Molecular Aspects of Medicine 69 (2019) 62–72

Contents lists available at ScienceDirect

Molecular Aspects of Medicine

journal homepage: www.elsevier.com/locate/mam

Somatic mutational signatures in polyposis and colorectal cancer

Judith E. Grolleman[a,1], Marcos Díaz-Gay[b,1], Sebastià Franch-Expósito[b], Sergi Castellví-Bel[b], Richarda M. de Voer[a,*]

[a] Department of Human Genetics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, the Netherlands
[b] Gastroenterology Department, Hospital Clínic de Barcelona, August Pi I Sunyer Biomedical Research Institute, CIBER of Hepatic and Digestive Diseases, University of Barcelona, Barcelona, Spain

ARTICLE INFO

Keywords:
Mutational process
Mutational signature
Polyposis
Colorectal cancer
Defective DNA repair
Genetic predisposition

ABSTRACT

The somatic mutation spectrum imprinted in the genome of a tumor represents the mutational processes that have been active in that tumor. Large sequencing efforts in various cancer types have resulted in the identification of multiple mutational signatures, of which several have been linked to specific biological mechanisms. Several pan-cancer mutational signatures have been identified, while other signatures are only found in specific tissue types. Research on tumors from individuals with specific DNA repair defects has led to links between specific mutational signatures and mutational processes. Studying mutational signatures in cancers that are likely the result of a genetic predisposition may represent an interesting strategy to identify constitutional DNA repair defects, including those underlying polyposis and colorectal cancer.

1. Introduction

Most cancers are characterized by multiple somatic mutations. These mutations can be classified into driver or passenger mutations according to their effects on cancer development (Stratton et al., 2009). Driver mutations are predominantly prioritized in most cancer sequencing studies because of the growth advantage that they confer, which causes their positive selection during cancer evolution (Stratton, 2011; Stratton et al., 2009). Passenger mutations have so far not been in the spotlight, essentially because they do not confer a selective advantage. However, passenger mutations are informative, since in particular the number of passenger and driver mutations allows to extract information both on the number of mitotic cell divisions that occurred in a cell lineage and on the mutation rate at each cell division (Stratton, 2011). In fact, passenger mutations are responsible of the emergence of a new exciting field of study. Based on the assumption that the patterns of these passenger mutations are invariable over time, these mutations can be used as a representative picture of the mutational mechanisms that are active during the tumorigenic process (Alexandrov et al., 2013b). Thus, the mutation pattern of driver and passenger mutations reflects the DNA damage and repair processes that cancer cells and their precursors underwent over time (Nik-Zainal et al., 2012).

Each specific mutational process leaves a particular imprint in the genome of a cell, also called *mutational signature* (Alexandrov et al., 2013a). Endogenous cellular mechanisms, such as DNA replication and repair, can generate mutations due to their intrinsic slight infidelity. However, mutations can also arise from exogenous mutagenic exposures. The final record of accumulated DNA damage is determined by the intensity and duration of all active mutational processes (Nik-Zainal et al., 2012). This genetic damage can emerge in the form of different classes of mutations, such us nucleotide substitutions, short insertions and deletions (indels), copy number alterations or structural variations. The current set of well-established mutational signatures consist of six different types of single nucleotide changes, based on the mutated pyrimidine base, including four transversions, C > A, C > G, T > A and T > G and two transitions, C > T and T > C, in context of the adjacent nucleotides. These six base substitutions, times the four possible preceding and posterior nucleotides, leaves a total of 96 possibilities. Thus, each mutational signature is composed by a particular distribution of these 96 potential trinucleotide mutations. This framework also allows that signatures composed by the same classes of substitutions, but in different sequence contexts, can be distinguished (Alexandrov et al., 2013a).

A formal mathematical approach is required in order to improve the quantification of the contribution of each process to the mutational profile in a specific cancer sample (Nik-Zainal et al., 2012). A theoretical model of mutational signatures has been built as a blind source separation problem and non-negative matrix factorization (NMF) was

* Corresponding author. Department of Human Genetics, Radboud university medical center, P.O. Box 9101, 6500, HB, Nijmegen, the Netherlands.
E-mail address: richarda.devoer@radboudumc.nl (R.M. de Voer).
[1] Shared first authors.

https://doi.org/10.1016/j.mam.2019.05.002
Received 14 January 2019; Received in revised form 13 May 2019; Accepted 16 May 2019
Available online 23 May 2019

implemented in order to define an appropriate computational framework (Alexandrov et al., 2013b). This unsupervised machine-learning algorithm is used to extract common features from multidimensional complex data and has primarily been used for face recognition and text mining (Berry et al., 2007; Lee and Seung, 1999). However, in recent years, NMF has been established as a common approach for different computational biologic applications (Devarajan, 2008), including mutational signature analyses.

The number of signatures extracted from mutational profiles of published cancer samples evolves as the number of these samples increases, i.e., the number of available genomes and their associated mutation burden mathematically constrain the number of signatures that can be retrieved by the model (Alexandrov et al., 2013b). A first attempt to extract somatic mutational signatures was performed on whole genome sequences (WGS) from 21 breast cancer samples. Five signatures were deciphered (Nik-Zainal et al., 2012), that subsequently were reduced to four after a further refinement of the approach and its computational implementation (Alexandrov et al., 2013b). Thereafter, in a seminal study, Alexandrov et al. (2013a, b) analyzed 4,938,362 somatic mutations from 7042 samples of 30 different cancer types, which led to the extraction of 21 different mutational signatures (Alexandrov et al., 2013a). Currently, 49 reference mutational signatures are extracted from a total of > 23,000 samples of most cancer types, including > 4500 whole cancer genomes (Alexandrov et al., 2018), of which 30 reference mutational signatures are indexed in the Catalogue of Somatic Mutations in Cancer (COSMIC) (Alexandrov et al., 2015a; Forbes et al., 2017).

