

# ITMO Cancer

## Extraits des Orientations

### Stratégiques

#### Aspects mathématiques et données

La variété des données étudiées pour comprendre et traiter les cancers ne cesse de croître. Il est maintenant admis que des données anatomopathologiques, « omiques », épigénétiques, immunologiques, pharmacologiques, microbiotiques, environnementales, comportementales, nutritionnelles, métaboliques, familiales ou relatives à l'activité physique à différents âges de la vie doivent être prises en compte conjointement pour décrire, comprendre, traiter les cancers. Cette révolution conceptuelle va de pair avec les avancées technologiques majeures des dernières années concernant les techniques à grande échelle et à haut débit qui permettent désormais :

- d'offrir une nouvelle dimension pour appréhender l'extraordinaire complexité des cancers ;
- de proposer de nouvelles approches thérapeutiques et de prédire leur validité ;
- d'analyser et de comprendre les résistances aux traitements ;
- d'identifier des biomarqueurs de diagnostic, de théranostic et d'évolution de la maladie.

Cependant, identifier les paramètres pertinents au sein de la masse des données produites, modéliser les phénomènes biologiques, intégrer des données multidimensionnelles restent des enjeux de taille qui nécessitent le développement de nouvelles approches mathématiques de modélisation, de fouille, d'apprentissage statistique et d'analyse de données. *In fine*, cela doit permettre à la fois de mieux comprendre les interactions à l'œuvre dans les grands processus biologiques, normaux ou pathologiques, et de mieux prédire l'évolution des tumeurs et celle de leur traitement.

#### A. En termes de recherche fondamentale

L'étude de l'hétérogénéité tumorale et de la résistance aux traitements montre clairement que les populations tumorales se comportent comme un système de type darwinien, sensible à la pression de sélection.

Il est donc nécessaire d'étudier la tumorigenèse sous l'angle de l'évolution, en caractérisant des populations de cellules cancéreuses depuis un état précancéreux en puissance jusqu'à un état cancéreux affirmé (évolution à court terme, échelle de la vie humaine). De plus, l'étude de l'évolution des génomes depuis les stades unicellulaires jusqu'à une multicellularité organisée de façon cohérente et bien différenciée devrait permettre

de mieux comprendre les étapes-clés de la multicellularité, ainsi que les fragilités qui, à ces étapes-clés, pourraient être à l'origine de cancers, qu'ils soient génétiquement favorisés ou sporadiques.

Il conviendrait ainsi de modéliser :

- L'évolution des populations de cellules, cancéreuses ou non, et de leurs interactions ;
- La représentation et le contrôle de la résistance, réversible ou non, aux traitements du cancer (en particulier par la prise en compte de l'hétérogénéité phénotypique/génotypique des tumeurs) ;
- Le contrôle thérapeutique exercé sur des cibles fonctionnelles (prolifération, mort cellulaire, différenciation) et le perfectionnement du contrôle par des stratégies thérapeutiques optimisées (échelle de la population de cellules) ;
- La représentation des interactions tumeur/stroma (compétitives ou mutualistes) ;
- La représentation de cibles moléculaires et de leur contrôle à l'échelle unicellulaire.

→ La modélisation mathématique (dynamique adaptative), en codant de façon continue la variabilité biologique (phénotypes de résistance, phénotypes de plasticité) des populations de cellules cancéreuses, devrait permettre de prédire leur évolution phénotypique et leur réponse aux traitements considérés comme pression de sélection. Grâce à ces modèles continus, l'origine de la dormance (apparente) de certains clones de cellules cancéreuses et les mécanismes de la résistance aux traitements et de la maladie résiduelle pourraient également être élucidés.

#### B. En termes de recherche translationnelle et clinique

L'un des enjeux majeurs de la médecine dite de « précision » est de stratifier les patients selon leurs capacités à répondre à tel ou tel traitement. Le corollaire consiste à identifier des biomarqueurs pertinents permettant de classer les patients. Compte tenu de la diversité des possibilités techniques d'identification de marqueurs pertinents, les panels de biomarqueurs comportent par nature plusieurs paramètres qu'ils convient de conjuguer, décrypter, intégrer.

