important de présenter brièvement le statut des laboratoires d'analyse médicale au sein du système de santé suisse.

Vous avez dit accréditation ?

- 22 Le système de santé suisse repose sur l'obligation de souscrire à l'une des nombreuses assurances privées établies dans le pays. Le catalogue des prestations remboursées par ces assurances obligatoires dites « de base » est régi par une loi sur l'assurance maladie, appelée LAMal, élaborée en 1999. La plupart des prestations standards fournies par un laboratoire d'analyses médicales en milieu clinique (par exemple, le test PCR mycoplasme) font ainsi partie du catalogue LAMal et sont prises en charge (avec des franchises) par les caisses maladie auxquelles les patient·es sont légalement tenus de souscrire.
- 23 En tant que principaux bailleurs de fonds des laboratoires d'analyses médicales, les assurances privées ont un droit de regard sur la qualité des prestations des laboratoires, qui est évaluée par une *certification* minimale appelée QUALAB⁸. En conséquence, tout laboratoire d'analyse médicale établi en Suisse – petit ou grand – qui propose des prestations remboursées par une caisse maladie doit être certifié QUALAB, ce qui implique de passer chaque année une série de contrôles de qualité internes et externes.
- 24 Si la certification QUALAB permet de contrôler effectivement la qualité des analyses biomédicales réalisées sous la juridiction de la LAMal, elle reste un label de qualité minimal qui n'implique pas d'audits systématiques. Et avec l'essor du séquençage nouvelle génération – processus sensibles, complexes et lourdement équipés –, l'Assemblée fédérale suisse a adopté une loi spécifique exigeant une *autorisation* nationale supplémentaire pour tout laboratoire proposant des analyses génomiques humaines (Loi fédérale sur les analyses génétiques humaines, 2004). Selon cette loi, tout laboratoire suisse – petit ou grand – souhaitant effectuer des analyses de tissus humains séquencés dans un contexte clinique doit se soumettre à des évaluations externes de contrôle de qualité afin d'obtenir une autorisation de l'Office fédéral de la santé publique (OFSP). Ce n'est qu'après l'octroi de cette autorisation de l'OFSP que ces laboratoires médicaux peuvent voir leurs analyses génomiques prises en charge par les assurances de base.
- 25 Cette loi fédérale suisse sur l'analyse génomique comporte également un article qui fait le lien avec une autre loi fédérale, votée elle en 1998 et qui concerne la procréation médicalement assistée (Loi fédérale sur la procréation médicalement assistée, 1998). L'article en question stipule que les laboratoires réalisant des analyses génomiques dans le cadre de la procréation médicalement assistée doivent impérativement obtenir une *accréditation* nationale délivrée par le Service d'accréditation suisse (SAS), l'organisme chargé notamment de l'application de la norme ISO 15189 relative à la qualité des laboratoires de biologie médicale. Ce texte de loi instaure ainsi un *troisième et ultime niveau de contrôle qualité* pour les laboratoires d'analyses médicales en Suisse, structurant de fait le système autour de trois niveaux de contrôle : un premier niveau, basique, qui regroupe les analyses couvertes par la certification QUALAB ; un deuxième

niveau, plus strict, qui englobe les analyses nécessitant une autorisation de l'OFSP ; et enfin un troisième niveau, le plus rigoureux (et coûteux), qui s'applique spécifiquement aux analyses génomiques, incluant – mais sans s'y limiter – celles effectuées dans le cadre de la procréation médicalement assistée.

- 26 Au moment de la présente enquête, il existait en Suisse une soixantaine de laboratoires de génomique humaine autorisés par l'OFSP, dont la grande majorité était affiliée à d'importants hôpitaux, le plus souvent universitaires⁹. Par conséquent, la grande majorité des analyses génomiques effectuées dans un cadre clinique en Suisse étaient réalisées par des laboratoires affiliés à ces mêmes grands hôpitaux, qui se trouvaient ainsi dans une situation de quasi-oligopole, souvent justifiée par des logiques de proximité géographique et institutionnelle¹⁰. Pour le Centre, en tant qu'organisation à but non lucratif faisant partie d'une fondation plus large qui n'est pas directement affiliée à une université ou à un hôpital, entrer sur ce marché de l'analyse génomique humaine n'était pas une tâche facile et attirer des clients ne pouvait se faire – du moins du point de vue de son comité de direction – qu'en *promouvant son expertise distinctive*. En bref, pour faire sa place au sein du marché restreint, mais prometteur, de l'analyse génomique clinique, le Centre devait se présenter comme un centre de génomique original, dont les compétences sont attestées par un *label de qualité supérieure*. Et c'est bien pour se distinguer des autres laboratoires, attirer des clients et gagner en indépendance financière (et aussi par défi) que le Centre a décidé d'*accréditer* son pipeline de séquençage et d'analyse, dans la foulée de son lancement en 2019.
- 27 Le projet d'accréditation du Centre comportait deux volets, dont le premier concernait les activités « humides » de la plateforme de séquençage, à savoir le séquençage de l'exome entier (WES), le séquençage du génome entier (WGS), et le séquençage de l'ARN (RNAseq)¹¹. Après avoir entrepris les démarches préliminaires exigées par le SAS – ce qui a conduit à la désignation d'un responsable qualité et le recours à un sous-traitant spécialisé –, une demande d'accréditation a été déposée début 2020. En décembre, un audit officiel s'est tenu au parc Biomed, où les représentants du SAS ont minutieusement évalué chaque étape, de la réception des échantillons à la production des données séquencées au format FASTQ¹², en se basant sur les exigences décrites au sein du standard ISO 15189. Malgré quelques « non-conformités » mineures relevées, cette épreuve de standardisation a conduit à l'accréditation officielle de la plateforme séquençage du Centre, qui a donc pu brandir fièrement le prestigieux logo « Swiss Accreditation », attestant son expertise unique (voir fig. 3).

Figure 3 : Logo officiel du SAS attestant d'une accréditation réussie



Source : www.sas.admin.ch

- 28 Si l'accréditation de la partie « humide » du Centre en 2020 s'est globalement déroulée avec succès, elle ne constituait que la première étape d'un projet plus ambitieux. Le second volet de l'accréditation du Centre – toujours dans sa volonté de se distinguer au sein d'un marché de l'analyse génomique clinique suisse dominé par les laboratoires affiliés à des hôpitaux universitaires – concernait sa partie « sèche », à savoir l'analyse informatique des données génomiques au sein de la plateforme computation. En effet, selon le comité directeur du Centre, seule l'*extension* de l'accréditation SAS au pipeline de la plateforme computation permettrait au Centre de vendre un service intégré, sérieux et haut de gamme, allant du séquençage des tissus biologiques à la restitution d'analyses sous forme de tableaux de type Excel, seuls vraiment à même d'intéresser les clinicien·nes désireuses de mobiliser rapidement des informations génomiques pour orienter et soutenir des options thérapeutiques. Autrement dit, l'extension de l'accréditation SAS à la partie sèche du Centre – et donc au travail de la plateforme computation – était perçue comme le meilleur moyen de rendre les données séquencées véritablement *utilisables* (Nelson *et al.*, 2013), c'est-à-dire capables d'être mobilisées rapidement, et avec confiance, par des clinicien·nes dans le *hic et nunc* des hôpitaux. Et c'est une partie de l'extension incertaine de cette accréditation – pièce maîtresse d'une ambitieuse stratégie de développement de la médecine de précision en Suisse – que je me proposais de décrire, entre mars et juin 2022, dans la tradition récente, mais maintenant établie des études de laboratoires informatiques.

Un certain besoin de vitesse et d'usabilité

- 29 Les applications pratiques du séquençage et de l'analyse de données génomiques sont nombreuses, allant de la description génétique de nouvelles espèces animales à l'immunothérapie du cancer chez les humains (Jaton, 2023). Pour diverses raisons, notamment l'existence d'acteurs privés, locaux et respectés dans le domaine de l'analyse génomique en oncologie, le Centre a très tôt décidé de se concentrer sur deux domaines spécifiques : 1) la surveillance des mutations génétiques du Sars-Cov-2 (non abordée ici) et 2) l'analyse de la « lignée germinale » (*germline analysis*), c'est-à-dire l'étude du matériel génétique humain transmis à la descendance (typiquement parents-enfants).
- 30 Le pari du Centre sur l'analyse de la lignée germinale peut sembler curieux, car les débouchés cliniques, et donc semi-commerciaux, ne sont a priori pas évidents. En effet, en Suisse, pour que l'analyse génomique de ce que l'on appelle un « trio » (c'est-à-dire le père biologique, la mère et la progéniture) soit prise en charge par l'assurance maladie de base, il faut une raison impérieuse que seul le spécialiste clinique travaillant le plus souvent pour un hôpital universitaire est en mesure d'attester. Par conséquent, les analyses germinales ont tendance à être réalisées dans les laboratoires autorisés par l'OFSP et le plus souvent affiliés aux hôpitaux universitaires (cf. section 3). Alors pourquoi le Centre a-t-il axé une partie importante de sa stratégie semi-commerciale sur l'analyse de la lignée germinale ? Une partie de la réponse réside dans une nouvelle tendance clinique qui a émergé dans les années 2015 et dont les caractéristiques suggèrent, potentiellement, le recours aux services d'une entité externe hautement spécialisée comme le Centre : le séquençage rapide du génome entier pour unité de soins intensifs néonataux (*rapid whole genome sequencing for neonatal intensive care unit*, ici noté SREN-USIN).
- 31 L'histoire du SREN-USIN mériterait un article entier, mais retenons ici l'essentiel. Lancée dans les années 2015, cette approche repose sur le séquençage *rapide* du génome entier en soins intensifs néonataux. Une étude fondatrice a montré que ce procédé, en fournissant un diagnostic en moins de cinq jours, pouvait réduire significativement la mortalité infantile (Willig *et al.*, 2015 ; Petrikin *et al.*, 2015). Ces résultats ont été confirmés en 2018 par une large enquête clinique qui a démontré l'impact du SREN-USIN sur la morbidité, la mortalité et les décisions de soins en fin de vie (Farnaes *et al.*, 2018). Depuis, plusieurs études de contrôle (Kingsmore *et al.*, 2019 ; Clark *et al.*, 2019 ; Dimmock *et al.*, 2021) ont affiné ces conclusions. Aujourd'hui, sous réserve d'une infrastructure adaptée, le SREN-USIN est reconnu comme une avancée fiable et bénéfique en médecine de précision, une réputation pour l'heure confortée par l'ensemble de la littérature spécialisée (Owen *et al.*, 2022).
- 32 En 2020, le développement modeste mais réel du SREN-USIN aux États-Unis, en Australie (Australian Genomics Health Alliance, 2020) et en Chine (Wu *et al.*, 2021) est considéré par le Centre comme une opportunité de poursuivre une cause noble, de générer des revenus et de souligner son expertise. En effet, pourquoi ne pas tenter de mettre en œuvre le SREN-USIN en Suisse ? Cela semble d'autant plus pertinent que le directeur de la plateforme computation – et membre du comité directeur du Centre – entretient depuis longtemps des relations professionnelles étroites avec plusieurs clinicien·nes du plus grand hôpital pédiatrique de Suisse. Cet établissement, où les nouveau·elles-nés suspectés de maladies génétiques sont le plus souvent transférés

et prises en charge, apparaît dès lors comme un partenaire idéal pour initier une collaboration.