The success of single-nucleotide signature studies has resulted from the possibility to identify multiple patterns created by different types of nucleotide changes in any given tumor type (Alexandrov et al., 2013b; Nesic et al., 2018). In contrast, the assessment of copy number alterations (CNA) only takes into consideration the absence (normal diploid state) or presence of duplications, deletions, unbalanced translocations or inversions. Efforts to identify signatures based on CNA are being made considering different approaches and accounting for diverse features, from plain subclassification and quantification of copy number events, to meta-feature studies in copy number profiles and data distributions (Macintyre et al., 2018; Nik-Zainal et al., 2016). The first six signatures based on CNA were generated using somatic data of 560 breast cancer genomes (Nik-Zainal et al., 2016). Interestingly, three of these copy number signatures represent rearrangement signatures, characterized by microhomology-mediated deletions that are associated with homologous recombination (HR) deficiency, and signatures for repeat-mediated deletions that are associated with mismatch repair (MMR) deficiency (Morganella et al., 2016). Macintyre et al. (2018) computed copy number signatures by analysis of shallow WGS data of 117 patients with high-grade serous ovarian cancer. Similar to the previous study, several of the copy number signatures were found to be linked to mutational processes such as DNA damage, cell cycle-failure or HR deficiency. Moreover, a positive survival correlation was found for three copy number signatures, while another copy number signature was selected as potential biomarker for a good prognosis and response to chemotherapy (Macintyre et al., 2018). Nonetheless, copy number signature analysis remains a computational challenge and, in order to be broadly implemented, consensus for their computation must be achieved. Therefore, in the near future, mutational signatures will likely not only consist of single nucleotide substitutions, but also incorporate different classes of mutations, such as dinucleotide substitutions, indels, and clustered mutations (Alexandrov et al., 2018). Additionally, expanded penta- and heptanucleotide contexts can also be used for point mutations, allowing the extraction of new specific mutational signatures (Alexandrov et al., 2018). In this case, not only the adjacent nucleotides are used for characterizing the mutation, but also the two or three possible preceding and posterior nucleotides.

Currently, several mutational signatures are linked to either one of the most common factors linked to mutational processes, such as aging, tobacco smoking, ultraviolet (UV) light exposure or defective MMR, whereas approximately half of the extracted signatures remains without an underlying etiology. Some mutational signatures are commonly found across different tumor types, while other signatures are predominant in tumors from unique tissue types. Here, we will review the current set of 30 single-nucleotide mutational signatures, which form the basis of most publications to date on mutational signatures. Furthermore, we hypothesize that mutational signature analysis can be used to discover (novel) polyposis and colorectal cancer (CRC) predisposition syndromes.

2. Mutational signatures associated with aging are present in normal and tumor tissues across tissue types

The process of aging is considered to be caused by a steady accumulation of DNA damage over time. DNA damage is generated by multiple mutational processes that are operative during the lifetime of a cell, which is reflected by the number of accumulated somatic mutations or epigenetic marks. Mutational signature analysis can be used to define the role and extent of aging in many different tissue types of individuals with a different age, gender, ethnicity and lifestyle. This analysis may facilitate our understanding of the mutational processes underlying aging in normal and cancerous cells.

Two of the initially identified mutational signatures, that were associated with aging in later studies, are signatures 1 and 5. Both signatures are observed across many tissue types (Alexandrov et al., 2015a). The first mutational signature, signature 1, is characterized by C > T mutations at CpG dinucleotides (Alexandrov et al., 2013b; Nik-Zainal et al., 2012), and is suggested to be the result of spontaneous deamination of 5-methylcytosine (5 mC) to thymine (Alexandrov et al., 2013a; Pfeifer, 2006). Alexandrov et al. (2015a, b) revealed a correlation between signature 1 and the number of mutations at the age of cancer diagnosis. It has been suggested that the deaminated 5 mC at CpG sites are not repaired before DNA replication, which leads to a higher number of mutations in each subsequent replication round (Alexandrov et al., 2015a). Hence, signature 1 is called a 'clock-like' signature. Moreover, signature 1 is found at varying levels in different tissue types, which has led to the hypothesis that cell types with high mitotic rates, such as epithelial cells of the colon, exhibit higher mutation rates as a result of frequent DNA replication without the repair of 5 mC deamination (Alexandrov et al., 2015a; Blokzijl et al., 2016). Mutations attributable to signature 1 are observed in paternal germ cells and normal adult stem cells (Blokzijl et al., 2016; Rahbari et al., 2016), thereby underscoring the underlying clock-like mutational mechanism of signature 1. Thus, signature 1 is suggested to be caused by a steady accumulation of mutations during a cells lifetime, both in germline and soma, indicating an underlying aging mechanism. Very recently, germline loss of *MBD4* has been identified as an initiator of 5 mC dependent hypermutation (Rodrigues et al., 2018; Sanders et al., 2018), suggesting a possible role for *MBD4* in the mutational mechanism underlying mutational signature 1. Other cytosine modifications can also be associated with specific patterns of somatic mutations, as is the case in cytosine hydroxymethylation, linked to an increase in C > G mutations (Supek et al., 2014). Thus far, none of the signatures described have been associated with this process, although it may be a mechanism potentially involved in a future deciphered signature or associated with one of unknown etiology.

The second mutational signature associated with aging, signature 5, is a relatively featureless 'flat' signature, with mutations distributed over all 96-mutation possibilities and a transcriptional strand bias for T > C transitions in an ApTpN context (Alexandrov et al., 2013a). The biological process that leads to the mutation types observed in signature 5 is not yet well understood. Similar to signature 1, a correlation has been observed between signature 5 and the number of mutations at the age of cancer diagnosis (Alexandrov et al., 2015a). A recent study extracted a novel signature (signature 40) that is highly similar to

signature 5, which also correlates with age (Alexandrov et al., 2018). Signature 5 is found in paternal germ cells, which indicates that mutations attributable to signature 5 may accumulate with age (Rahbari et al., 2016). However, in urothelial cell carcinomas, the amount of mutational signature 5 is associated with somatic non-silent *ERCC2* mutations (Kim et al., 2016), possibly indicating that diverse mechanisms underlie signature 5. While both signatures 1 and 5 are associated with aging, the signatures themselves do not closely correlate with each other (Alexandrov et al., 2015a). The lack of a correlation between signature 1 and 5 indicates that the process underlying the somatic mutations associated with signature 5 is different from the mechanism underlying signature 1 (Alexandrov et al., 2015a). Thus, signature 5 is associated with aging in a clock-like manner that is yet to be elucidated.

