



HAL
open science

To identify the panoply of somatic mutations in each cell

Bernard Dujon

► **To cite this version:**

Bernard Dujon. To identify the panoply of somatic mutations in each cell. *Comptes Rendus Biologies*, 2020, 343 (3), pp.211-213. 10.5802/crbiol.25 . hal-03151187

HAL Id: hal-03151187

<https://hal.sorbonne-universite.fr/hal-03151187v1>

Submitted on 24 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus

Biologies


Bernard Dujon

To identify the panoply of somatic mutations in each cell

Volume 343, issue 3 (2020), p. 211-213.

<<https://doi.org/10.5802/crbio.25>>

© Académie des sciences, Paris and the authors, 2020.
Some rights reserved.

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



*Les Comptes Rendus. Biologies sont membres du
Centre Mersenne pour l'édition scientifique ouverte*
www.centre-mersenne.org



News and Views / *C'est apparu dans la presse*

To identify the panoply of somatic mutations in each cell

Identifier la panoplie des mutations somatiques de chaque cellule

Bernard Dujon^a

^a Sorbonne Université, Paris & Département “Genomes and Genetics”, Institut Pasteur, 25, rue du Docteur Roux, F75015 Paris, France
E-mail: bdujon@pasteur.fr

Manuscript received 20th October 2020, accepted 23rd October 2020.

While advances in DNA sequencing have already revealed many mutations in somatic tissues and cancers, the genetic heterogeneity of the cells that make them up has so far remained difficult to access. This difficulty limits our understanding of the phenomena that take place during aging or cancer proliferation. It results from the limitations of single cell genome sequencing because there is, in a diploid cell, only one DNA molecule corresponding to each parental allele [1]. Two methods partially solve the problem: *in vitro* culture of isolated cells or amplification of the DNA to be sequenced by PCR. But each has its limitations. Not all differentiated cell types are suitable for *in vitro* culture. And DNA amplification is a source of random mutations that must be distinguished from those that we are trying to identify in the genome of the cell under study.

In a particularly well-conducted recent work, the authors have made significant progress in addressing this problem [2]. Having chosen to focus on melanocytes, they reveal a rich mutational landscape

in cells derived from elderly (63 to 85 years old) people of European origin, living or recently deceased, with or without melanoma. For each of the 6 donors, 4 men and 2 women, skin samples taken from different points of the body exposed to varying degrees of solar radiation (face, back, shoulders, soles of the feet, etc.) were studied. Unlike keratinocytes, melanocytes only represent a small proportion of skin cells, making their individual analysis essential. Several critical steps are used to identify the mutations present in the genomes of these cells and to distinguish them from technical artifacts.

First, isolated cells are cultured *in vitro* for a limited number of generations resulting in colonies of only a few hundred cells from which DNA and RNA molecules are extracted and then amplified separately for sequencing. The interest of sequencing RNA is twofold. It ensures that the colonies come from melanocytes and not from other skin cells (the transcriptomes are distinct) and it facilitates the distinction between the mutations sought (which will

be coincident between the RNA and DNA sequences) and the numerous random mutations due to the amplifications. Next, the authors use two genetic screens to sort the mutations sought from those produced during amplification: haplotypic phasing and allelic frequency within DNA sequence readings. The first screen is applicable to any mutation sufficiently close to a heterozygous variant (true mutations are always associated with one of the two parental alleles, never with the other). For other mutations, a frequency of 50% of the readings is required. For the 133 melanocytes studied in this way from normal skin samples flanking or not a melanoma, the average number of mutations found is 7.9 per Mb, with a very large individual variation (from 0.8 to 32).

By focusing on the sequences of more than 500 genes known to be involved in cancers, as well as whole exomes in about a third of the melanocytes studied, the authors offer us an original panoply of somatic mutations (37 on average in each cell for the 500 genes selected), some of which are potentially pathogenic. Surprisingly, melanocytes from regions of the body most regularly exposed to the sun, such as the face, have a lower average number of mutations than those from regions of the body irregularly

exposed to the sun. While the explanation of this phenomenon remains to be discovered, this difference corroborates the fact that melanomas occur most often in the latter regions. More importantly, while many of the observed mutations bear the expected signature of UV induction, some melanocytes have far fewer mutations than others in the same biopsy as if they were protected from mutations for an unknown reason (the hypothesis of radiation protection by hair follicles is being considered). According to the authors, the presence of cells with few mutations among others that are abundantly mutated is reminiscent of what is observed in the lung cells of some former smokers. Finally, while about 20% of melanocytes show at least one mutation likely to lead to cancerous development, the authors noted that several cells from the same biopsy showed the same type of mutation, suggesting some scattered multiplication of precancerous cells in healthy tissues.

Beyond the great interest of these results for subsequent studies on the appearance of melanomas, this work opens a new path for the study of somatic mutations in healthy tissues that could rapidly bring us many essential discoveries.