- 33 Comme son nom l'indique, le SREN-USIN impose la production *rapide* d'analyses génomiques sur des trios de patients. Par ailleurs, le statut du Centre en tant qu'entité externe exige de fournir des résultats lisibles et directement exploitables par les cliniciens potentiels de cet hôpital pédiatrique de premier plan. Cela s'avère d'autant plus crucial que les bioinformaticien·nes de la plateforme computation ne peuvent rester aux côtés des clinicien·nes pour interpréter les résultats via des lignes de commande. Comme nous le verrons, ces exigences complexes, alliant rapidité et usabilité (Nelson *et al.*, 2013), représentent un défi majeur pour l'accréditation de la plateforme computationnelle.
- 34 En ce qui concerne la rapidité, la détermination des variants génomiques d'un·e patient·e à partir de l'ADN de ses deux géniteur·rices est un processus complexe, nécessitant de nombreuses opérations de calcul non triviales¹³. Lors de la création du Centre, la décision a été prise d'adopter la suite logicielle libre GATK¹⁴ pour cette étape dite d'appel de variants (*variant calling*). Bien que ce choix ait initialement semblé pertinent – GATK est une référence en matière d'analyse et d'interprétation des données génomiques et est largement maîtrisée par les bioinformaticien·nes – il a rapidement montré ses limites, en raison des dizaines d'heures requises pour l'exécution des opérations, complexes, d'appel de variants.
- 35 En 2020, pour remédier au problème de lenteur de GATK, qui le rendait *de facto* impropre à son utilisation pour du SREN-USIN, le comité directeur du Centre a décidé d'acquérir la suite logicielle propriétaire DRAGEN, et son *hardware* dédié, développés et maintenus par la société Illumina¹⁵. Ce fut un succès notable, car une fois DRAGEN installé et ajusté – ce qui a nécessité de longs mois de travail expert en interne – la durée de l'appel de variants pour un génome entier a été réduite à 45 petites minutes.
- 36 Cependant, si les composants logiciels internes de GATK sont largement connus et accessibles, il en va différemment pour DRAGEN. Contrairement à GATK, qui est entièrement *open source*, largement distribué et abondamment discuté, DRAGEN – malgré sa puissance et sa réputation – demeure coûteux et, pour l'essentiel, propriétaire. Ceci limite considérablement sa documentation, qui se résume aux seules informations fournies par son distributeur, Illumina. Pour le Centre, ce problème doit d'être abordé de front : afin de poursuivre son espoir de SREN-USIN dans le cadre de son désir d'accréditation, il devient crucial pour le Centre de mieux comprendre (et rendre des comptes sur) le fonctionnement interne et relativement opaque de DRAGEN ainsi que la manière dont il interagit avec les autres composants du pipeline du Centre.
- 37 Une deuxième problématique liée au désir d'accréditation du Centre, en rapport avec son espoir de SREN-USIN, concerne l'usabilité des résultats générés. Le format d'appel de variants (VCF pour *Variant Call Format*) généré par des logiciels comme DRAGEN ou GATK est en effet assez aride et nécessite des compétences informatiques avancées pour être adroitement manipulé et interprété. Et en vertu de sa connaissance plus ou moins fine des environnements cliniques, le comité directeur du Centre – et donc aussi de la plateforme computation – considère qu'il est crucial de veiller à la bonne restitution des résultats, c'est-à-dire à la transmission d'analyses correctes mais épurées, à même d'être examinées par des clinicien·nes potentiellement dépourvus de compétences informatiques afin d'étayer leurs décisions cliniques. Ce deuxième

élément oblige le Centre à imaginer une solution pour *encore davantage traduire* les VCF générés par DRAGEN afin de les rendre mobilisables par les clinicien·nes.

- 38 Début 2022, lorsque j'arrive au Centre pour mener la présente enquête, la solution imaginée par le comité directeur afin de pallier ce problème d'usabilité est d'acquérir la suite logicielle propriétaire Congenica¹⁶, une plateforme intégrée d'aide à l'analyse génomique, qui présente l'avantage d'être certifiée comme dispositif médical par la législation européenne (et donc facilement intégrable à la demande d'extension d'accréditation du Centre). Bien que cette solution présente des avantages indéniables – malgré son coût relativement élevé – elle place DRAGEN dans une position encore plus centrale : puisque les résultats VCF générées par DRAGEN seront ensuite transmis à Congenica pour les rendre davantage intelligibles et utilisables par des clinicien·nes, il devient plus important que jamais de comprendre ce qui se passe à l'intérieur de DRAGEN. En somme, début 2022, avec l'intégration imminente de Congenica pour faciliter l'interprétation des résultats des processus d'appel de variants, DRAGEN devient plus que jamais l'un des fragiles *points de passage obligés* (Callon, 1984) du pipeline du Centre. D'où l'importance, dans le cadre de sa stratégie d'accréditation (et donc de standardisation), de connaître les capacités de DRAGEN dans les moindres détails, c'est-à-dire de savoir où se situent ses forces, mais aussi ses limites.

DRAGEN versus GATK

- 39 Logiciel récent, propriétaire et onéreux – car nécessitant un *hardware* dédié –, DRAGEN constitue une *boîte grise* (Latour, 2012, pp. 207-232) : si ses constituants internes sont plutôt bien documentés, son code source n'est pas accessible et il ne dispose pas – au moment de l'enquête – d'une grande communauté d'utilisateur·rices, contrairement à son principal concurrent, pourtant plus lent, GATK. Comment dès lors s'assurer que les fichiers VCF rapidement produits par DRAGEN soient robustes pour le SREN-USIN et à même d'être inclus dans la demande d'extension d'accréditation au SAS ? En mars 2022, la seule solution qui s'offre au Centre est de réaliser une expérience comparative entre DRAGEN à GATK, ce dernier étant déjà implémenté et bien maîtrisé par les bioinformaticien·nes.
- 40 Pour ce faire, fin mars 2022, Noah et Emily – deux bioinformaticien·nes de la plateforme computation du Centre – conçoivent une expérience comparative. Celle-ci s'appuie sur un trio récemment reçu à titre exploratoire et sur une série de scripts écrits par leurs soins en langage *shell*¹⁷, permettant d'identifier toutes les différences d'appel de variants entre les deux suites logicielles. C'est cette expérience comparative – centrale dans la stratégie du Centre – discutée lors d'une réunion de travail tenue l'après-midi du lundi 28 mars 2022 que nous allons suivre en détail dans la suite de cette section¹⁸.

Initier une comparaison qualitative

- 41 Tous·tes les collaborateur·rice·s de la plateforme computation sont présent·es à cette réunion de travail qui se déroule – comme chaque semaine – dans une salle de séminaire du Centre au sein du parc Biomed. Vue l'importance stratégique du principal sujet à traiter – la comparaison entre DRAGEN et GATK pour préparer l'extension d'accréditation – le responsable de la plateforme séquençage, Liam, est également présent. Comme d'habitude, l'atmosphère est à la fois sérieuse et détendue.

- 42 Au moment d'aborder le sujet intitulé « DRAGEN versus GATK » à l'ordre du jour, Noah projette sur l'écran de la salle un document .html qui servira de fil conducteur à la discussion sur l'expérience comparative qu'il a menée avec Emily :

Noah : On a donc écrit un petit script pour comparer le nombre de variants encodés par DRAGEN et par GATK, avec différents seuils de qualité [fig. 4]¹⁹. Et la première chose qu'on peut voir, c'est que le nombre de variants reste assez proche entre les deux. Mais c'est des chiffres absolus ; on parle pas encore de variants communs.

Figure 4 : Les deux premiers scripts shell écrits par Noah et Emily leur permettant de lister le nombre absolu de variants pour DRAGEN et GATK (avec différents seuils de qualité)

DRAGEN with QUAL threshold = 40

```
# Number of retained variants
project=WGS_████████_Jan22
analysisId=panelApp
cat $project/$project\_variantSelection_$analysisId\.tsv | grep -v "^#" | wc -l
```

```
## 1490
```

GATK with QUAL threshold 40<=QUAL<=3000

```
for qual in 40 100 500 1000 1500 2000 2500 3000; do
  project=WGS_████████_Jan22_GATK_QUAL$qual
  analysisId=panelApp
  nb=$(cat $project/$project\_variantSelection_$analysisId\.tsv | grep -v "^#" | wc -l)
  echo -e "Qual_threshold=$qual\t$nb"
done
```

```
## Qual_threshold=40      1568
## Qual_threshold=100    1553
## Qual_threshold=500    1518
## Qual_threshold=1000   1492
## Qual_threshold=1500   1480
## Qual_threshold=2000   1468
## Qual_threshold=2500   1455
## Qual_threshold=3000   1444
```

- 43 Le script shell conçu par Noah et Emily (fig. 4) offre une première analyse comparative du nombre de variants identifiés par DRAGEN et GATK, selon différents seuils de qualité²⁰. Sans surprise, ajuster le seuil de qualité dans GATK – une option non disponible dans DRAGEN – modifie le nombre de variants détectés, créant un écart avec ceux identifiés par DRAGEN. Comme le souligne Noah, bien que cette première observation soit rassurante – les deux logiciels détectant un nombre globalement similaire de variants –, elle reste incomplète. En effet, elle se limite à un simple décompte absolu, sans préciser si ces variants sont communs aux deux outils. Pour approfondir cette question, Noah introduit un second script :

Noah : Et là vous avez le script qui donne l'évolution des variants communs en fonction des seuils de qualité de GATK [fig. 5]. En gros, ils commencent à diverger plus sérieusement lorsque GATK est réglé sur 2000. C'est pas une surprise car GATK commence à être très sélectif à ce moment-là. Encore une fois, c'est une bonne chose car le nombre de variants communs reste assez élevé jusqu'à là. Par contre, ça nous donne pas d'informations sur les variants uniques.

Figure 5 : Script shell écrit par Noah et Emily (partie supérieure) produisant un graphique montrant l'évolution des variants communs en fonction de différents seuils de qualité GATK (partie inférieure)



- 44 En commentant le graphique de la fig. 5, Noah explique que DRAGEN et GATK commencent à diverger significativement à partir du moment où le seuil de qualité de GATK est fixé à 2000. Avant d'atteindre ce point – c'est-à-dire jusqu'au seuil plutôt élevé de 1500 – les deux programmes semblent globalement alignés, ce qui constitue un signe encourageant, GATK étant la référence la plus fiable et documentée. Cependant, cette information n'est qu'indicative : il demeure essentiel d'examiner le nombre total de variants exclusifs à DRAGEN ou à GATK, ainsi que de confirmer que le seuil de 1500 marque véritablement le point de rupture entre les deux outils :

Noah : Et ici [fig. 6] vous voyez un autre script plus précis qui montre que c'est en fait au seuil 1500 qu'on a les variants les plus communs, mais aussi les variants les moins uniques. Quarante-huit pour ce cas.

Sam : Excuse-moi mais je comprends pas pourquoi la courbe DRAGEN diminue à un certain moment.

Emily : Ça c'est parce que la courbe GATK augmente. Et la troisième [en vert], c'est juste la somme.

Sam : Ah, d'accord, je comprends.

Noah : Et comme on l'a appris la dernière fois avec le SAS, on doit se pencher sur ce qui ne fonctionne pas, mais aussi mettre en évidence ce qui fonctionne. Donc à partir de là, on a décidé de se concentrer sur la qualité de ces quarante-huit variants et sur pourquoi ils sont uniques à chaque logiciel. Et voir ce qui se passe, en gros.

Ethan : Bien.

Liam : Ok.