Three other mutational signatures associated with the age of cancer diagnosis have been identified in liver cancer. Letouzé et al. (2017) found a strong correlation between age of diagnosis and signatures 4, 12 and 16 in liver tumors. However, it has been suggested that this correlation may be based on the use of tobacco and alcohol over the years, which affects the liver, rather than aging itself (Letouzé et al., 2017). In summary, the mutational processes underlying the aging-related signatures 1 and 5 appear to be active in the germline as well as in (normal) somatic tissues (Blokzijl et al., 2016; Rahbari et al., 2016) and may be observed across various cancer types (Alexandrov et al., 2015a). Whereas signature 1 is probably influenced by the proliferation rate of a cell, which explains the differences in signature 1 contribution across cancer types, signature 5 may be activated by non-clock-like factors (Alexandrov et al., 2015a). It remains to be elucidated whether the prominent presence of one of the aging-related signatures in a tumor genome may be indicative of a sporadic origin, which can be of clinical relevance to the patient.

3. Tissue-specific mutational processes and exogenous mutagens associated with signatures

Different tissues present different cell replacement turnover ratios (Alexandrov et al., 2015a). As mentioned above, the aging signatures 1 and 5 are found across tissue types with varying contributions. Additionally, specific endogenous (e.g. DNA replication or repair deficiencies) and exogenous (e.g. lifestyle habits or anti-cancer treatments) factors may vary widely between tissue types, resulting in the presence of tissue-specific mutational signatures.

3.1. Mutational signatures in sporadic forms of gastrointestinal cancer

In sporadic CRCs, a number of mutational signatures have hitherto been validated. Next to the signatures that correspond to the process of aging (signatures 1 and 5), in a substantial proportion of sporadic CRCs signatures 6 and 10 have been observed, which are associated with defective MMR and defective exonuclease activity of polymerase epsilon (*POLE*), respectively (The Cancer Genome Atlas, 2012). Recently, additional signatures (20 and 26) have been associated with somatic MMR defects and these are often found in a combination with signature 6 in the same samples (Nagahashi et al., 2016). More recently, other signatures have been identified in sporadic CRCs with an as yet unknown etiology (Roerink et al., 2018) that require further confirmation. Next to these signatures, several signatures have been identified in tumors originating from the colorectum that result from constitutional DNA repair deficiencies, which will be described in more detail in section 4.

Apart from CRCs, signatures in other gastrointestinal cancers are widely studied, such as in liver, esophageal and head and neck cancers. Liver cancers exhibit a set of signatures commonly found in other cancers, such as those associated with aging (with a particular dominant profile of signature 5), tobacco smoking (signature 4), alkylating chemotherapy (signature 11) and MMR (signatures 6) (Dow et al., 2018; Letouzé et al., 2017; Ng et al., 2017; Schulze et al., 2015). In addition, other signatures have been found to be mainly specific for liver cancer, which are mostly associated with exogenous mutagens affecting the liver. Signature 24, which is associated with aflatoxin exposure (Dow et al., 2018; Letouzé et al., 2017; Ng et al., 2017; Schulze et al., 2015), and platinum compound chemotherapy signatures (Boot et al., 2018; Liu et al., 2017) are mainly found in the liver. Signature 22, linked to aristolochic acid exposure, has thus far only been found in liver and urothelial carcinomas (Ng et al., 2017; Nik-Zainal et al., 2015; Poon et al., 2015). Interestingly, although signature 16 was previously thought to be exclusive for liver cancer (Alexandrov et al., 2015a; Letouzé et al., 2017; Schulze et al., 2015), it has recently been associated with alcohol consumption in esophageal and head and neck squamous cell carcinomas (Chang et al., 2017; Li et al., 2018). This notion strengthens the observation that this signature reflects the mutations induced by tissue-specific exposure to alcohol. Thus far, several signatures associated with exogenous mutagen exposure have been observed in gastrointestinal cancers, mainly in the liver. A recent analysis of signatures involving clustered mutations suggested that some of these exogenous exposures, mainly alcohol, could exert their influence not directly by increasing the mutation rate but by redistributing the mutations to key parts of the genome (Supek and Lehner, 2017).

3.2. Mutational signatures in tumors developed in tissues exposed to exogenous mutagens

Lung cancer and melanoma have been the first cancer types in which the somatic mutation spectra have been associated with mutational processes, like exposure to benzo [a]pyrene, a compound in tobacco smoke, and UV light, respectively (Pleasance et al., 2010a, 2010b). These exogenous mutagens have been linked to mutational signatures 4 and 7, respectively (Alexandrov et al., 2013a). Next to melanomas, signature 7 is prominently found in head and neck squamous cell carcinomas (Alexandrov et al., 2013a; Nik-Zainal et al., 2015). Additional mutational signatures commonly observed in melanoma are signatures linked to aging (signatures 1 and 5) and APOBEC protein family members (signatures 2 and 13) (Robles-Espinoza et al., 2016). Signature 11 is thought to be associated with alkylating chemotherapy and has, as yet, only been found in liver cancers, malignant melanomas and glioblastomas (Alexandrov et al., 2013a; Campbell et al., 2017; Dow et al., 2018; Robles-Espinoza et al., 2016). It was first thought that signature 11 was associated with temozolomide treatment. A recent paper, in which induced pluripotent stem cells (iPSCs) are exposed to several exogenous mutagens, shows that temozolomide treatment induces a completely different signature and thus suggests that signature 11 is likely caused by another alkylating agent (1,2-dimethylhydrazine) (Kucab et al., 2019). This study demonstrates that functional validation of a proposed mutational process underlying an associated mutational signature(s) is necessary.

