



**Laboratoire de Mathématiques et Informatique pour la Complexité et les
Systèmes
MICS**

Présente

L'AVIS DE SOUTENANCE

De Monsieur Yoann PRADAT

Laboratoire MICS, CentraleSupélec, Université Paris Saclay, soutiendra
publiquement ses travaux de thèse de doctorat intitulés :

***“Analyzes of genomic and transcriptomic profiles of metastatic tumors from precision
medicine clinical trials.”***

Sous la Direction de Monsieur Paul-Henry COURNÈDE et Monsieur Fabrice ANDRÉ

Le 14 février 2024 à 14h00

À l'institut Gustave Roussy, dans la **Salle 21- 23 (ONYX)** – Bâtiment B2M.

Membres du jury :

Eric LETOUZÉ, Rapporteur & Examineur, Université de Nantes

Tatiana POPOVA, Rapportrice & Examinatrice, Université Paris Sciences et Lettres

Macha NIKOLSKI, Examinatrice, Université de Bordeaux

Elli PAPAEMMANUIL, Examinatrice, Memorial Sloan Kettering Cancer Center

Isidro CORTÉS CIRIANO, Examineur, EMBL

Edwin CUPPEN, Examineur, UMC Utrecht

Résumé :

À l'ère de l'acquisition et de l'analyse massives de données, notre compréhension de l'apparition et de l'évolution du cancer s'est améliorée à la lumière des résultats dérivés des analyses des portraits moléculaires de dizaines de milliers de tumeurs à travers le monde. L'avènement des technologies de séquençage de nouvelle génération dans les années 2000 a révolutionné la façon dont nous étudions les cellules tumorales des patients atteints de cancer. Ces technologies permettent désormais le profilage systématique de l'ensemble de l'exome, du transcriptome et même du génome entier. Étant donné la place sans cesse grandissant du séquençage dans la recherche et la pratique clinique, une compréhension complète des différents aspects de l'analyse des données de séquençage est primordiale. Bien que le séquençage à haut débit ne fasse pas encore partie du parcours clinique de tous les patients atteints de cancer, des profilages moléculaires ont été proposés à nombre de patients participant à des essais cliniques et est aujourd'hui utilisé en routine pour un certaines indications. Cette grande quantité de données est désormais disponible pour étayer de nombreuses recherches, allant de l'établissement de portraits moléculaires détaillés de groupes particuliers de patients au déchiffrement des liens entre les génotypes des tumeurs et les parcours cliniques des patients.

Cette thèse couvre de nombreux aspects de l'analyse rétrospective d'une large cohorte de patients atteints de cancer, ainsi que des revues détaillées de concepts et d'outils importants de l'oncologie moderne. Les premiers chapitres présentent des concepts généraux sur la biologie et la classification du cancer, qui sont fondamentaux pour les décisions thérapeutiques mais également pour l'orientation et l'organisation de la recherche. La dernière section de ce premier chapitre présente en outre au lecteur des réflexions sur la place croissante du profilage moléculaire et son impact sur les classifications et les conceptions d'essais. Le deuxième chapitre passe en revue en détail les outils informatiques et les bases de données utilisés pour analyser les données de séquençage et extraire des informations cliniquement pertinentes. Les deux premiers chapitres servent de base à l'analyse d'une grande cohorte de patients pan-cancer présentée dans le troisième chapitre. Cette cohorte, META-PRISM, comprend 1031 patients issus de deux de grands essais de médecine de précision menés à Gustave Roussy dans la décennie 2010-2020, dont un tiers ont bénéficié de séquençage de l'exome entier ou de l'ARN de leurs tumeurs. Comparée à d'autres analyses de grandes cohortes pan-cancer, cette étude se distingue par l'accent mis sur des patients réfractaires aux traitements, tous étant considérés comme affectés par un cancer incurable selon un comité multidisciplinaire. Les analyses comparatives avec une cohorte internationale de tumeurs primaires non traitées ont mis au jour des différences génétiques générales ainsi que de multiples différences spécifiques à certains types de tumeur. De plus, la modélisation prédictive de la survie des patients à l'aide de biomarqueurs moléculaires a montré que même les patients à un stade avancé peuvent bénéficier du séquençage pour des décisions thérapeutiques importantes, en particulier l'éligibilité aux essais. Le quatrième et dernier chapitre se concentre sur l'analyse des marqueurs génomiques connus et émergents de résistance aux traitements dans la cohorte META-PRISM, mais également dans deux autres cohortes issues d'études cliniques récentes menées à Gustave Roussy, dont l'une s'intéresse à un anticorps conjugué récemment approuvé pour les cancers du sein et l'autre à une classe particulière d'inhibiteurs pour les cancers de la vessie. Ces deux études ont démontré que les altérations de l'expression ou de la structure de la cible, ou l'activation de voies alternatives induite par une mutation, contribuent grandement à la résistance aux médicaments.