Figure 6 : Graphique issu du script shell de Noah et d'Emily montrant l'évolution des variants uniques à GATK et à DRAGEN en fonction de différents seuils de qualité de GATK



- 45 La nouvelle inscription (fig. 6) confirme que les deux suites logicielles convergent lorsque GATK est réglé sur le seuil de qualité 1500, avec « seulement » quarante-huit variants uniques (vingt-neuf uniques pour DRAGEN et dix-neuf uniques pour GATK). C'est là qu'intervient une leçon – exprimée par Noah – tirée de la précédente épreuve d'accréditation du Centre pour sa partie « humide » : au cours d'un processus d'accréditation SAS, les candidat·es doivent montrer que les résultats problématiques sont traités de front, *tout en mettant en évidence les résultats opératoires et positifs*. Et c'est en vertu de cette leçon – confirmée par Ethan (responsable de la plateforme computation) et Liam (responsable de la plateforme séquençage) – que Noah et Emily ont décidé de focaliser leur expérience sur ces quarante-huit variants uniques, expressions d'une discordance *a priori* faible, mais néanmoins réelle et potentiellement problématique, entre DRAGEN et GATK.
- 46 À ce stade de l'analyse, une première observation élémentaire mais importante peut être faite. Avant les scripts shell de Noah et d'Emily, DRAGEN et GATK étaient des logiciels qui ne pouvaient être comparés que quantitativement : ils traitaient le même trio à des rythmes très différents et produisaient des résultats – en l'occurrence des variants génomiques – légèrement dissemblables. Cependant, afin d'initier une évaluation à des fins d'accréditation, il apparaît ici nécessaire de générer les conditions d'une comparaison *qualitative*, non seulement axée sur le nombre d'éléments détectés, mais aussi sur les caractéristiques de leurs constituants internes. Ce n'est que lorsque les quarante-huit variants uniques sont connus – en fait, *construits* par des scripts, des graphiques et des décisions plus ou moins stratégiques – qu'il devient possible pour Noah et Emily d'entamer véritablement l'expérience comparative susceptible d'identifier les différences entre les deux suites logicielles.

et Emily est donc de visualiser, via le logiciel IGV, chacun des emplacements des quarante-huit variants uniques de GATK et DRAGEN afin de tenter d'inférer les raisons de leurs encodages respectifs.

- 49 En gardant ces éléments en tête, continuons à suivre la réunion :

Noah : Ce qu'on a remarqué, c'est qu'il y a deux catégories de variants uniques. La première, c'est celle qu'on voit ici [fig. 7]. C'est pas facile à voir, mais il y a beaucoup de points de couleur un peu partout, ce qui veut dire qu'il y a beaucoup de *mismatches* un peu partout. Et c'est un exemple typique de région problématique [à cartographier]. En fait, ici, les *callers* [c'est-à-dire DRAGEN et GATK] font ce qu'ils peuvent ; l'un appelle quelque chose, et l'autre quelque chose d'autre. Mais les variants qui sont appelés sont de toute façon problématiques, vu la région.

Ethan : C'est vrai que c'est une région notoirement problématique, là.

Sam : Ouais, et il y a beaucoup de ces régions peu complexes.

Emily : Du coup c'est pas vraiment une erreur. Il y a pas grand-chose qu'on puisse faire ici.

- 50 Certaines régions du génome sont difficiles à cartographier – c'est-à-dire à décrire en termes de nucléotides – parce qu'elles sont peu complexes – c'est-à-dire constituées de répétitions des mêmes nucléotides (par exemple, 'AAAAAA'). De façon contre-intuitive, ces régions sont sujettes à de nombreuses erreurs de description, car l'absence de points de référence limite les opérations de contrôle et de correction informatiques²³. En conséquence, ces régions – bien connues des bioinformaticien·nes qui travaillent en génomique – ont tendance à conduire à davantage de *mismatches*, qui, par accumulation, peuvent parfois être appelés variants. Pour ce qui intéresse ici le Centre, cette information, bien qu'importante, ne nécessite pas un examen plus approfondi : qu'ils soient encodés uniquement par DRAGEN ou par GATK, ces variants doivent en tous les cas être considérés avec méfiance par l'analyste en charge de la restitution des résultats, car il y a de fortes chances qu'il s'agisse simplement d'artefacts résultant d'erreurs d'alignement dues à la faible qualité des séquences « brutes ». En bref, il s'agit d'un problème classique, commun à l'ensemble de la génomique computationnelle actuelle et sur lequel le Centre ne peut, en l'état, pas faire grand-chose.

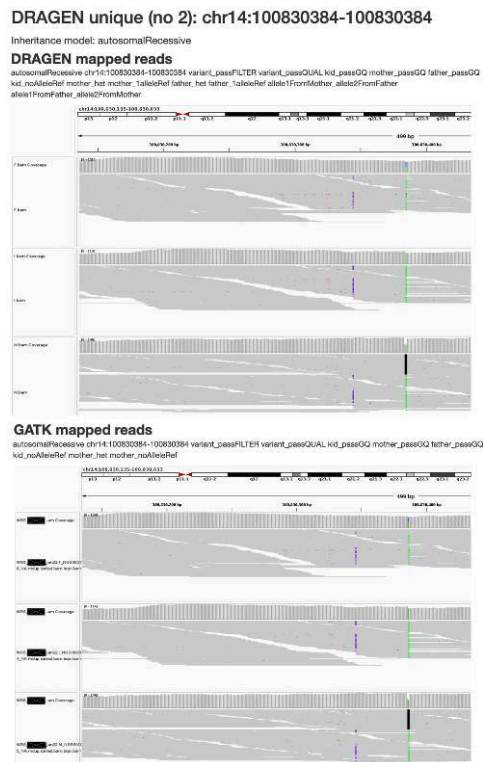
- 51 Jusqu'à présent, l'expérience de comparaison se déroule donc plutôt bien : le nombre de variants uniques est relativement faible et, pour une partie d'entre eux, ces différences sont dues à des problèmes de séquençage « régionaux » bien connus qui ne peuvent être résolus en l'état. Cependant, la deuxième raison évoquée par Noah et Emily s'avérera plus problématique :

Noah : L'autre point qu'on doit discuter est lié à la manière dont le VCF est encodé. Parce qu'il semble qu'il y ait parfois une différence entre GATK et DRAGEN qui pourrait être un peu trompeuse. Vous pouvez voir un exemple ici [fig. 8], mais vous trouverez la liste complète dans le document. Mais donc ici [fig. 8], on regarde un variant autosomal récessif, et on peut voir qu'il y a des substitutions marquées en vert chez le père [première ligne d'IGV], la mère [dernière ligne d'IGV] et l'enfant [ligne centrale d'IGV]. Mais chez la mère, il y a aussi toute une partie qui n'est pas verte, qui est en fait une délétion. Mais DRAGEN ne mentionne pas ça dans le VCF.

Ethan : Et qu'est-ce que dit GATK ?

Noah : C'est là qu'il y a un problème parce que GATK cartographie l'emplacement de la même manière, appelle la substitution mais mentionne aussi la possibilité d'une délétion. C'est pour ça qu'il donne deux variants différents à la même position. Mais je suis d'accord, c'est difficile de voir ça sur cette image.

Figure 8 : Représentation tirée d'IGV qui illustre les *mismatches* (points colorés) appelés sur les nombreux *reads* (couches grises) des positions des gènes 100830100 (tout à gauche) à 100830500 (tout à droite). La position 100830384 (succession horizontale de points verts) est celle où le variant unique à DRAGEN est appelé



- 52 La deuxième raison des différences entre GATK – la norme de référence – et DRAGEN semble plus délicate. Dans le cas examiné par Noah (fig. 8), il semble que DRAGEN considère uniquement la possibilité d'une substitution et ignore la possibilité d'une délétion, ce qui est pourtant évident – pour des bioinformaticien·nes averties – lorsqu'on examine les détails des données BAM via IGV. Cette omission est d'autant plus problématique que le format VCF permet de rapporter les deux possibilités (substitution ou délétion), ce que GATK fait de manière prévisible. Que se passe-t-il dans DRAGEN pour qu'il « oublie » la possibilité de délétion, quand bien même cela pourrait s'avérer problématique dans un contexte de SREN-USIN où chaque variant compte ?²⁴ Difficile de le savoir à ce stade, car la visualisation d'IGV reste à un niveau trop élevé. Il semble donc important de zoomer davantage :

Noah : Mais quand on zoome sur IGV [fig. 9], on voit que la délétion commence en fait à la position précédente, c'est-à-dire la « 383 » [chr14 100830383]. On le voit aussi dans le détail du VCF [fig. 10] où c'est noté « ACC/A ». Mais à la position suivante [chr14 100830384], DRAGEN mentionne seulement la substitution « CA » alors que GATK mentionne aussi la possible délétion « * ». Le problème c'est qu'on a trouvé ça dans vingt cas, et on ne sait pas pourquoi DRAGEN fait ça.

Figure 9 : Zoom IGV sur la position 100830384 (tout à droite) telle que cartographiée par DRAGEN

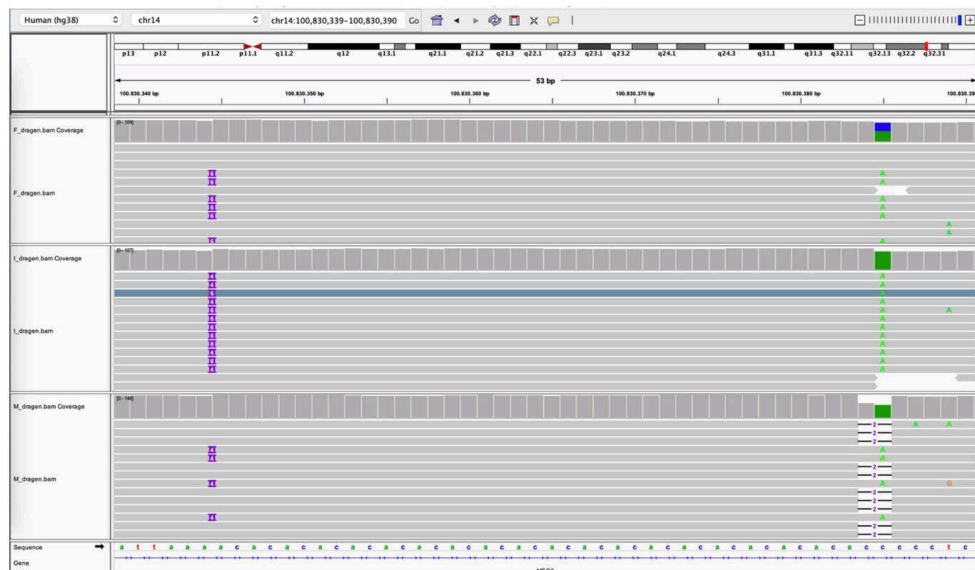


Figure 10 : Extraits de fichiers VCF tels qu'encodés par DRAGEN (moitié supérieure) et GATK (moitié inférieure) entre les positions 100830374 et 100830394. Pour les deux programmes, seules les positions 100830383 à 100830385 sont concernées par cet appel, mais de manière légèrement différente

```

@frt2 ICU_variantSelection]$ bcftools view -H -r chr14:100830374-100830394 /data/UHTS/us
| cut -f1-7,10- | sed 's:/[^\t]*//g'
chr14 100830383 . ACC A 45.23 PASS 0/0 0/0 0/1
chr14 100830384 . C CA 363.69 PASS 0/1 1/1 1/0
chr14 100830385 . C CACA,CACACA 243 PASS 0/2 1/2 1/0
@frt2 ICU_variantSelection]$ #GATK
@frt2 ICU_variantSelection]$ bcftools view -H -r chr14:100830374-100830394 /scratch/perm
8631309.genotyped.vcf.gz | cut -f1-7,10- | sed 's:/[^\t]*//g'
chr14 100830383 rs1491366505 ACC A 944.25 PASS 0/0 0/0 0/1
chr14 100830384 rs373934045 C CA,* 5481.75 PASS 0/1 1/1 1/2
chr14 100830385 . C CACACA,CACA,* 5196.29 PASS 0/1 1/2 2/3
@frt2 ICU_variantSelection]$