The presence of melanoma-specific signatures differs between subtypes. Mucosal and acral melanoma subtypes are mainly dominated by the aging and APOBEC-associated signatures, which are shared by most cancer types, whereas cutaneous melanoma (influenced by sunlight exposure) is mainly characterized by the UV-associated mutational signature 7 (Hayward et al., 2017). Interestingly, signature 7 is also observed in a set of lung cancers. These cancers are now thought to represent lung metastases derived from skin carcinomas, which have erroneously been classified as lung squamous cell carcinomas (Campbell et al., 2016). This observation exemplifies the great power of mutational signature analysis in tumor classification.

Interestingly, the presence of a specific mutational signature may have clinical implications. Chronic exposure to the immunosuppressive drug azathioprine (signature 32), which is used to treat conditions like arthritis and inflammatory bowel syndrome, is only found in cutaneous squamous cell carcinomas (Inman et al., 2018). The basis for this signature is proposed to result from the combined action of azathioprine and UV light exposure (Inman et al., 2018). These results demonstrate

the usefulness of mutational signature analyses in clinical practice, as advice on sun protection, skin surveillance, early diagnosis and lesion removal can now be offered to patients treated with azathioprine.

Signature 4 has been encountered in several subtypes of lung cancer: small-cell-, squamous cell-, neuroendocrine-, and adeno-carcinomas (LUAD) (Campbell et al., 2016; George et al., 2018). Noteworthy, it has been found that LUAD patients without a smoking history present with a higher contribution of aging-related signature 5 compared to LUAD patients with a smoking history (Campbell et al., 2016). This observation implicates that the presence of an aging-related signature indicates a sporadic origin of the tumor, and that the presence of signature 4 indicates a driving role of DNA damage induced by tobacco smoking. Besides lung cancer, signature 4 is mainly observed in cancers of the larynx, oral cavity, pharynx, esophagus, head and neck and liver (Alexandrov et al., 2013a; Jia et al., 2014; Nik-Zainal et al., 2015; Schulze et al., 2015). Interestingly, while smoking is associated with an elevated risk for CRC, 'smoking' signature 4 is not found in CRCs. Therefore, it is likely that other compounds than benzo [a]pyrene in tobacco smoke are contributing to the risk for developing CRC, possibly causing another signature that is yet to be determined.

4. The role of somatic and germline DNA repair defects on mutational signatures

DNA damage that accumulates in a cell lineage over time can be of exogenous or endogenous origin. Of the endogenous processes, DNA repair machineries play a crucial role in the maintenance of DNA integrity and genomic stability. The mutational signatures associated with DNA repair defects can both be found in tumors with somatic or germline DNA repair deficiencies, the latter of which are often linked to cancer predisposition syndromes. Currently, a large proportion of known mutational signatures has been linked to specific DNA repair defects.

As mentioned in section 3, DNA repair defects caused by the APOBEC protein family are associated with mutational signatures 2 and 13. Both signatures show lagging strand bias (Alexandrov et al., 2013a, 2013b) and are marked by regional hypermutation events called *kataegis* (Nik-Zainal et al., 2012). The APOBEC-mediated repair system induces the same type of DNA damage as is observed in signatures 2 and 13, and APOBEC3A and APOBEC3B activities have been correlated with these signatures (Alexandrov et al., 2013a; Jia et al., 2014; Nik-Zainal et al., 2012, 2014; Taylor et al., 2013). Although the reason for increased APOBEC activity in cancers remains unknown, it is hypothesized that signatures 2 and 13 represent collateral DNA damage from a response that is originally directed at retrotransposing DNA elements (Alexandrov et al., 2013a; Petljak et al., 2019). The APOBEC3A/B-related signatures are found in many tissue types, but are most commonly observed in bladder and cervical cancers (Alexandrov et al., 2013a). Moreover, APOBEC3A/B expression positively correlates with a higher mutation burden and an enhanced immune response in urothelial carcinomas (Glaser et al., 2018). Starrett et al. describe that a combination of APOBEC3B and APOBEC3H haplotype I underlies signatures 2 and 13 in breast and lung cancer (Starrett et al., 2016), while another recent study proposes APOBEC3A to be responsible for the APOBEC-associated signatures. Moreover, the latter study showed that APOBEC-associated mutagenesis can be episodic *in vitro* (Petljak et al., 2019). The exact mechanism of the APOBEC-associated signatures remains to be elucidated.

Another DNA repair defect that is associated with unique mutational signatures is caused by mutations in the *BRCA1* and *BRCA2* genes, which are associated with signature 3 and, to a lesser extent, signature 8 (Nik-Zainal et al., 2016). Next to a single nucleotide pattern of mutation, signature 3 has been found to be associated with a substantial number of (> 3bp) deletions with microhomology at rearrangement breakpoints, whereas signature 8 has been found to be associated with double nucleotide substitutions, especially CC > AA (Nik-Zainal et al., 2012). Due to the observed mutation types and association with *BRCA1/2* mutations, the mutations attributable to signature 3 are thought to result from a failure of DNA double-strand break-repair by HR (Nik-Zainal et al., 2012). In an independent study, signature 3 has been associated with defects in other genes of the HR machinery that are linked to breast cancer, such as mutations in *PALB2* or hypermethylation of the *RAD51C* gene promoter, but not in *ATM* or *CHEK2* (Polak et al., 2017). Mutational signature 3 is found in breast, ovarian, pancreatic, gastric and esophageal cancer, while *BRCA1/2* mutations have also been associated with an increased prostate cancer risk (Alexandrov et al., 2013a, 2015b; Connor et al., 2017; Secrier et al., 2016; Waddell et al., 2015). Recently, a tool (HRDetect) has been reported that enables the detection of HR deficient tumors based on the presence of signatures 3 and 8, in combination with other indel and structural variant features (Davies et al., 2017). The presence of the HR signatures has potential clinical implications linked to PARP inhibitor treatment responses (Davies et al., 2017; Fong et al., 2009).