Abstract:

In the era of massive data acquisition and analysis, our understanding of cancer onset and evolution has improved in light of the results derived from the analyses of the molecular portraits of tens of thousands of tumors across the globe. The advent of next-generation sequencing technologies in the 2000s has revolutionized how we investigate the tumor cells of patients with cancer. These technologies were first used to characterize specific genomic regions but have matured over time to allow for the systematic profiling of the whole exome, transcriptome, and even whole genome of patients enrolled in clinical trials. As sequencing technologies' continued development will fuel new research areas and discoveries for many years, a complete understanding of the different aspects of sequencing data analysis is of utmost importance. Although high-throughput sequencing is not yet part of the standard of care for all cancer patients, extensive molecular profiling has been offered to significant numbers of patients participating in clinical trials and is now used in routine for an increasing number of indications. This large quantity of data is now available to support many research opportunities, ranging from drawing detailed molecular portraits of particular groups of patients to deciphering the links between tumors' genotypes and patients' clinical outcomes and, eventually, to advancing precision oncology.

This thesis covers many aspects involved in the retrospective analysis of a large cohort of cancer patients, along with detailed reviews of important concepts and tools from modern oncology. The first chapter introduces general concepts about cancer biology and classification, which are fundamental to the immediate treatment decisions but also to research wherein patients are organized according to clinically defined groups. The last section of this first chapter additionally introduces the reader to considerations about the growing place of molecular profiling and their impact on classifications and trial designs. The second chapter extensively reviews the computing tools and databases employed to analyze sequencing data and extract clinically meaningful information. The first two chapters serve as the laying bricks behind the analysis of a large pan-cancer cohort of patients presented in the third chapter. This cohort, META-PRISM, comprises 1,031 patients from two large precision medicine trials led at Gustave Roussy in the decade 2010-2020. About a third of the enrolled patients benefited from whole-exome and RNA sequencing technologies at trial entry and, for a subset of them, at resistance. Compared to other analyses of large pan-cancer cohorts, this study stands out for its focus on patients that were refractory to treatments and the derivation of highly detailed and curated clinical histories. Notably, all patients shared the common characteristic of being in the advanced stages of their disease, deemed incurable by a multidisciplinary board. The comparative analyses against an international cohort of primary untreated tumors have shed light on a few tumor-type-agnostic genetic differences and multiple tumor-type-specific differences. Additionally, predictive modeling of patients' survival using molecular biomarkers has shown that even late-stage patients can benefit from sequencing for important therapeutic decisions, particularly trial eligibility. The fourth and last chapter focuses on the analysis of known and emerging genomic markers of treatment resistance in the META-PRISM cohort, but also in two other cohorts from recent clinical studies led by Gustave Roussy, one interested in a recently approved antibody-drug conjugate for breast cancers and another in a particular class of inhibitors for bladder cancers. These two studies demonstrated that alterations in the expression or structure of the target, or mutation-induced activation of alternative pathways are important contributors to drug resistance.