```

- 53 Le problème se précise, tout en restant ouvert. En effet, en zoomant dans IGV, il apparaît que la délétion potentielle se manifeste – pour certains *reads* – à la position « 100830383 » (fig. 9). En examinant ensuite le fichier VCF généré par DRAGEN (fig. 10, partie supérieure), seule la substitution de « C » – la référence – par « A » est rapportée. En revanche, GATK (fig. 10, partie inférieure) rapporte non seulement cette substitution, mais signale également la potentielle délétion, représentée par un « * » à la position « 100830384 ». C'est là, semble-t-il, une divergence notable qui mérite d'être examinée plus en détail.
- 54 À ce stade plus avancé de l'analyse, une seconde observation peut être formulée. Grâce à un dispositif expérimental combinant l'examen détaillé de fichiers VCF et la visualisation précise de fichiers BAM via IGV, Noah et Emily parviennent à identifier un phénomène problématique : dans de rares, mais réelles, conditions, DRAGEN échoue à encoder les délétions là où GATK y parvient. En conséquence, vingt variants spécifiques à DRAGEN présentent une *qualité inférieure* à ceux de GATK, car ils omettent l'intégration des délétions potentielles en complément des substitutions probables. Un problème est ainsi *réalisé* – au sens de rendu réel – tout en étant *marginalisé* – au sens de mis en relation avec de très nombreux alignements. Ce problème existait-il avant cette

analyse comparative ? Très probablement, mais il demeurerait latent et inaperçu. L'expérience, en révélant son existence, offre l'occasion de mieux le caractériser.

Élaborer une opportunité

- 55 Quelle est donc la raison pour laquelle DRAGEN omet de signaler d'éventuelles délétions, expliquant ainsi la présence d'une proportion significative – bien que modeste en nombre absolu – de variants uniques (vingt sur quarante-huit) ?

Noah : En fait, on dirait que lorsqu'une délétion se produit à une position antérieure à une substitution, DRAGEN l'inclut pas dans le VCF.

Ethan : C'est peut-être un bug ?

Sam : Peut-être. Mais ce qui est bizarre, c'est qu'on dirait que c'est un problème dans la façon dont DRAGEN encode le VCF. Parce que le *mapping* semble être le même que celui de GATK.

Noah : C'est aussi ce qu'on pense, oui.

Ethan : Mais alors est-ce qu'il y aurait un moyen de demander à DRAGEN de prendre en compte la position antérieure avant l'encodage ?

Emily : Certainement, c'est imaginable. Mais il faudrait creuser beaucoup plus pour mieux cadrer et spécifier le problème. Par exemple, est-ce qu'on parle d'une seule position antérieure ou de plusieurs ? Et est-ce que ça se passe seulement dans des régions spécifiques ?

- 56 La raison exacte pour laquelle DRAGEN ne prend pas en compte les potentielles délétions qui commencent à des positions antérieures aux substitutions semble difficile à isoler. Il pourrait s'agir d'un bug ou d'un mauvais paramétrage. Mais serait-il possible de mettre en place une solution *ad hoc* – par exemple au moyen d'un script dédié – pour demander à DRAGEN de prendre en compte ces cas de délétions possibles ? Pour Emily, qui a participé à l'installation et au réglage de DRAGEN, cela semble faisable, pour autant que le problème soit bien mieux défini, ce qui nécessiterait des tests et des expériences supplémentaires. Et est-ce vraiment utile ?

Ethan : Bon, c'est important de se rappeler ce dont nous avons parlé la dernière fois : on nous demande de connaître nos limites.

Liam : Voilà, exact. Ils [les SAS] s'attendent pas à ce que nous soyons parfaits. Ils attendent de nous des résultats convaincants, mais aussi qu'on souligne les imperfections de nos résultats. Et à partir du moment où on montre qu'on a aussi des idées pour améliorer les choses, c'est considéré comme positif.

Ethan : Et c'est exactement ce qu'on a ici, au final. On a des résultats similaires pour la grande majorité des cas, avec quelques exceptions sur lesquelles on peut travailler, parce qu'on a les compétences pour. En fait, c'est presque une bonne chose pour la suite, parce que ça nous donne un axe d'amélioration.

...

Liam : Oui. Et c'est aussi ça qu'ils [le SAS] recherchent, parce que l'ISO leur donne pas grand-chose en termes de geekerie.

Ethan : On a de la marge du coup, mais il faut bien choisir quoi dire.

- 57 Le compromis proposé par Ethan – fondé sur les échanges au sein de l'équipe – consiste à s'en tenir là pour le moment, en intégrant l'expérience comparative et ses résultats dans la demande d'extension d'accréditation. Ce que l'on pourrait qualifier de « problème des délétions de DRAGEN » affecte moins de 2 % des variants, et le Centre dispose des ressources et des compétences nécessaires pour y remédier à terme. Sur le plan stratégique, cette décision semble judicieuse : il est en effet crucial pour le Centre de se projeter dès maintenant au-delà de l'extension d'accréditation, en tenant compte notamment des audits bisannuels qui vérifieront la conformité du pipeline tout en

évaluant les améliorations réalisées ou non. Et ce problème de DRAGEN, précédemment rendu réel et marginal, pourrait même constituer une prise pour travailler au maintien de l'accréditation (si elle est accordée à la partie sèche du Centre).

- 58 Il est intéressant de constater que l'expérience acquise lors de la précédente procédure d'accréditation se révèle ici précieuse pour transformer ce qui pourrait sembler être un problème – celui des délétions de DRAGEN – en une *opportunité*. En effet, selon Ethan et Liam – qui ont tous deux participé à l'épreuve d'accréditation de la partie humide du Centre – une idée reçue fréquente consiste à percevoir une demande d'accréditation comme une exigence de résultats parfaits. En réalité, la démarche est plus subtile (et intéressante) : le SAS, comme sûrement d'autres organismes d'accréditation, accorde davantage d'importance aux moyens mobilisés pour surveiller et améliorer la fiabilité de résultats perfectibles dans la durée. Ce retour d'expérience, hérité de l'accréditation précédente, permet au Centre de renverser la perspective. Ce qui émerge de l'expérience comparative menée par Noah et Emily apparaît moins comme un obstacle délicat que comme une opportunité stratégique : en adoptant une vision plus large, permise par l'expérience antérieure, le problème des délétions devient une preuve de la capacité du Centre à identifier les limites subtiles de son pipeline, tout en esquissant des pistes concrètes d'amélioration, démarche typiquement perçue de manière positive par le SAS.
- 59 C'est même pour le Centre d'autant plus important de suggérer des pistes concrètes d'amélioration que le standard ISO 15189 – celui qui sert de base à l'accréditation SAS – reste passablement allusif en ce qui concerne les critères qualité des éléments computationnels des laboratoires biomédicaux, comme le souligne Liam en fin d'extrait (« l'ISO leur donne pas grand-chose »). La stratégie consistant à être généreux envers le SAS en leur présentant un problème bien construit, car améliorable, semble ainsi prometteuse, sans pour autant offrir aucune garantie, comme le souligne immédiatement Ethan (« mais il faut bien choisir quoi dire »).
- 60 En définitive, le suivi attentif de cette expérience comparative semble indiquer qu'un enjeu – certainement parmi d'autres – de la préparation à une épreuve d'accréditation en génomique computationnelle peut résider dans la construction adéquate de problèmes qualitatifs suffisamment importants pour être identifiés et évalués, mais aussi suffisamment mineurs (voire *minorés*) pour permettre des résolutions à court terme. Plutôt qu'un exercice de démonstration (Rosental, 2009), il s'agit d'un exercice visant à mettre en avant des limites surmontables, témoignant ainsi d'un potentiel d'amélioration continue.
- 61 En ce sens, loin d'être une procédure mécanique ou purement administrative, la préparation à une épreuve d'accréditation en génomique computationnelle s'apparente à un subtil travail créatif qui implique de mettre en évidence les limites de son système en vertu des améliorations ultérieures qui pourraient y être apportées. Une partie de ce que l'on pourrait appeler l'art de l'accréditation, et peut-être des processus de standardisation qualité en général, semble résider dans cette capacité à façonner et à valoriser des problèmes qui sont à la fois réels, marginaux et corrigeables. Ce travail d'ajustement, de recadrage et de reformulation des problèmes fait écho à ce Joan Fujimura (1987) nomme la *doability*, entendue comme la capacité à rendre un projet faisable dans un contexte sociotechnique donné. L'accréditation devient ainsi non seulement une épreuve de conformité, mais aussi un moment où l'on reconfigure

stratégiquement les conditions de faisabilité d'un système en le rendant à la fois problématisable et améliorable.

Accélération et influence

- 62 La séance s'achève ainsi sur le compromis proposé par Ethan : l'expérience préparatoire, moyennant quelques ajustements mineurs, pourra être intégrée presque telle quelle dans le dossier de demande d'accréditation. Elle constituera ainsi l'élément central – certes imparfait, mais c'est bien l'un des objectifs – du chapitre consacré à l'appel de variants. Cela signifie-t-il que cette réunion a représenté l'étape clé du processus d'accréditation de la plateforme computationnelle ? Non, car l'accréditation – et donc l'épreuve par laquelle elle passe et qui requiert tant de préparation – concernait également d'autres éléments importants, notamment en lien avec la protection des données et la cybersécurité, qui eux-mêmes ont nécessité d'importants efforts préparatoires (non documentés ici). Mais en un sens oui, car parmi les chapitres essentiels à l'élaboration de ce rapport préparatoire d'accréditation – couvrant la cybersécurité, la traçabilité, la protection des données, etc. – la question de la qualité de l'appel rapide de variants occupait une place centrale au sein des efforts semi-commerciaux du Centre (cf. *supra*), tout en étant sans conteste la moins bien définie au sein du standard ISO, et donc aussi des critères d'évaluation du SAS. En d'autres termes, le chapitre, nécessaire, dédié à l'appel de variants était à la fois celui qui disposait du moins de lignes directrices et celui qui jouait l'un des premiers rôles dans la stratégie future du Centre, et cette contradiction même représentait un défi théorique et pratique, finalement surmonté en situation lors de cette séance.
- 63 Mais tout cela a-t-il fonctionné ? Les efforts préparatoires des bio-informaticien·nes de la plateforme computation afin de patiemment définir les atouts et les limites de DRAGEN pour l'appel de variants ont-ils porté leurs fruits ? Absolument, à tel point que l'ensemble de la procédure d'extension d'accréditation a même été accélérée, afin notamment de permettre à certaines membres de la plateforme computation de participer à des audits d'autres centres suisses de génomiques. En somme, les nouveautés produites par le Centre en termes de contrôle qualité, et notamment en ce qui concerne l'appel de variants, ont fini par être intégrées dans les critères du SAS, notamment pour aider à spécifier et renforcer la norme ISO 15189 quant à ces aspects, très techniques mais aussi très communs, de génomiques computationnels.
- 64 Si les données de la présente enquête ne permettent hélas pas d'assigner des motivations précises du côté du SAS, elles permettent tout de même de constater qu'une accélération s'est effectivement produite, puisque le SAS s'est finalement empressé d'étendre l'accréditation du Centre à sa partie sèche²⁵. Mais aussi, en faisant maintenant partie « physiquement » des critères d'évaluation du SAS, le Centre a accru son influence au sein de l'écosystème suisse de la génomique. En somme, le Centre est parvenu à articuler le standard qualité ISO 15189 aux aspérités de ses équipements computationnels, anticipant avec justesse les attentes du SAS et sa marge de manœuvre sur les questions bio-informatiques de pointe.