Next to the previous two DNA repair defects that are associated with mutational signatures, multiple signatures have been found to be specifically associated with DNA repair defects that are involved in CRC predisposition. A large study on colon and rectal cancers from the Cancer Genome Atlas (TCGA) revealed two distinctive groups, namely hypermutated and non-hypermutated tumors. The non-hypermutated group included mostly tumors with no or low microsatellite instability (MSI) and microsatellite stable (MSS) tumors. The hypermutated tumors mostly consisted of tumors with high levels of MSI, due to *MHL1* promoter methylation or mutations in one of the MMR genes, or tumors that harbored somatic mutations in the exonuclease domain of *POLE* (The Cancer Genome Atlas, 2012). Mutational signatures 6, 15, 20 and 26 are associated with MMR deficiency. All four signatures have been found to be associated with a high number of indels at nucleotide repeats, which is in concordance with the observed MSI caused by defects in MMR genes (Alexandrov et al., 2013a, 2015a). Germline mutations in one of the MMR genes (*MLH1, MSH2, MSH6, PMS2*) are known to predispose to CRC, a genetic condition called Lynch syndrome (Leach et al., 1993; Ligtenberg et al., 2009; Miyaki et al., 1997; Nicolaides et al., 1994; Papadopoulos et al., 1994). Moreover, somatic mutations in *MLH1* and *MSH2* are a frequent cause of MMR deficiency in Lynch syndrome-like tumors that lack causal germline mutations and/or promoter hyper-methylation (Mensenkamp et al., 2014). Signature 6 is mainly associated with germline mutations in one of the MMR genes, while signatures 15, 20 and 26 are associated with somatic MMR deficiency (Alexandrov et al., 2015a). Signature 6 is observed in many tissue types (Alexandrov et al., 2013a), including CRC. Signature 15 has, as yet, only been detected in stomach and lung cancer (Alexandrov et al., 2013a), while signatures 20 and 26 have, next to CRC, among others been observed in breast and stomach cancer (Alexandrov et al., 2015a; Drost et al., 2017; Nagahashi et al., 2016).

The polymerase proofreading genes *POLE* and *POLD1*, when their exonuclease domain is mutated in the germline, predispose to polyposis and CRC (Palles et al., 2013). Defective *POLE* causes a hypermutated cancer genome and is strongly associated with signature 10 (Alexandrov et al., 2013a; The Cancer Genome Atlas, 2012). *POLE* mutations are known to induce C > A transversions in a TpCpT > TpApT context (Church et al., 2013; Shinbrot et al., 2014). Moreover, a replication strand bias has been observed for C > A mutations at TpCpT context and T > G mutations at TpTpT context (Shinbrot et al., 2014). A recent study showed that somatic *POLE* mutations are likely initiating events in CRC and endometrium cancer development, as the mutation types associated with signature 10 have been observed in cancer driver genes (Temko et al., 2018). Similar findings were made in a simultaneous study in *POLE*-mutated uterus- and colon tumors, in which both *POLE* and *PTEN* mutations were predicted to influence early tumorigenesis (Hatakeyama et al., 2018). Besides the colon, signature 10 is observed in cancers that developed in breast, bladder, uterus and cervix (Alexandrov et al., 2013a). Thus far, no clear signature has been

described for *POLD1* mutated CRCs.

Interestingly, tumors with concurrent activity loss of *POLE* or *POLD1* and one of the MMR genes display a unique mutational signature that is not simply the sum of the signatures of the individual DNA repair defects (Campbell et al., 2017; Haradhvala et al., 2018). Tumors with activity loss of one of the MMR genes (e.g. *MLH1*) in combination with somatic *POLE* exonuclease mutations are associated with signature 14 (Campbell et al., 2017; Castellsague et al., 2018; Haradhvala et al., 2018), whereas MMR deficiency in combination with a somatic *POLD1* exonuclease mutation is associated with signature 20 (Campbell et al., 2017; Haradhvala et al., 2018). Signature 20 has also been observed in a *MLH1* knockout organoid and is, thus, associated with MMR deficiency by *MLH1* inactivation (Drost et al., 2017). It remains to be elucidated whether signature 20 is the result of mutations attributable to *MLH1* inactivation alone, or any MMR deficiency in combination with *POLD1* exonuclease mutations. These observations demonstrate that multiple distinct mutational signatures may result from different combinations and/or orders of mutational processes.

Next to the hypermutated CRCs, some of the non-hypermutated MSS CRCs have also been found to be induced by DNA repair defects. At present, two distinct mutational signatures have been associated with mutations in one of the base-excision repair (BER) genes. The first BER gene associated with a unique mutational signature was *MUTYH*. *MUTYH* is a DNA glycosylase encoding gene playing a fundamental role in the repair of oxidative DNA damage and of which biallelic germline inactivation leads to the well-known predisposition syndrome *MUTYH*-associated polyposis (MAP) (Al-Tassan et al., 2002). Two studies have shown that MUTYH deficiency leads to a specific signature that is characterized by C > A mutations, mainly in a CpA context (Pilati et al., 2017; Viel et al., 2017). The *MUTYH* signature was first identified as signature 18 in human CRCs (Pilati et al., 2017). Thereafter, others associated MUTYH deficiency in CRCs and CRC stem cells with signature 36 (Viel et al., 2017). The mutation patterns of signature 36 closely resembles signature 18 (Pearson correlation coefficient of 0.77) (Viel et al., 2017). Signature 18 has also been found in *in vitro* cell models, possibly as the result of induced oxoG damage (Blokzijl et al., 2016; Drost et al., 2017). It remains to be determined whether signatures 18 and 36 are truly two distinct signatures, and which of the two is truly correlated with MUTYH deficiency.

The second BER gene associated with a unique mutational signature is *NTHL1*. NTHL1 deficiency predisposes to polyposis and CRC, and is characterized by C > T mutations at non-CpG sites (Drost et al., 2017; Grolleman et al., 2019; Weren et al., 2015). Deletion of *NTHL1* using CRISPR-Cas9 in colon organoids revealed a causal link between NTHL1 deficiency and mutational signature 30, as the accumulation of mutations in these organoids revealed a mutational footprint that is similar to signature 30 (Drost et al., 2017). Signature 30 was previously identified in a single breast cancer case, where it was the only signature that could explain the majority of mutations in this tumor (Alexandrov et al., 2013a; Nik-Zainal et al., 2016). Retrospective analysis of that single breast cancer sample revealed an *NTHL1* germline nonsense mutation with loss of heterozygosity (LOH) in the tumor (Drost et al., 2017; Nik-Zainal et al., 2016). Very recently, Grolleman et al. (2019) found that, next to polyposis and CRC, individuals with biallelic germline *NTHL1* mutations also develop extracolonic tumors (Grolleman et al., 2019). Using mutational signature analysis they revealed that signature 30 underlies the main mutational process in all but one of the tumors studied (93%). The one tumor without signature 30 was an urinary cell carcinoma (UCC) in which signature 2 (commonly observed in sporadic UCCs) was the most prominent signature, which suggests that this tumor developed sporadically (Grolleman et al., 2019). These data provide a proof-of-principle for a novel approach in which mutational signatures in tumors can be used as a tool to corroborate a genetic predisposition.

5. Software and web tools available to perform mutational signature analysis

The mutational signatures described in sections 2, 3 and 4 have been extracted by employing various software packages. In recent years, different software packages and web-based applications have been released, in order to practically implement mutational signature analyses even without a broad bioinformatic expertise (for a comprehensive review on this topic see (Baez-Ortega and Gori, 2019). It is important to distinguish between tools that allow the deciphering of *de novo* signatures for a specific set of samples from those that reconstruct the known mutational profiles using a collection of well-established mutational signatures. Most studies use the former strategy, since it is essential to decipher which signatures are contributing to the mutational catalogue of every specific tissue types.

With respect to the extraction of *de novo* mutational signatures, the original Wellcome Trust Sanger Institute's (WTSI) computational framework (SigProfiler) (Alexandrov et al., 2013b), used in most of the studies presented in this review, is available at http://www.mathworks.com/matlabcentral/fileexchange/38724-sigprofiler. This framework, based on NMF, has been developed in MATLAB, which is a proprietary software that limits widespread use. Therefore, different software packages are being built using open source programming languages and platforms. The R programming language is used by SomaticSignatures (Gehring et al., 2015), MutationalPatterns (Blokzijl et al., 2018), mutSignatures (Fantini et al., 2018), Palimpsest (Shinde et al., 2018) and Maftools (Mayakonda et al., 2018), whereas the MutSpec toolbox uses the well-known Galaxy platform (Ardin et al., 2016). Alternatively, other computational approaches are being developed using different strategies outside the NMF paradigm or adding by additional concepts, such as the probabilistic models of both EMu (Fischer et al., 2013) and pmsignature (probabilistic mutation signature) (Shiraishi et al., 2015), the Bayesian variants of NMF of SignatureAnalyzer (Alexandrov et al., 2018; Kasar et al., 2015; Kim et al., 2016) and signeR (Rosales et al., 2017).

A different computational challenge arises with respect to the reconstruction of mutational profiles using a set of known signatures. The first available tool of this class was deconstructSigs (Rosenthal et al., 2016), which has been widely used since its publication (Bruna et al., 2016; Goh et al., 2016; Hao et al., 2016; Kanu et al., 2016; Nagahashi et al., 2016) (Table 1). Additionally, MutationalPatterns allows for *de novo* deciphering as well as signature reconstruction (Blokzijl et al., 2018), which is used to extract the contributions of known signatures in some recent reports (Blokzijl et al., 2016; Drost et al., 2017). DeconstructSigs and MutationalPatterns are applied in the user friendly web-based interactive applications in mSignatureDB (Huang et al., 2018a) and MuSiCa (Diaz-Gay et al., 2018), respectively, which are suitable for non-specialized bioinformatic researchers. In addition, some other web tools and software packages for signature reconstruction have recently been developed, including MutaGene (Goncearenco et al., 2017), Mutalisk (Lee et al., 2018), SignatureEstimation (Huang et al., 2018b), decompTumor2Sig (Krüger and Piro, 2019), and SigProfilerSingleSample which is implemented within the SigProfiler package (Alexandrov et al., 2018) (Table 1).

In the near future, the possibility to perform a sample by sample analysis according to a set of consensus signatures may become of interest, especially considering potential clinical applications of this methodology where computational resources and timely processing may represent limiting factors. Moreover, several user friendly webtools are available that allow for *de novo* signature analyses, which makes mutational signature analysis accessible for non-bioinformaticians as well.

Table 1
Available software packages and (online) tools for the analysis of mutational signatures.

Software	*De novo* signature extraction	Profile reconstruction based on defined signatures	Web-based graphical user interface	Mathematical framework	Implementation	Reference
SigProfiler (WTSI computational framework)	Yes	Yes (SigProfilerSingleSample)	No	NMF	MATLAB	Alexandrov et al. (2018); Alexandrov et al. (2013a,b)
Somatic Signatures	Yes	No	No	NMF	R	Gehring et al. (2015)
MutationalPatterns	Yes	Yes	No	NMF/Non-negative least squares	R	Blokzijl et al. (2018)
mutSignatures	Yes	No	No	NMF	R	Fantini et al. (2018)
Palimpsest	Yes	Yes	No	NMF	R	Shinde et al. (2018)
Maftools	Yes	No	No	NMF	R	Mayakonda et al. (2018)
MutSpec	Yes	No	Yes	NMF	Galaxy	Ardin et al. (2016)
EMu	Yes	No	No	Probabilistic	C++	Fischer et al. (2013)
pmsignature (probabilistic mutation signature)	Yes	No	Yes	Probabilistic	R, C++	Shiraishi et al. (2015)
signeR	Yes	No	No	Bayesian NMF	R, C++	Rosales et al. (2017)
SignatureAnalyzer	Yes	Yes	No	Bayesian NMF	R	Alexandrov et al. (2018); Kasar et al. (2015); Kim et al. (2016)
deconstructSigs	No	Yes	No	Multiple linear regression	R	Rosenthal et al. (2016)
mSignatureDB	Yes	Yes	Yes	NMF/Multiple linear regression	R, Shiny	Huang et al. (2018a)
MuSiCa	No	Yes	Yes	Non-negative least squares	R, Shiny	Diaz-Gay et al. (2018)
MutaGene	No	Yes	Yes	NMF/non-smooth NMF/Non-negative least squares	Python	Goncearenco et al. (2017)
Mutalisk	No	Yes	Yes	Linear regression/Multinomial test	R, PHP	Lee et al. (2018)
SignatureEstimation	No	Yes	No	Quadratic programming/Simulated annealing	R	Huang et al. (2018b)
decompTumor2Sig	No	Yes	No	Quadratic programming	R	Krüger and Piro (2019)

NMF: non-negative matrix factorization; WTSI: Wellcome Trust Sanger Institute.