French version

Si les progrès du séquençage de l'ADN ont déjà révélé de nombreuses mutations dans les tissus somatiques et les cancers, l'hétérogénéité génétique des cellules qui les composent est restée jusqu'à présent difficile d'accès. Cette difficulté limite notre compréhension des phénomènes qui s'y déroulent durant le vieillissement ou la prolifération cancéreuse. Elle résulte des limites du séquençage des génomes de cellules uniques car il n'existe, chez une cellule diploïde, qu'une seule molécule d'ADN correspondant à chaque allèle parental [1]. Deux méthodes permettent de résoudre partiellement le problème : cultiver *in vitro* les cellules isolées ou amplifier l'ADN à séquencer par PCR. Mais chacune a ses limitations. Tous les types de cellules différenciées ne se prêtent pas à la culture *in vitro*. Et l'amplification de l'ADN est source de mutations aléatoires qu'il faudra distinguer de celles que l'on cherche à identifier dans le génome de la cellule étudiée.

Dans un travail récent particulièrement bien conduit, les auteurs font progresser ce problème de manière significative [2]. Ayant choisi de s'intéresser aux mélanocytes, ils révèlent un riche paysage mutationnel chez des cellules de personnes âgées (63 à 85 ans) d'origine européenne, vivantes ou récemment décédées et atteintes ou non de mélanomes. Pour chacun des 6 sujets, 4 hommes et 2 femmes, des prélèvements cutanés faits en différents points du corps exposés à des degrés variables au rayonnement solaire (visage, dos, épaules, plante des pieds, ...) ont été étudiés. Contrairement aux kératinocytes, les mélanocytes ne représentent qu'une faible proportion des cellules cutanées, rendant indispensable leur analyse individuelle. Plusieurs étapes critiques sont utilisées pour identifier les mutations présentes dans les génomes de ces cellules et les distinguer d'artefacts techniques.

D'abord, les cellules isolées sont cultivées *in vitro* pour un nombre limité de générations aboutissant à des colonies de quelques centaines de cellules seulement à partir desquelles sont extraites les molécules d'ADN et d'ARN qui sont ensuite amplifiées séparément pour le séquençage. L'intérêt de séquencer les ARN est double. Il permet de s'assurer que les colonies proviennent bien de mélanocytes et non d'autres cellules cutanées (les transcriptomes sont distincts) et il facilite la distinction entre les mutations recherchées (qui seront coïncidentes entre les séquences d'ARN et d'ADN) et les nombreuses mutations aléatoires dues aux amplifications. Ensuite, les auteurs utilisent deux cribles génétiques pour trier les mutations recherchées de celles produites pendant l'amplification : le phasage haplotypique et la fréquence allélique au sein des lectures de séquence d'ADN. Le premier crible est applicable à toute mutation suffisamment proche d'un variant hétérozygote (les vraies mutations sont toujours associées à l'un des deux allèles parentaux, jamais à l'autre). Pour les autres mutations, une fréquence de 50% des lectures est demandée. Pour les 133 mélanocytes étudiés de cette façon issus d'échantillons de peau normale à proximité ou non de mélanomes, la moyenne du nombre de mutations trouvées est de 7,9 par Mb, avec une très grande variation individuelle (de 0,8 à 32).

En se focalisant sur les séquences de plus de 500 gènes connus pour leur implication dans les cancers, ainsi que sur les exomes entiers chez environ un tiers des mélanocytes étudiés, les auteurs nous offrent une panoplie originale de mutations somatiques (37 en moyenne dans chaque cellule pour les 500 gènes sélectionnés) parmi lesquelles certaines sont potentiellement pathogènes. De façon surprenante, les mélanocytes des régions du corps les plus

régulièrement exposées au soleil, comme le visage, ont un nombre moyen de mutations moins élevé que ceux provenant de régions du corps irrégulièrement exposées au soleil. Alors que l'explication de ce phénomène reste à découvrir, ceci corrobore le fait que les mélanomes apparaissent le plus souvent dans ces dernières régions. Mais surtout, alors que nombre des mutations observées portent la signature attendue d'une induction par les rayons UV, certains mélanocytes ont beaucoup moins de mutations que d'autres dans la même biopsie comme s'ils étaient protégés des mutations pour une raison inconnue (l'hypothèse d'une protection des rayonnements par les follicules pileux est envisagée). La présence de cellules peu mutées parmi d'autres qui le sont abondamment rappelle, selon les auteurs, ce que l'on observe au niveau des cellules pulmonaires de certains anciens fumeurs. Enfin, alors qu'environ 20% des mélanocytes montrent au moins une mutation susceptible d'entraîner un développement cancéreux, les auteurs ont remarqué que plusieurs cellules issues de la même biopsie montraient une même mutation de ce type, suggérant une certaine multiplication dispersée de cellules précancéreuses dans des tissus sains.

Au-delà du grand intérêt de ces résultats pour les études ultérieures sur l'apparition des mélanomes, ce travail ouvre une nouvelle voie pour l'étude des mutations somatiques dans les tissus sains susceptible de nous apporter rapidement de nombreuses découvertes essentielles.

References

- [1] C. Gawad *et al.*, "Single-cell genome sequencing: current state of the science", *Nature Rev. Genet.* **17** (2016), p. 175-188.
- [2] J. Tang *et al.*, "The genomic landscapes of individual melanocytes from human skin", *Nature* (2020), doi:10.1038/s41586-020-2785-8 (october 7 on-line).