Discussion et conclusion

- 65 De façon relativement surprenante, peu a été dit dans la littérature socio-anthropologique sur les épreuves situées de standardisation dans le domaine de la bio-informatique, quand bien même ces épreuves constituent autant de points de passage obligés pour quiconque souhaite développer une médecine clinique ‘personnalisée’ ou ‘de précision’, elle-même objet de nombreux discours utopiques et dystopiques (e.g., Roth & Bruni, 2021 ; Topol, 2019 ; Zuboff, 2019, chap. 8). Peut-être est-ce dû à la posture ethnographique que requiert ce genre d’analyse, qui implique d’entrer en contact étroit avec des bio-informaticien·nes pour observer – de manière assez intrusive, il faut bien le reconnaître – leurs pratiques quotidiennes, notamment celles consistant à se préparer en vue d’examens coûteux et souvent couperets ? Dans la perspective d’une sociologie de la préparation aux épreuves de standardisation en bio-informatique, cet article entend proposer quelques bases fragiles mais réelles pour une telle exploration. En guise de conclusion, au terme de cette plongée au cœur de la bio-informatique en action, je souhaiterais discuter trois observations.
- 66 La première observation confirme un phénomène déjà documenté dans la littérature : l’inévitable besoin de « bricoler » autour du standard (Timmermans & Epstein, 2010 ; Lampland & Star, 2008). Cela est particulièrement manifeste dans le cas du Centre, qui n’a d’autre choix que d’adopter une posture active vis-à-vis du standard ISO 15189, dont la conformité est inspectée par le SAS. En effet, ce standard ne précise pas explicitement les critères de qualité applicables aux éléments bio-informatiques de pointe liés à l’appel de variants. Ce flou, hérité d’une époque antérieure aux avancées de la génomique computationnelle, oblige le Centre à modeler le standard tout en s’y conformant – un travail d’artisan autant qu’un exercice de rigueur. Cette apparente contradiction, résolue dans la pratique, fait écho à l’observation de Timmermans et Epstein : « pour coordonner des intérêts et des activités divers, les standards délèguent nécessairement une partie du travail résiduel, qui exige une participation et soumission active des personnes aux directives du standard » (2010, p. 81, ma traduction). Ainsi, loin d’être seulement une contrainte réglementaire, l’appropriation du standard par le Centre illustre la dynamique plus large d’adaptation et de négociation propre à la standardisation en contexte technoscientifique et biomédical.
- 67 La deuxième observation met en lumière l’ambivalence inhérente aux processus de standardisation. Ceux-ci ne se résument pas à des directives imposées d’en haut, exerçant une domination sur les praticien·nes, comme on pourrait intuitivement le penser. Ils peuvent aussi être saisis d’en bas, offrant aux acteur·rices une véritable opportunité stratégique, ce que Timmermans et Almeling (2009) avaient par ailleurs déjà souligné dans le contexte des soins hospitaliers. Dans le cas du Centre, la dimension stratégique du processus d’accréditation apparaît clairement, du moins lorsqu’on l’examine de près. Elle lui permet en effet, au moins théoriquement, d’intégrer un marché dont les retombées économiques semi-commerciales sont cruciales pour sa pérennisation et son développement. Cette réalité du secteur de la génomique computationnelle en Suisse invite même à quelque peu nuancer l’argument avancé par Webster et Erikson (2008) sur la gouvernance par les standards. En effet, s’il existe bien une forme de gouvernance, elle ne découle pas tant ici des organismes de standardisation – ISO ou SAS, qui restent relativement flous sur les aspects computationnels des laboratoires médicaux en Suisse – que de certains acteurs du

secteur, comme le Centre lui-même, qui s'appuient sur un standard et sur le label qualité associé à son accréditation pour asseoir leur position dans un environnement concurrentiel. L'extension de l'accréditation du Centre à ses analyses génomiques de trio en est une illustration frappante, et réussie : elle parvient à placer le Centre en position de référence en établissant un précédent. Désormais, tout laboratoire souhaitant obtenir une accréditation pour des analyses génomiques devra s'inspirer du travail accompli par le Centre (et donc, *in fine*, de ses instruments), qui servira de modèle initial conservé par le SAS.

- 68 Enfin, sur un plan davantage anthropologique, cette enquête enrichit également l'étude sociale sur la préparation (Riom & Tanferri, 2024), en documentant un mode d'action préparatoire particulier, qui s'articule autour de trois moments : un moment proprement épistémique, un moment de critique négative et un moment d'orientation positive. Si l'on reprend rapidement le fil de la séance préparatoire du Centre, le rôle des bioinformaticien·nes Noah et Emily consistait au départ essentiellement à mettre en place une expérience comparative afin de mieux comprendre le comportement de la nouvelle suite logicielle propriétaire DRAGEN, telle qu'intégrée au pipeline du Centre. Les pratiques mises en œuvre à ce stade préliminaire – obtenir un trio d'échantillons, l'analyser à la fois avec DRAGEN et GATK (la référence, plus lente, du domaine), puis rédiger plusieurs scripts shell pour comparer leurs résultats – peuvent ainsi être qualifiées d'*épistémiques*, dans la mesure où elles visaient à éclairer les rapports constitutifs de DRAGEN. Cependant, comme nous l'avons vu, une fois établie l'expérience visant à informer le comportement de DRAGEN, il a fallu mettre entre parenthèses les résultats positifs pour mieux se concentrer sur les résultats problématiques. En l'occurrence, si le nombre élevé de variants communs entre DRAGEN et GATK a été un soulagement pour le Centre, ce qui importait vraiment étaient les quelques dizaines de variants uniques, expressions de discordances marginales mais réelles. L'attention portée aux dysfonctionnements apparaît ici comme un enjeu central : peu importe que le dispositif fonctionne dans la grande majorité des cas, ce qui compte réellement, c'est la capacité à identifier les anomalies et les défaillances. Ici, les bioinformaticien·nes semblent devoir opérer un basculement et passer d'un mode épistémique initial – sans doute indispensable – à un mode que je qualifie de critique négative, tout aussi nécessaire mais sans doute moins intuitif, où l'attention se porte majoritairement sur les limites et les dysfonctionnements. Mais le seul fait de signaler les dysfonctionnements ne suffirait pas à répondre aux exigences d'une préparation à une épreuve de standardisation en génomique computationnelle. Au vu des éléments présentés, il est encore important de parvenir à présenter ce qui résiste comme quelque chose qui peut être amélioré et, idéalement, résolu. Et c'est précisément là qu'il semble nécessaire d'adopter une approche positive, qui vise à tirer des avantages de la situation problématique attestée. Dans le cas présenté ici, DRAGEN a en effet éprouvé quelques difficultés à inclure dans ses VCF la possibilité de délétions lorsqu'elles précédaient et chevauchaient d'éventuelles substitutions. Mais les bioinformaticien·nes du Centre, bien aidés par l'expérience accumulée lors de la précédente démarche d'accréditation, parviennent à retourner la situation à leur avantage en la présentant comme une opportunité de rendre leur pipeline encore plus performant, dans leur quête désormais perpétuelle – à la demande du SAS – d'amélioration processuelle. Il s'avère donc qu'une partie de ce que nous pouvons nommer l'art des processus de standardisation réside, peut-être et parfois, dans cette

capacité à façonner et à valoriser des problèmes *doable* (Fujimura, 1987) qui sont à la fois réels, problématiques, marginaux et corrigeables.

- 69 Ces trois éléments – le bricolage des standards, leur ambivalence, ainsi que les modes d'action propres aux préparations d'épreuves – constituent des résultats encore provisoires, issus d'une étude de cas menée dans un contexte national particulier. D'autres enquêtes restent évidemment nécessaires pour mieux comprendre les ressorts de la préparation aux épreuves de standardisation, en bio-informatique et ailleurs. Mais la présente analyse laisse entrevoir, je l'espère, quelques aspects concrets, quoique subreptices, des liens étroits entre la bio-informatique et les processus de standardisation, au-delà des abstractions qui nourrissent les discours utopiques comme dystopiques sur l'avènement de la médecine dite personnalisée ou de précision.

Ce travail a été soutenu par le fonds de recherche SNSF Sinergia 180350, « Development of Personalized Health in Switzerland: Social Sciences Perspectives ». Je souhaite exprimer ma profonde reconnaissance aux membres du centre de génomique computationnelle qui m'ont permis de suivre leurs activités et, plus généralement, m'ont aidé dans mes recherches. Ma gratitude va également à Geoffrey Bowker, Nolwenn Bühler, Luca Chiapperino, Nils Graber, Francesco Panese, Dominique Vinck, les évaluateur-rices anonymes et les éditrices de cette revue pour leur soutien, leur aide et leurs suggestions pertinentes. Toutes les erreurs et approximations éventuelles demeurent, bien sûr, de ma seule responsabilité.

BIBLIOGRAPHIE

- Australian Genomics Health Alliance Acute Care Flagship. (2020). Feasibility of Ultra-Rapid Exome Sequencing in Critically Ill Infants and Children With Suspected Monogenic Conditions in the Australian Public Health Care System. *JAMA*, 323(24), 2503-2511. <https://doi.org/10.1001/jama.2020.7671>
- Bowker, G. C. & Star, S. L. (2000). *Sorting Things Out: Classification and Its Consequences*. Cambridge, Mass. : The MIT Press.
- Busch, L. (2013). *Standards: Recipes for Reality*. Cambridge, Mass. : MIT Press.
- Callon, M. (1984). Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Briec Bay. *The Sociological Review*, 32(1), 196-233. <https://doi.org/10.1111/j.1467-954X.1984.tb00113.x>
- Callon, M. (2017). *L'emprise des marchés*. Paris : La Découverte.
- Callon, M., Cohendet, P., Curien, N. & Dalle, J.-M. (1999). *Réseau et coordination*. Paris : Economica.
- Cambrosio, A. & Keating, P. (1992). A Matter of FACS: Constituting Novel Entities in Immunology. *Medical Anthropology Quarterly*, 6(4), 362-384. <https://doi.org/10.1525/maq.1992.6.4.02a00040>
- Cambrosio, A., Keating, P., Schlich, T. & Weisz, G. (2006). Regulatory objectivity and the generation and management of evidence in medicine. *Social Science & Medicine*, 63(1), 189-199. <https://doi.org/10.1016/j.socscimed.2005.12.007>

- Casper, M. J. & Clarke, A. E. (1998). Making the Pap Smear into the 'Right Tool' for the Job: Cervical Cancer Screening in the USA, circa 1940-95. *Social Studies of Science*, 28(2), 255-290. <https://doi.org/10.1177/030631298028002003>
- Chiapperino, L., Graber, N. & Panese, F. (2024). A precision immuno-oncology turn? Hybridizing cancer genomics and immunotherapy through neoantigens-based adoptive cell therapies. *Social Studies of Science*, 03063127241303720. <https://doi.org/10.1177/03063127241303720>
- Clark, M. M., Hildreth, A., Batalov, S., Ding, Y., Chowdhury, S., Watkins, K., Ellsworth, K., Camp, B., Kint, C. I., Yacoubian, C., Farnaes, L., Bainbridge, M. N., Beebe, C., Braun, J. J. A., Bray, M., Carroll, J., Cakici, J. A., Caylor, S. A., Clarke, C., ... Kingsmore, S. F. (2019). Diagnosis of genetic diseases in seriously ill children by rapid whole-genome sequencing and automated phenotyping and interpretation. *Science Translational Medicine*, 11(489), eaat6177. <https://doi.org/10.1126/scitranslmed.aat6177>
- Conley, T. (1998). Mapping in the Folds: Deleuze « Cartographe ». *Discourse*, 20(3), 123-138. <https://www.jstor.org/stable/41389502> (consulté le 10/07/2025)
- Dear, P. & Jasanoff, S. (2010). Dismantling boundaries in science and technology studies. *Isis; an International Review Devoted to the History of Science and Its Cultural Influences*, 101(4), 759-774.
- Desrosières, A. (2000). *La politique des grands nombres. Histoire de la raison statistique*. Paris : La Découverte.
- Dimmock, D., Caylor, S., Waldman, B., Benson, W., Ashburner, C., Carmichael, J. L., Carroll, J., Cham, E., Chowdhury, S., Cleary, J., D'Harlingue, A., Doshi, A., Ellsworth, K., Galarreta, C. I., Hobbs, C., Houtchens, K., Hunt, J., Joe, P., Joseph, M., ... Farnaes, L. (2021). Project Baby Bear: Rapid precision care incorporating rWGS in 5 California children's hospitals demonstrates improved clinical outcomes and reduced costs of care. *The American Journal of Human Genetics*, 108(7), 1231-1238. <https://doi.org/10.1016/j.ajhg.2021.05.008>
- Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C. & Borgman, C. L. (2011). Science friction: Data, metadata, and collaboration. *Social Studies of Science*, 41(5), 667-690. <https://doi.org/10.1177/0306312711413314>
- Farnaes, L., Hildreth, A., Sweeney, N. M., Clark, M. M., Chowdhury, S., Nahas, S., Cakici, J. A., Benson, W., Kaplan, R. H., Kronick, R., Bainbridge, M. N., Friedman, J., Gold, J. J., Ding, Y., Veeraraghavan, N., Dimmock, D. & Kingsmore, S. F. (2018). Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. *npj Genomic Medicine*, 3(1), 1-8. <https://doi.org/10.1038/s41525-018-0049-4>
- Foucault, M. (1963). *Naissance de la clinique. Une archéologie du regard médical*. Paris : Presses universitaires de France.
- Foucault, M. (1975). *Surveiller et punir : Naissance de la prison* (French Language édition). Paris : Gallimard.
- Fujimura, J. H. (1987). Constructing 'Do-able' Problems in Cancer Research: Articulating Alignment. *Social Studies of Science*, 17(2), 257-293. <https://doi.org/10.1177/030631287017002003>
- GATK Team. (2024, 25 juin). *Phred-scaled quality scores*. GATK. <https://gatk.broadinstitute.org/hc/en-us/articles/360035531872-Phred-scaled-quality-scores> (consulté le 10/07/2025)
- Hà, T.-D. & Chow-White, P. A. (2021). The cancer multiple: Producing and translating genomic big data into oncology care. *Big Data & Society*, 8(1), 2053951720978991. <https://doi.org/10.1177/2053951720978991>

- Jarvis, E. D., Formenti, G., Rhie, A., Guarracino, A., Yang, C., Wood, J., Tracey, A., Thibaud-Nissen, F., Vollger, M. R., Porubsky, D., Cheng, H., Asri, M., Logsdon, G. A., Carnevali, P., Chaisson, M. J. P., Chin, C.-S., Cody, S., Collins, J., Ebert, P., ... Miga, K. H. (2022). Semi-automated assembly of high-quality diploid human reference genomes. *Nature*, 611(7936), 519-531. <https://doi.org/10.1038/s41586-022-05325-5>
- Jaton, F. (2017). We get the algorithms of our ground truths: Designing referential databases in digital image processing. *Social Studies of Science*, 47(6), 811-840. <https://doi.org/10.1177/0306312717730428>
- Jaton, F. (2019). « Pardonnez cette platitude » : de l'intérêt des ethnographies de laboratoire pour l'étude des processus algorithmiques. *Zilsel*, 5(1), 315-339. <https://www.cairn.info/revue-zilsel-2019-1-page-315.htm> (consulté le 10/07/2025)
- Jaton, F. (2021). Assessing biases, relaxing moralism: On ground-truthing practices in machine learning design and application. *Big Data & Society*, 8(1), 20539517211013569. <https://doi.org/10.1177/20539517211013569>
- Jaton, F. (2021). *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating*. Cambridge, Mass. : MIT Press.
- Jaton, F. (2022). Éléments pour une sociologie de l'activité de programmation. *RESET. Recherches en sciences sociales sur Internet*, 11, 1-32. <https://doi.org/10.4000/reset.3829>
- Jaton, F. (2023). Groundwork for AI: Enforcing a benchmark for neoantigen prediction in personalized cancer immunotherapy. *Social Studies of Science*, 53(5), 787-810. <https://doi.org/10.1177/03063127231192857>
- Jaton, F. & Vinck, D. (2016). Processus frictionnels de mises en bases de données. *Revue d'anthropologie des connaissances*, 10(4), 489-504. <https://doi.org/10.3917/rac.033.0489>
- Jaton, F. & Vinck, D. (2023). Politicizing Algorithms by Other Means: Toward Inquiries for Affective Dissensions. *Perspectives on Science*, 31(1), 84-118. https://doi.org/10.1162/posc_a_00582
- Kingsmore, S. F., Cakici, J. A., Clark, M. M., Gaughran, M., Feddock, M., Batalov, S., Bainbridge, M. N., Carroll, J., Caylor, S. A., Clarke, C., Ding, Y., Ellsworth, K., Farnaes, L., Hildreth, A., Hobbs, C., James, K., Kint, C. I., Lenberg, J., Nahas, S., ... Yamada, C. (2019). A Randomized, Controlled Trial of the Analytic and Diagnostic Performance of Singleton and Trio, Rapid Genome and Exome Sequencing in Ill Infants. *The American Journal of Human Genetics*, 105(4), 719-733. <https://doi.org/10.1016/j.ajhg.2019.08.009>
- Kohler, R. E. (1994). *Lords of the Fly: Drosophila Genetics and the Experimental Life*. Chicago : University of Chicago Press.
- Lampland, M. & Star, S. L. (dir.). (2008). *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Ithaca : Cornell University Press.
- Latour, B. (1987). *Science in Action – How to Follow Scientists & Engineers Through Society*. Cambridge, Mass. : Harvard University Press.
- Latour, B. (2012). *Enquête sur les modes d'existence : Une anthropologie des Modernes*. Paris : La Découverte.
- Lefebvre, M. (2013). L'infra-ordinaire de la recherche. Écritures scientifiques personnelles, archives et mémoire de la recherche. *Sciences de la société*, 89, 3-17. <https://doi.org/10.4000/sds.203>

- Loi fédérale sur l'analyse génétique humaine, 810.12, CH, (2004). Disponible à <https://www.fedlex.admin.ch/eli/cc/2007/131/fr> (consulté le 08/10/2023)
- Loi fédérale sur la procréation médicalement assistée, 810.11, CH, (1998). Disponible à <https://www.fedlex.admin.ch/eli/cc/2000/554/fr> (consulté le 08/10/2023)
- Mallard, A. (1998). Compare, Standardize and Settle Agreement: On some Usual Metrological Problems. *Social Studies of Science*, 28(4), 571-601. <https://doi.org/10.1177/030631298028004003>
- Mitchell, K., Brito, J. J., Mandric, I., Wu, Q., Knyazev, S., Chang, S., Martin, L. S., Karlsberg, A., Gerasimov, E., Littman, R., Hill, B. L., Wu, N. C., Yang, H. T., Hsieh, K., Chen, L., Littman, E., Shabani, T., Enik, G., Yao, D., ... Mangul, S. (2020). Benchmarking of computational error-correction methods for next-generation sequencing data. *Genome Biology*, 21(1), 71. <https://doi.org/10.1186/s13059-020-01988-3>
- Nelson, N. C., Keating, P. & Cambrosio, A. (2013). On being “actionable”: clinical sequencing and the emerging contours of a regime of genomic medicine in oncology. *New Genetics and Society*, 32(4), 405-428. <https://doi.org/10.1080/14636778.2013.852010>
- O’Connell, J. (1993). Metrology: The Creation of Universality by the Circulation of Particulars. *Social Studies of Science*, 23(1), 129-173. <https://doi.org/10.1177/030631293023001005>
- Owen, M. J., Lefebvre, S., Hansen, C., Kunard, C. M., Dimmock, D. P., Smith, L. D., Scharer, G., Mardach, R., Willis, M. J., Feigenbaum, A., Niemi, A.-K., Ding, Y., Van Der Kraan, L., Ellsworth, K., Guidugli, L., Lajoie, B. R., McPhail, T. K., Mehtalia, S. S., Chau, K. K., ... Kingsmore, S. F. (2022). An automated 13.5 hour system for scalable diagnosis and acute management guidance for genetic diseases. *Nature Communications*, 13(1), 4057. <https://doi.org/10.1038/s41467-022-31446-6>
- Peerbaye, A. & Vinck, D. (2023). Les ethnographies de laboratoire : par-delà un mythe fondateur des STS. Dans S. Boudia & A. Peerbaye (dir.) *Sciences et techniques en sociétés* (p. 35-58). Paris : ISTE Éditions.
- Petrikin, J. E., Willig, L. K., Smith, L. D. & Kingsmore, S. F. (2015). Rapid whole genome sequencing and precision neonatology. *Seminars in Perinatology*, 39(8), 623-631. <https://doi.org/10.1053/j.semperi.2015.09.009>
- Pfeifer, S. P. (2017). From next-generation resequencing reads to a high-quality variant data set. *Heredity*, 118(2), 111-124. <https://doi.org/10.1038/hdy.2016.102>
- Riom, L. & Tanferri, M. (2024). Sur la piste des préparations. *Revue d'anthropologie des connaissances*, 18(2). <https://doi.org/10.4000/11r9q>
- Rip, A. (1986). Controversies as Informal Technology Assessment. *Knowledge*, 8(2), 349-371. <https://doi.org/10.1177/107554708600800216>
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29(1), 24-26. <https://doi.org/10.1038/nbt.1754>
- Rosental, C. (2009). Anthropologie de la démonstration. *Revue d'anthropologie des connaissances*, 32(2), 233-252. <https://doi.org/10.3917/rac.007.0233>
- Roth, P. H. & Bruni, T. (2021). Participation, Empowerment, and Evidence in the Current Discourse on Personalized Medicine: A Critique of “Democratizing Healthcare”. *Science, Technology, & Human Values*, 01622439211023568. <https://doi.org/10.1177/01622439211023568>
- Stevens, H. (2013). *Life Out of Sequence: A Data-Driven History of Bioinformatics*. Chicago : University of Chicago Press.

- Stevens, H. (2016). Hadooping the genome: The impact of big data tools on biology. *BioSocieties*, 11(3), 352-371. <https://doi.org/10.1057/s41292-016-0003-6>
- Thévenot, L. (2009). Postscript to the Special Issue: Governing Life by Standards: A View from Engagements. *Social Studies of Science*, 39(5), 793-813. <https://doi.org/10.1177/0306312709338767>
- Timmermans, S. (2015). Trust in standards: Transitioning clinical exome sequencing from bench to bedside. *Social Studies of Science*, 45(1), 77-99. <https://doi.org/10.1177/0306312714559323>
- Timmermans, S. & Almeling, R. (2009). Objectification, standardization, and commodification in health care: A conceptual readjustment. *Social Science & Medicine*, 69(1), 21-27. <https://doi.org/10.1016/j.socscimed.2009.04.020>
- Timmermans, S. & Berg, M. (1997). Standardization in Action: Achieving Local Universality through Medical Protocols. *Social Studies of Science*, 27(2), 273-305. <https://doi.org/10.1177/030631297027002003>
- Timmermans, S. & Berg, M. (2003). *The Gold Standard: The Challenge Of Evidence-Based Medicine*. Philadelphia, Pa : Temple University Press.
- Timmermans, S. & Epstein, S. (2010). A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology*, 36(1), 69-89. <https://doi.org/10.1146/annurev.soc.012809.102629>
- Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York, NY : Basic Books.
- Turnbull, C., Scott, R. H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F. B., Halai, D., Baple, E., Craig, C., Hamblin, A., Henderson, S., Patch, C., O'Neill, A., Devereau, A., Smith, K., Martin, A. R., Sosinsky, A., McDonagh, E. M., Sultana, R., ... Caulfield, M. J. (2018). The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. *BMJ*, 361, k1687. <https://doi.org/10.1136/bmj.k1687>
- Vinck, D., Camus, A., Jaton, F. & Oberhauser, P.-N. (2018). Localités distribuées, globalités localisées : Actions, actants et médiations au service de l'ethnographie du numérique. *Symposium*, 22(1), 41-60.
- Wagner, J., Olson, N. D., Harris, L., McDaniel, J., Cheng, H., Fungtammasan, A., Hwang, Y.-C., Gupta, R., Wenger, A. M., Rowell, W. J., Khan, Z. M., Farek, J., Zhu, Y., Pisupati, A., Mahmoud, M., Xiao, C., Yoo, B., Sahraeian, S. M. E., Miller, D. E., ... Sedlazeck, F. J. (2022). Curated variation benchmarks for challenging medically relevant autosomal genes. *Nature Biotechnology*, 40(5), 672-680. <https://doi.org/10.1038/s41587-021-01158-1>
- Webster, A. & Eriksson, L. (2008). Governance-by-standards in the field of stem cells: managing uncertainty in the world of "basic innovation". *New Genetics and Society*, 27(2), 99-111. <https://doi.org/10.1080/14636770802077009>
- Willig, L. K., Petrikin, J. E., Smith, L. D., Saunders, C. J., Thiffault, I., Miller, N. A., Soden, S. E., Cakici, J. A., Herd, S. M., Twist, G., Noll, A., Creed, M., Alba, P. M., Carpenter, S. L., Clements, M. A., Fischer, R. T., Hays, J. A., Kilbride, H., McDonough, R. J., ... Kingsmore, S. F. (2015). Whole-genome sequencing for identification of Mendelian disorders in critically ill infants: a retrospective analysis of diagnostic and clinical findings. *The Lancet Respiratory Medicine*, 3(5), 377-387. [https://doi.org/10.1016/S2213-2600\(15\)00139-3](https://doi.org/10.1016/S2213-2600(15)00139-3)
- Wu, B., Kang, W., Wang, Y., Zhuang, D., Chen, L., Li, L., Su, Y., Pan, X., Wei, Q., Tang, Z., Li, Y., Gao, J., Cheng, R., Zhou, W., Wang, Z., Qiu, G., Wang, J., Yang, L., Zhang, P., ... Zhou, W. (2021). Application of Full-Spectrum Rapid Clinical Genome Sequencing Improves Diagnostic Rate and

- Clinical Outcomes in Critically Ill Infants in the China Neonatal Genomes Project*. *Critical Care Medicine*, 49(10), 1674. <https://doi.org/10.1097/CCM.0000000000005052>
- Yates, J. & Murphy, C. N. (2007, 1^{er} janvier). Coordinating International Standards: The Formation of the ISO. *MIT Sloan Research Paper*. 4638-07. <https://doi.org/10.2139/ssrn.962455>
- Zerubavel, E. (1982). The Standardization of Time: A Sociohistorical Perspective. *American Journal of Sociology*, 88(1), 1-23. <https://doi.org/10.1086/227631>
- Zook, J. M., Chapman, B., Wang, J., Mittelman, D., Hofmann, O., Hide, W. & Salit, M. (2014). Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. *Nature Biotechnology*, 32(3), 246-251. <https://doi.org/10.1038/nbt.2835>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York : PublicAffairs.

NOTES

1. Je souscris ici à Peter Dear et Sheila Jasanoff, lorsqu'ils affirment que ce qui relie les chercheur-euses de cette communauté hétérogène labellisée STS est la conviction que « depuis les travaux de Ludwik Fleck, Thomas Kuhn, et David Bloor, la science n'est pas le froid royaume de l'empirisme logique et des principes premiers, que nous préexistons nos connaissances du monde (tout comme le monde préexiste nos connaissances), que la matérialité est centrale à la formation et à la mise à l'épreuve des vérités scientifiques, et que les sciences et les dynamiques des pratiques scientifiques et techniques sont un terrain fertile à l'analyse sociale, politique et éthique » (Dear & Jasanoff, 2010 : 761. Ma traduction).
2. Sur les processus frictionnels de définition de normes, notamment informatiques, dans les pratiques scientifiques, voir notamment Edwards *et al.* (2011) et Jatton et Vinck (2016).
3. Comme le résume bien Lawrence Busch : « Au fur et à mesure de l'utilisation des standards, les personnes et les choses sont testées, ce qui permet de déterminer ce qui compte véritablement. Les personnes et les choses qui réussissent les tests [...] sont ensuite en mesure d'intégrer des réseaux variés » (2013 : 12. Ma traduction).
4. Comme nous le verrons par la suite, dans ce cas d'étude, l'accréditation constitue une reconnaissance officielle, fondée sur une base légale, attestant de la conformité à une norme ISO. Elle se distingue de la 'simple' certification tant par la nature des épreuves qu'elle implique que par les opportunités qu'elle ouvre.
5. L'enquête elle-même a été financée par un projet qui s'attachait explicitement à documenter le développement de la médecine personnalisée en Suisse.
6. Tous les noms de personnes et d'institutions sont fictifs, pour des raisons de confidentialité.
7. L'International Standard Organization est une institution indépendante dont le siège se situe à Genève. À la suite de la seconde guerre mondiale, l'ISO a émergé comme une institution clé dans la coordination des normes internationales. Un temps limité aux produits manufacturés, les standards émanant de l'ISO s'étendent aujourd'hui aux pratiques organisationnelles, à la gestion environnementale ou encore aux droits de humains. Sur l'histoire, qui reste encore largement à faire, de cette institution centrale à la grande accélération de l'après-guerre, voir Yates et Murphy (2007).
8. Plus précisément, la QUALAB est une association qui regroupe des représentants de l'Association suisse des hôpitaux, de l'Association des laboratoires d'analyses, de l'Association des professionnels de la santé, de l'Association faîtière des pharmaciens, de l'Association faîtière des caisses-maladie et de la Commission des tarifs médicaux.

9. La liste des laboratoires d'analyse autorisés par l'OFSP est disponible ici : <https://backend.bag.admin.ch/fileservice/sdweb-docs-prod-bagadminch-files/files/2025/03/18/8b574501-5ab6-4d7f-8db0-5db05073db4d.pdf>

10. Plusieurs laboratoires privés agréés par l'OFSP et fondés par d'anciens ou d'actuels professeurs d'université semblent également bénéficier de cette logique de proximité géographique et institutionnelle.

11. Le séquençage du génome entier (WGS) définit tous les nucléotides – base par base – de l'ADN d'un individu, tandis que le séquençage de l'exome entier (WES) se concentre sur les régions codant principalement pour les protéines (exons). Dans le cas spécifique du Centre, le séquençage de l'ARN (RNAseq) était principalement utilisé pour l'étude génétique des virus, en particulier le SARS-CoV-2.

12. FASTQ est le principal format textuel utilisé pour stocker les séquences de nucléotides (par exemple, « TTACCG ») ainsi qu'un score de qualité pour chaque nucléotide. Dans le cas des dispositifs de séquençage distribué par l'entreprise Illumina – acteur important de ce marché –, les fichiers FASTQ sont généralement générés à partir de données au format Binary Base Call (BCL) au cours de ce qui est souvent appelé la « conversion BCL-FASTQ ».

13. Après la génération des données FASTQ, qui proviennent elles-mêmes de données séquencées « brutes » au format BCL, il est nécessaire de les aligner sur un génome de référence. Les données résultant de ce processus d'alignement complexe – souvent appelé « mapping » – sont d'abord encodées dans un format texte délimité par des tabulations appelé *Sequence Alignment/Map* (SAM), puis converties en format binaire – *Binary Alignment/Map* (BAM) – afin de faciliter leur traitement ultérieur. Mais en raison du très grand nombre d'éléments inscrits dans ces fichiers BAM, les erreurs d'alignement sont potentiellement nombreuses. D'où la nécessité d'opérations de post-traitement plus ou moins automatisées, qui incluent par exemple des suppressions ou des recalibrages de scores de qualité de base. Et ce n'est qu'une fois que les données BAM ont été post-traitées que l'analyse d'appel de variants (*variant calling*) peut effectivement commencer, conduisant à la génération de jeux de données VCF (*Variant Call Format*) qui répertorient ce que l'on appelle souvent des « candidats variants ». Le plus souvent, ces listes de candidats variants encodés en que données VCF sont ensuite affinées – généralement en fonction des phénotypes des patient·es (lorsqu'ils sont disponibles) – pour aboutir finalement à une liste de « variants d'intérêt ». Et c'est bien cette liste de variants d'intérêt qui, dans l'idéal, sera communiquée dans un tableau de type Excel aux clinicien·nes généticien·nes afin d'étayer certains diagnostics et décisions thérapeutiques. Pour plus de détails sur le processus (théorique) qui permet de passer des données FASTQ aux données VCF, voir Pfeifer (2017).

14. GATK est l'acronyme de *Genome Analysis Toolkit* (boîte à outils d'analyse du génome) et constitue un vaste ensemble modulaire de programmes d'analyse de la lignée germinale. Initialement publié en 2010 par un petit groupe de bioinformaticien·nes du Broad Institute du MIT et de Harvard, il s'est depuis fortement développé pour devenir la référence industrielle open-source pour l'appel de variants génétiques.

15. DRAGEN a été développé dans les années 2015 par la startup de biotechnologie Edico Genome. Après son acquisition par Illumina en 2018 (pour un montant non divulgué), DRAGEN a été intégré au portefeuille de séquençage d'Illumina.

16. Congenica est une entreprise de santé numérique basée au Royaume-Uni qui commercialise un environnement logiciel intégré pour l'analyse, l'interprétation et la distribution de données génomiques. Issue de *Genomics England* – la société détenue par le ministère britannique de la santé pour mener à bien son *100'000 Genomes Project* (Turnbull *et al.*, 2018) – Congenica a l'avantage d'être certifié comme dispositif médical (norme ISO 13485) par l'Union européenne, ce qui rend les données qu'elle génère mobilisables dans la plupart des contextes cliniques de l'Espace économique européen.

17. Un script shell est un petit programme informatique écrit en shell, un langage de programmation informatique conçu pour interagir avec les interpréteurs de ligne de commande des systèmes d'exploitation Unix tels que Linux ou macOS.

18. Toutes les figures qui suivent sont extraites d'un document de travail .html, rédigé par Noah et Emily, qui présente la procédure et les résultats de leur expérience comparative. Toutes les citations sont des reconstitutions de notes prises à la volée par l'ethnographe.

19. Pour les illustrations ne comportant pas de crédits, nous publions les données en accord avec l'institution dont elles émanent, en revanche, pour des raisons de confidentialité nous préservons l'identité de cette dernière.

20. Le score de qualité des variants de GATK fait référence à la probabilité que le programme reconnaisse correctement une position donnée du génome en tant que variation par rapport au génome de référence dans au moins une position. Comme expliqué sur la page web dédiée à GATK (GATK Team, 2024), cette estimation dite de Phred (*Phred-scaled estimate*, initialement utilisée pour quantifier la probabilité d'erreurs d'identification des nucléases dans les données FASTQ) peut varier de 0 à l'infini, un score plus élevé indiquant une plus grande probabilité qu'une décision soit correcte, un score plus faible indiquant une plus grande probabilité que la décision soit incorrecte.

21. *Genome in a Bottle* fait référence à la séquence ADN humaine faisant autorité, établie par le National Institute of Standards and Technology, basé aux États-Unis, à partir d'un assemblage de plusieurs individus (Zook *et al.*, 2014). Cette norme de référence est toujours en cours d'élaboration, notamment pour mieux définir ses régions de faible complexité (Wagner *et al.*, 2022) et la rendre plus apte à intégrer la diversité génétique humaine (Jarvis *et al.*, 2022).

22. IGV correspond à *Integrative Genomics Viewer*, un logiciel open-source de visualisation interactive du génome développé et maintenu par des équipes travaillant à l'UC San Diego et au Broad Institute du MIT et de Harvard. Pour plus de détails sur l'histoire d'IGV, voir Robinson *et al.* (2011).

23. Les nombreuses techniques informatiques de correction d'erreurs actuellement employées en génomique computationnelles sont le plus souvent des variations de modèles bien connus (par exemple, l'analyse du spectre k-mer). La plupart d'entre elles, cependant, nécessitent une variété décente de nucléotides afin d'ancrer leurs opérations algorithmiques. Pour un aperçu récent des techniques de correction d'erreurs, voir Mitchell *et al.* (2020).

24. Une délétion constitue *in fine* une perte d'information génétique, non problématique si la région déléetée est non codante (comme dans la grande majorité des cas), mais potentiellement pathogène lorsqu'elle touche une région déterminante.

25. Le Centre n'a pas manqué de mettre en avant cet atout, dans la continuité de sa stratégie de distinction. Cette reconnaissance a notamment été soulignée dans une série d'articles – certains publiés par la société Illumina – qui ont pris soin de le présenter comme un établissement accrédité ISO.

RÉSUMÉS

Depuis plusieurs décennies, les *Science & Technology Studies* (STS) ont mis en évidence le rôle central des normes et standards dans la construction et la circulation du savoir technoscientifique et biomédical. Cependant, deux angles morts persistent : d'une part, la standardisation des pratiques bio-informatiques, essentielles en biomédecine contemporaine,

reste sous-explorée ; d'autre part, les préparations aux *épreuves de standardisation* (tests, évaluations, audits) sont rarement étudiées. Cet article comble partiellement ces lacunes en documentant la préparation à une épreuve de standardisation d'un pipeline bio-informatique au sein d'un centre suisse spécialisé en génomique computationnelle. À partir d'une enquête ethnographique de quatre mois, il analyse la manière dont une équipe de bio-informaticien·nes compare deux suites logicielles – GATK et DRAGEN – pour l'appel de variants, dans le cadre d'une démarche de conformité au standard ISO 15189 qui spécifie les exigences de qualité pour les laboratoires biomédicaux. L'étude met en lumière les stratégies mises en œuvre pour identifier, qualifier et négocier les écarts entre ces outils logiciels, en soulignant la marge d'interprétation propre aux processus de standardisation en bio-informatique. Elle propose également une réflexion plus large sur l'articulation entre standardisation et préparation en bio-informatique, suggérant que ces épreuves ne se limitent pas à des procédures techniques, mais impliquent également un travail créatif d'optimisation et d'anticipation. En révélant les dynamiques de pouvoir, de négociation et d'adaptation qui nourrissent les pratiques de standardisation en bio-informatique, cette analyse contribue à une meilleure compréhension du rôle des standards dans l'évolution des infrastructures biomédicales qui soutiennent la mise en place d'une médecine dite personnalisée ou de précision.

For several decades, Science & Technology Studies (STS) have highlighted the central role of norms and standards in the construction and circulation of technoscientific and biomedical knowledge. However, two blind spots persist: on the one hand, the standardization of bioinformatics practices, essential in contemporary biomedicine, remains underexplored; on the other hand, preparations for standardization trials (tests, evaluations, audits) are rarely studied. This article partially fills these gaps by documenting the preparation for a standardization trial of a bioinformatics pipeline within a Swiss center specialized in computational genomics. Based on a four-month ethnographic study, it analyzes how a team of bioinformaticians compares two software suites – GATK and DRAGEN – for variant calling, as part of their efforts to meet the ISO 15189 standard for quality in biomedical laboratories. The study highlights the strategies implemented to identify, qualify, and negotiate discrepancies between these software tools, highlighting the room for interpretation inherent in bioinformatics standardization processes. The study also offers a broader reflection on the articulation between standardization and preparation in bioinformatics, suggesting that these trials are not limited to technical procedures but also involve creative work in optimization and anticipation. By revealing the dynamics of power, negotiation and adaptation that feed standardization practices in bioinformatics, this analysis contributes to a better understanding of the role of standards in the evolution of biomedical infrastructures that support the implementation of so-called personalized or precision medicine.

Desde hace varias décadas, los estudios sobre ciencia y tecnología (Science & Technology Studies, STS) destacan el papel central de las normas y estándares en la construcción y circulación del conocimiento tecnocientífico y biomédico. Sin embargo, persisten dos puntos ciegos: por un lado, la normalización de las prácticas bioinformáticas, esencial en la biomedicina contemporánea, sigue siendo poco explorada; por otro, rara vez se estudia la preparación de ensayos de normalización (pruebas, evaluaciones, auditorías). Este artículo colma parcialmente estas lagunas documentando la preparación de un ensayo de normalización de un proceso bioinformático en un centro suizo especializado en genómica computacional. Basándose en un estudio etnográfico de cuatro meses, analiza cómo un equipo de bioinformáticos compara dos suites de software – GATK y DRAGEN – para la llamada de variantes, como parte de sus esfuerzos por cumplir la norma ISO 15189 de calidad en laboratorios biomédicos. El estudio pone de relieve las estrategias aplicadas para identificar, calificar y negociar las discrepancias entre estas herramientas de software, destacando el margen de interpretación inherente a los procesos de normalización de

la bioinformática. El estudio también ofrece una reflexión más amplia sobre la articulación entre normalización y preparación en bioinformática, sugiriendo que estos ensayos no se limitan a procedimientos técnicos, sino que también implican un trabajo creativo de optimización y anticipación. Al revelar las dinámicas de poder, negociación y adaptación que alimentan las prácticas de normalización en bioinformática, este análisis contribuye a una mejor comprensión del papel de las normas en la evolución de las infraestructuras biomédicas que apoyan la implantación de la llamada medicina personalizada o de precisión.

Die Wissenschafts- und Technologiestudien (Science & Technology Studies, STS) betonen seit mehreren Jahrzehnten die zentrale Rolle von Normen und Standards bei der Konstruktion und Verbreitung von technowissenschaftlichem und biomedizinischem Wissen. Zwei blinde Flecken bleiben jedoch bestehen: Einerseits bleibt die Standardisierung bioinformatischer Praktiken, die in der zeitgenössischen Biomedizin von wesentlicher Bedeutung sind, unerforscht; andererseits werden die Vorbereitungen für Standardisierungsversuche (Tests, Evaluierungen, Audits) nur selten untersucht. Dieser Artikel füllt diese Lücken teilweise, indem er die Vorbereitungen für einen Standardisierungsversuch einer Bioinformatik-Pipeline in einem Schweizer Zentrum für Computational Genomics dokumentiert. Auf der Grundlage einer viermonatigen ethnografischen Studie wird analysiert, wie ein Team von Bioinformatikern zwei Software-Suiten – GATK und DRAGEN – für die Variantenerkennung vergleicht, um die ISO-Norm 15189 für Qualität in biomedizinischen Labors zu erfüllen. Die Studie hebt die Strategien hervor, die zur Identifizierung, Qualifizierung und Verhandlung von Diskrepanzen zwischen diesen Softwaretools eingesetzt wurden, und verdeutlicht den Interpretationsspielraum, der den Standardisierungsprozessen in der Bioinformatik innewohnt. Die Studie bietet auch eine umfassendere Reflexion über die Verbindung zwischen Standardisierung und Vorbereitung in der Bioinformatik und legt nahe, dass diese Versuche nicht auf technische Verfahren beschränkt sind, sondern auch kreative Arbeit bei der Optimierung und Antizipation beinhalten. Durch die Offenlegung der Dynamik von Macht, Verhandlung und Anpassung, die die Standardisierungspraktiken in der Bioinformatik speisen, trägt diese Analyse zu einem besseren Verständnis der Rolle von Standards in der Entwicklung biomedizinischer Infrastrukturen bei, die die Umsetzung der sogenannten personalisierten oder Präzisionsmedizin unterstützen.

INDEX

Schlüsselwörter : computergestützte genomik, variantenaufwurf, standardisierung, präzisionsmedizin, personalisierte medizin, ethnographie

Palabras claves : genómica computacional, llamada de variantes, normalización, medicina de precisión, medicina personalizada, etnografía

Keywords : computational genomics, variant calling, standardization, precision medicine, personalized medicine, ethnography

Mots-clés : génomique computationnelle, appel de variants, standardisation, médecine de précision, médecine personnalisée, ethnographie

AUTEUR

FLORIAN JATON

Faculté des sciences sociales et politiques, Université de Lausanne & Collège des Humanités, EPFL
Florian Jaton est enseignant-chercheur à l'Université de Lausanne (Faculté des sciences sociales et politiques) et à l'EPFL (Collège des Humanités). Il a précédemment travaillé à la Donald Bren

School of Information and Computer Science à l'Université de Californie, Irvine, au Centre de sociologie de l'innovation à Mines Paris, Université PSL et au Geneva Graduate Institute of International and Development Studies. Ses intérêts de recherche sont la sociologie des algorithmes, la philosophie des mathématiques et l'histoire de l'informatique. Il est l'auteur de *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating* (MIT Press, 2021 ; en accès libre). ORCID : <https://orcid.org/0000-0002-5001-9098>

Adresses : Faculté des sciences sociales et politiques, Institut des sciences sociales, 1015 Lausanne, Suisse ; Collège des Humanités, EPFL, 1015 Lausanne, Suisse.

florian.jaton@unil.ch ; florian.jaton@epfl.ch