→ Des approches mathématiques et de nouveaux algorithmes pour intégrer cette dimension « multiparamétrique » devront donc être développés.

Si l'harmonisation des procédures entre les laboratoires pour la caractérisation d'un profil multiparamétrique de ces biomarqueurs constitue un défi majeur, la production massive de données nécessite également d'être capitalisée de manière optimale grâce à une organisation précise de l'information.

→ L'expertise multidisciplinaire et l'interaction des chercheurs, biologistes moléculaires, cliniciens, pathologistes et bio-informaticiens se révèlent incontournables pour faire face à ces défis et permettre la mise en place de la médecine de précision en routine.

Face à la toxicité potentielle des nouvelles thérapies ou de leurs combinaisons, il sera également crucial de mettre au point des modèles de prédiction du bénéfice clinique, soit :

- Par des études de populations de patients (en particulier de PK-PD populationnelle) utilisant des études statistiques de type santé publique ;
- Par des études de modélisation de la réponse au traitement des organismes représentés par des populations de cellules cancéreuses, stromales transformées et saines en interaction.

→ Cette évaluation du bénéfice clinique fait intervenir la modélisation mathématique *via* des études de dynamique adaptative des populations de cellules concernées.

## C. En termes de valorisation de la recherche

L'évolution des technologies numériques et des outils de communication autorise actuellement la collecte de banques de données numériques (prélèvements, images, dossiers cliniques) potentiellement très riches, en particulier si elles sont mises en commun, pour atteindre des tailles de cohortes qui seraient autrement inaccessibles à un seul centre de recherche, voire à un seul pays.

Face à ces évolutions, il est plus que jamais nécessaire de produire des données de qualité contrôlée, assorties d'une description des méthodes permettant de les reproduire, d'encourager l'utilisation des méthodologies de référence, d'accompagner leur amélioration pour la production de bases de données.

→ Ces banques de données constituent un matériel expérimental permettant l'émergence de connaissances originales au moyen de méthodes dédiées à leur exploration, fondées sur l'intelligence artificielle, l'apprentissage statistique et l'enrichissement sémantique. La recherche de telles méthodes de fouilles de données (*datamining*) biomédicales par les technologies de l'information les plus modernes doit être encouragée. La mise en relation des équipes de recherche sur le cancer avec des laboratoires de recherche intégrative en oncologie permettrait de favoriser le développement de modèles explicatifs et prédictifs.

Par ailleurs, dans un contexte de séquençage génomique massif des tumeurs grâce aux technologies à haut débit, il est crucial que les données qui seront accumulées puissent être exploitées à des fins de recherche, car elles constituent une richesse considérable pour :

- Améliorer notre connaissance des cancers ;

- Mieux comprendre la fonction des gènes, dont très peu sont annotés (réseau moléculaire, localisation, fonction, etc.) ;
- Produire de nouvelles hypothèses testables en laboratoire, utiles dans le futur en clinique humaine.

→ Le *datamining* des bases de données génomiques de santé doit être libre d'accès pour les personnels de recherche.

En conclusion, les défis majeurs actuels de la recherche en cancérologie visent à obtenir une caractérisation approfondie de la signature biologique de chaque tumeur dans l'espace (son environnement) et dans le temps (tout au long de son évolution, depuis son initiation jusqu'à sa rechute éventuelle). Une telle signature permettrait en effet la mise en place d'un programme de prévention et de surveillance adapté à la tumeur, et de proposer un traitement optimisé et personnalisé à chaque patient. Ces objectifs ne peuvent être atteints sans l'apport significatif de disciplines scientifiques non biomédicales, mathématiques et informatiques en particulier.

*Extraits du texte « Orientations stratégiques » de l'ITMO Cancer d'Aviesan (décembre 2017) recomposés par François Castagner, Emmanuelle Chollet, Muriel Altabef <https://itcancer.aviesan.fr/index.php?pagendx=980>*