6. Mutational signatures as a tool for the identification of a genetic predisposition for colorectal cancer

At present, in only 20–30% of individuals suspected of hereditary polyposis and/or CRC a genetic diagnosis is made. The most frequent occurring CRC predisposition syndrome with a Mendelian inheritance is Lynch syndrome (see above). Other polyposis and CRC predisposition syndromes in which DNA replication or repair is impaired are MAP, polymerase proofreading associated polyposis (PPAP) and, more recently, *NTHL1*-associated tumor syndrome (NATS) (Al-Tassan et al., 2002; Grolleman et al., 2019; Palles et al., 2013; Weren et al., 2015). Epidemiologic studies suggest that novel polyposis and CRC predisposition syndromes remain to be discovered, but their prevalence is likely rare, which hampers their identification. We propose mutational signature analysis to facilitate and support the detection of novel genetic predispositions to colorectal (and other) cancers, even in absence of overt clinical parameters.

Mutational signatures with a known etiology can be used to identify somatic or germline defects that have played a role in the development of a particular tumor. As described in section 4, mutational signatures can be observed across cancer types or, conversely, are tissue specific. Six signatures have been observed in sporadic CRCs (Fig. 1), which are associated with aging (signatures 1 and 5), somatic inactivation of MMR genes (signatures 6, 20 and 26) and somatic *POLE* deficiency (signature 10) (Alexandrov et al., 2013a; Jia et al., 2014; Nagahashi et al., 2016). More recently, mutational signatures have been identified in tumors from individuals with a germline predisposition syndrome. Several tumors from different patients with biallelic germline mutations in either one of the MMR genes, *POLE*, *MUTYH* or *NTHL1*, have been identified and a main contribution of the mutational signatures associated with each genetic defect has been revealed (Campbell et al., 2017; Castellsague et al., 2018; Grolleman et al., 2019; Viel et al., 2017). We hypothesize that individuals suspected of hereditary polyposis and/or CRC harbor known and novel mutational signatures hidden in their tumor genomes, that can be used to identify the underlying (genetic) cause (Fig. 1).

As we are able to detect somatic and/or germline defects in known CRC predisposition genes based on mutational signature analysis, this approach can be used to detect CRCs that have developed due to genetic predisposition in a cohort of unexplained CRCs. In such a cohort, next to signatures associated with MMR deficiency, signatures 10, 18/36 and 30, which are associated with PPAP, MAP and NATS, respectively, are anticipated. Subsequent analysis of tumor and germline data may reveal pathogenic variants in the corresponding gene. Besides the detection of known pathogenic variants, the identification of a mutational signature associated with a DNA repair defect can provide evidence for the pathogenicity of a novel variant detected in that patient, as has been shown for *POLE* (Castellsague et al., 2018). Interestingly, the identification of signatures that have, as yet, not been associated with CRC can provide a first hint towards a genetic diagnosis of a patient. The

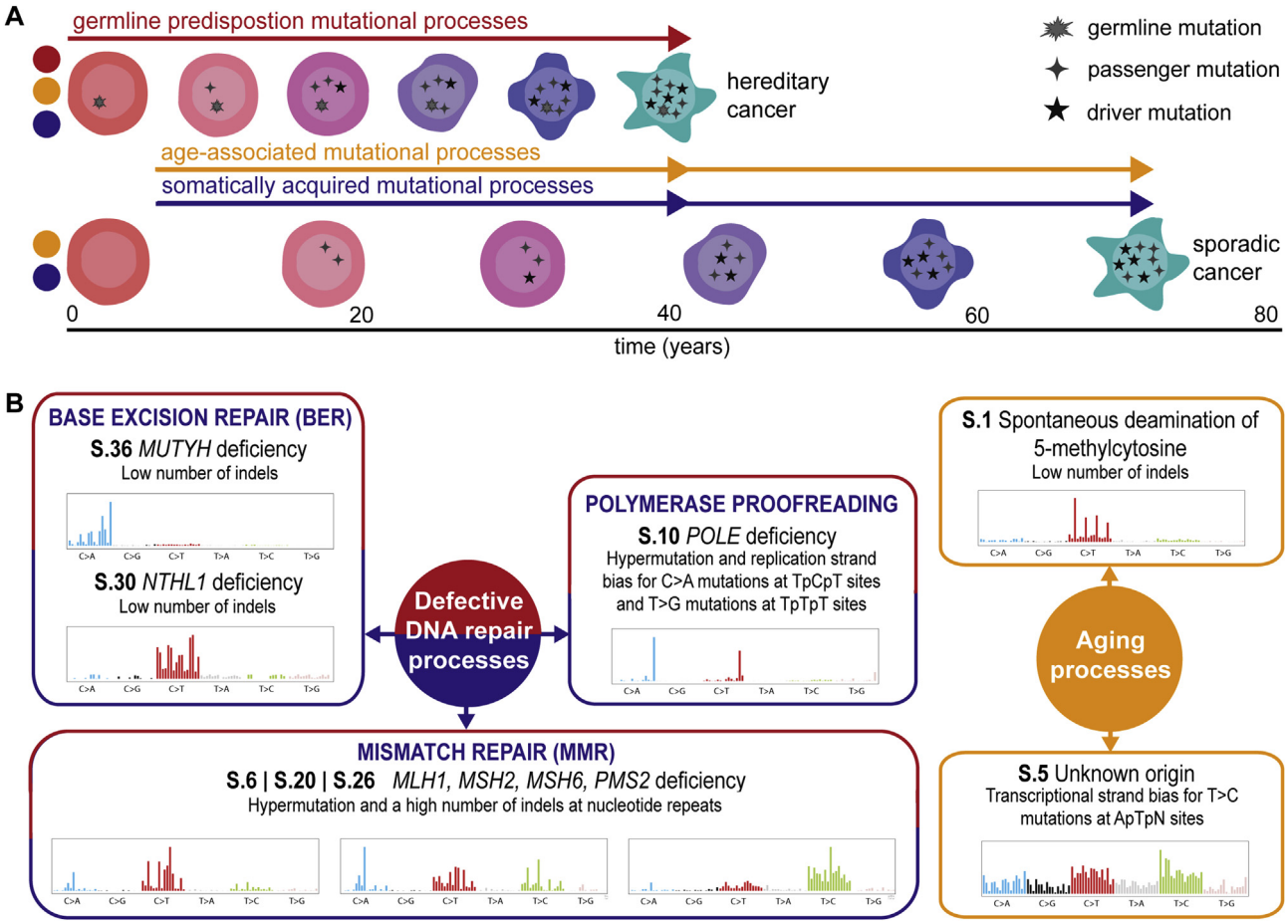

Fig. 1. Mutational processes and signatures in hereditary and sporadic colorectal cancer (CRC). A) Several mutational processes are involved in the development of CRC. Hereditary cancers develop more rapidly than sporadic cancers due to a genetic predisposition (in red). Age-associated (in yellow) and somatically acquired (in blue) mutational processes may have an effect on the development of both hereditary and sporadic CRCs. B) Eight mutational signatures observed in CRC that are associated with either defective DNA repair or aging-associated mutational processes. The mutational signatures resulting from defective DNA repair may have a somatic or a germline origin. Colours correspond to the colours related to the mutational processes in panel A. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

identification of signature 3 (HR) in CRCs may, for example, be related to a defect in *PALB2* (Polak et al., 2017), which has recently been suggested as a CRC predisposition gene (AlDubayan et al., 2018).

Besides these known mutational signatures, the identification of a novel mutational signature that is shared between multiple individuals suggests a common process (i.e., genetic predisposition) underlying its development. To detect such signatures, a preselected cohort with colon tumors from multiple individuals with a suspected shared hereditary cause based on their clinical characteristics (e.g. age of tumor diagnosis, polyp/cancer type, multiple primary tumors) may be used. The pattern of somatic mutations attributable to that novel identified mutational signature may provide insight in the putative function of the (as yet to be determined) underlying genetic defect, similarly to the link between signatures 2 and 13 and APOBEC3A/B activity (Nik-Zainal et al., 2012). The association between signatures 2 and 13 and APOBEC3A/B has been functionally validated in yeast and human, and a recent study characterized APOBEC signatures in cancer cell lines and xenografts (Petljak et al., 2019).

Several other mutational signatures have also been functionally validated in *in vitro* models. For example, the association between > 70 exogenous mutagens and mutational signatures in iPSCs (Kucab et al., 2019), and the association between the MMR signatures and *MLH1, PMS2* and *MSH6* defects in human colon organoids, *C. elegans* and haploid cells, respectively (Drost et al., 2017; Meier et al., 2018; Zou et al., 2018). Especially for rare genetic disorders, for which the number of available tumors and/or patients is low, functional validation of putative associations between genetic defects and unique mutational signatures to prove causality is warranted.

7. Conclusions and perspectives

Signatures associated with aging are found across many tumor types, while mutational signatures associated with germline DNA repair defects and exogenous mutagen exposure are mainly tissue specific. Furthermore, metastases of distinct tissue types can be distinguished from second primary tumors based on the presence of particular signatures. Similarly, therapy-induced tumors can be identified based on known associations between mutational signatures and specific therapeutics. It has been shown that tumors with a sporadic origin can be distinguished from hereditary tumors based on the observed mutational signatures. Thus, the existing DNA repair signatures can assist in the recognition of patients with known DNA repair predisposition syndromes, as WES or WGS of tumors will likely become standard in the diagnosis of patients with a suspected genetic predisposition in the years to come. Moreover, it is anticipated that the therapeutic value of mutational signatures will increase to predict therapy responses. For example, the tool HRDetect enables the detection of HR deficient tumors and, thus, to identify patients that may benefit from PARP inhibitor treatment (Davies et al., 2017; Fong et al., 2009). In case of CRC, it is already known that patients with hypermutated tumors respond well to immune checkpoint inhibitors (Llosa et al., 2015). Hypermutated tumors are currently recognized based on their MSI status, but MSS tumors with a high tumor mutational burden (TMB) may be missed. Several mutational signatures have been linked to hypermutation, thereby facilitating the recognition of tumors with a high TMB. Additionally, the underlying cause of a high TMB in CRCs (often MMR or *POLE* defects) may be used to provide personalized treatment or predict treatment outcomes. Therefore, we expect that mutational signature analysis will not only become a powerful tool for the identification of a genetic predisposition, but also for the design of personalized treatment strategies for CRC patients.

Declarations of interest

None.

Acknowledgements

We thank prof. dr. Ad Geurts van Kessel for critical reviewing of the manuscript. MDG is supported by a contract from Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR; Generalitat de Catalunya, 2018FI_B1_00213). JEG is supported by a grant from the Dutch Cancer Society (KWF; KUN2015-7740). SCB acknowledges support from CIBEREHD (Instituto de Salud Carlos III), Fondo de Investigación Sanitaria/FEDER (17/00878), Ministerio de Economía y competitividad (SAF 2014–54453-R), Asociación Española Contra el Cáncer (GCB13131592CAST), CERCA Programme (Generalitat de Catalunya), AGAUR, Generalitat de Catalunya, 2017SGR21), and PERIS (Salut, Generalitat de Catalunya, SLT002/16/00398). The work was partly carried out at the Esther Koplowitz Centre, Barcelona. RMdV holds a Fellowship from the Dutch Cancer Society (KWF; KUN 2014-6666) and is supported by the Solve-RD project. The Solve-RD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 779257. Furthermore, we acknowledge the networking support of the European Cooperation in Science and Technology (COST) Action CA17118.

References

Al-Tassan, N., Chmiel, N.H., Maynard, J., Fleming, N., Livingston, A.L., Williams, G.T., Hodges, A.K., