# Introduction:

The mutational catalog of a given cancer genome is an archeological record of the endogenous and exogenous mutational processes operative during the lifetime of a cancer patient. From the very first fertilized egg and until the development of a cancerous cell, DNA is continuously accumulating mutations (Figure 01; Stratton et al., 2009).

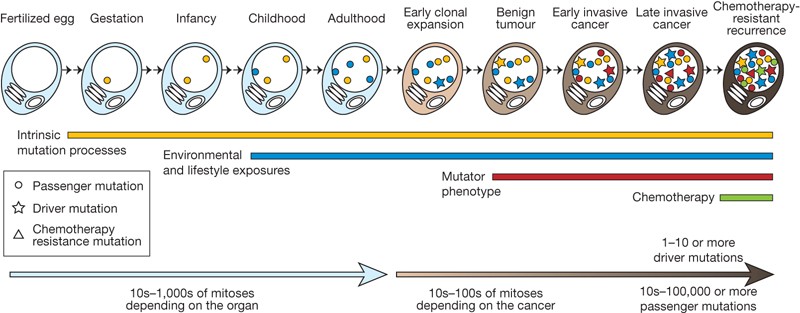


Figure 1

The vast majority of mutations occurring are effectively repaired. However, in some cases the mutation cannot profit from a proper repair which could lead to cell apoptosis, senescence or challenge genome stability leading to various outcomes (Hakem, 2008). Previous studies were able to elucidate several mutations’ consequences (Chin et al., 2011; Poduri et al., 2013; Vuong et al., 2014), some of which are currently used as prognostics (Olivier et al., 2011; Griffith et al., 2018) or predictive (Martin et al., 2013) markers. As an example, breast cancer patients with BRCA1 mutations were found to exhibit a significantly worse survival rate over a period of ten years compared to patients without BRCA1 mutations (Huszno et al., 2019). Despite the advances made in terms of understanding the consequences of mutations, there is still a lack of understanding of the reasons behind these mutations in the first place.

Owing to large scale sequencing initiatives such as The Cancer Genome Atlas -TCGA- and Pan-Cancer Analysis of Whole Genome -PCAWG- (ICGC/TCGA, 2020), thousands of cancer mutational catalogs were obtained spanning most of the cancer types, thus offering the unique opportunity to extract the mutational patterns left by the mutational processes operative in cancer. These mutational patterns are usually referred to as mutational signatures, which can be viewed as a probability mass function of a multinomial distribution for context-based mutation types. Frequently six mutation types are used [C>A, C>G, C>T, T>A, T>C, T>G] plus the flanking 5’ and 3’ bases yielding 96 mutation type (Alexandrov et al., 2013). It is hypothesized that every signature is an imprint of a mutational process that was operative in the cancer cell. Thus, a mutational catalog is regarded as the combination of mutations generated by these mutational processes that are represented by the mutational signatures (Alexandrov et al., 2013). In 2013 Alexandrov and colleagues published the first mathematical approach for deciphering the mutational signatures operating in each cancer type, this method is currently known as SigProfiler (Alexandrov et al., 2013) and is the one used to extract the reference signatures that are available at the Catalog Of Somatic Mutations In Cancer -COSMIC- (Tate et al., 2019). Since then, several other tools were developed for the same purpose (See Omichessan et al., 2019 for a comprehensive list of tools). The largest study in this sense was also performed by Alexandrov and colleagues in 2020. It yielded 49 single base substitutions signatures (SBS), 11 doublet base substitutions signatures (DBS), and 17 insertion-deletion signatures (INDELS). Some of these signatures are of known aetiology while many remain ambiguous (Alexandrov et al., 2020).

Replication stress is a phenomenon that can challenge the genome stability through interference with replication fork progression. Replication stress can be originated from different events that could be endogenous or exogenous. Namely, DNA lesions, misincorporation of ribonucleotides, unusual DNA structures, conflicts between replication and transcription machineries as well as other sources (Zeman and Cimprich, 2014).

Replication stress response involve many pathways themselves include several components. A depletion of one of these components or a malfunctioning will lead to an unresolved replication stress which could increase the chances of mutation occurrence (Zeman and Cimprich, 2014). Such perturbations could impact the mutational catalog; thus, they could leave mutational imprints, and they could be considered as potential aetiologies.

Previous studies have identified an asymmetrical distribution of mutations along the strands. This asymmetry can be divided into transcriptional asymmetry observed mainly with mutations originating from UV lights and smoking, and a replicative asymmetry observed with POLE, APOBEC and MSI associated mutations (Nicholas et al., 2015).

Furthermore, studies in the team have shown that R-loops tend to form at the promoters and terminators of highly transcribed genes and when TOP1 is depleted this tendency increases at TTS, especially for converging genes (Promonet et al., 2020).

To the best of our knowledge there is no study that has addressed the issue of signatures associated with replication stress. Yet, there is mounting evidence that there could be different mutational processes operating during replication phase. All of this let us hypothesis that there might be signatures specific to replication stress.

The goal of this project is to search for mutational signatures associated with replication stress. More specifically, if there are different mutational processes operating at TSS and TTS of a given set of genes or if there is a mutational asymmetry between the two regions.

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J., & Stratton, M. R. (2013). Deciphering signatures of mutational processes operative in human cancer. *Cell reports*, *3*(1), 246–259. <https://doi.org/10.1016/j.celrep.2012.12.008>

Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., Wu, Y., Boot, A., Covington, K. R., Gordenin, D. A., Bergstrom, E. N., Islam, S., Lopez-Bigas, N., Klimczak, L. J., McPherson, J. R., Morganella, S., Sabarinathan, R., Wheeler, D. A., Mustonen, V., PCAWG Mutational Signatures Working Group, Getz, G., … PCAWG Consortium (2020). The repertoire of mutational signatures in human cancer. *Nature*, *578*(7793), 94–101. <https://doi.org/10.1038/s41586-020-1943-3>

Chin, L., Andersen, J. N., & Futreal, P. A. (2011). Cancer genomics: from discovery science to personalized medicine. *Nature medicine*, *17*(3), 297–303. <https://doi.org/10.1038/nm.2323>

Griffith, O. L., Spies, N. C., Anurag, M., Griffith, M., Luo, J., Tu, D., Yeo, B., Kunisaki, J., Miller, C. A., Krysiak, K., Hundal, J., Ainscough, B. J., Skidmore, Z. L., Campbell, K., Kumar, R., Fronick, C., Cook, L., Snider, J. E., Davies, S., Kavuri, S. M., … Ellis, M. J. (2018). The prognostic effects of somatic mutations in ER-positive breast cancer. *Nature communications*, *9*(1), 3476. <https://doi.org/10.1038/s41467-018-05914-x>

Hakem R. (2008). DNA-damage repair; the good, the bad, and the ugly. *The EMBO journal*, *27*(4), 589–605. <https://doi.org/10.1038/emboj.2008.15>

Huszno, J., Kołosza, Z., & Grzybowska, E. (2019). *BRCA1* mutation in breast cancer patients: Analysis of prognostic factors and survival. *Oncology letters*, *17*(2), 1986–1995. <https://doi.org/10.3892/ol.2018.9770>

ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020). Pan-cancer analysis of whole genomes. *Nature*, *578*(7793), 82–93. <https://doi.org/10.1038/s41586-020-1969-6>

Martin, P., Leighl, N. B., Tsao, M. S., & Shepherd, F. A. (2013). KRAS mutations as prognostic and predictive markers in non-small cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*, *8*(5), 530–542. <https://doi.org/10.1097/JTO.0b013e318283d958>

Haradhvala, N. J., Polak, P., Stojanov, P., Covington, K. R., Shinbrot, E., Hess, J. M., Rheinbay, E., Kim, J., Maruvka, Y. E., Braunstein, L. Z., Kamburov, A., Hanawalt, P. C., Wheeler, D. A., Koren, A., Lawrence, M. S., & Getz, G. (2016). Mutational Strand Asymmetries in Cancer Genomes Reveal Mechanisms of DNA Damage and Repair. *Cell*, *164*(3), 538–549. <https://doi.org/10.1016/j.cell.2015.12.050>

Olivier, M., & Taniere, P. (2011). Somatic mutations in cancer prognosis and prediction: lessons from TP53 and EGFR genes. *Current opinion in oncology*, *23*(1), 88–92. <https://doi.org/10.1097/CCO.0b013e3283412dfa>

Omichessan, H., Severi, G., & Perduca, V. (2019). Computational tools to detect signatures of mutational processes in DNA from tumours: A review and empirical comparison of performance. *PloS one*, *14*(9), e0221235. <https://doi.org/10.1371/journal.pone.0221235>

Promonet, A., Padioleau, I., Liu, Y., Sanz, L., Biernacka, A., Schmitz, A. L., Skrzypczak, M., Sarrazin, A., Mettling, C., Rowicka, M., Ginalski, K., Chedin, F., Chen, C. L., Lin, Y. L., & Pasero, P. (2020). Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. *Nature communications*, *11*(1), 3940. <https://doi.org/10.1038/s41467-020-17858-2>

Poduri, A., Evrony, G. D., Cai, X., & Walsh, C. A. (2013). Somatic mutation, genomic variation, and neurological disease. *Science (New York, N.Y.)*, *341*(6141), 1237758. <https://doi.org/10.1126/science.1237758>

Stratton, M. R., Campbell, P. J., & Futreal, P. A. (2009). The cancer genome. *Nature*, *458*(7239), 719–724. <https://doi.org/10.1038/nature07943>

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S. C., Kok, C. Y., Noble, K., Ponting, L., Ramshaw, C. C., Rye, C. E., Speedy, H. E., … Forbes, S. A. (2019). COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic acids research*, *47*(D1), D941–D947. <https://doi.org/10.1093/nar/gky1015>

Vuong, H., Cheng, F., Lin, C. C., & Zhao, Z. (2014). Functional consequences of somatic mutations in cancer using protein pocket-based prioritization approach. *Genome medicine*, *6*(10), 81. <https://doi.org/10.1186/s13073-014-0081-7>

Zeman, M. K., & Cimprich, K. A. (2014). Causes and consequences of replication stress. *Nature cell biology*, *16*(1), 2–9. <https://doi.org/10.1038/ncb2897>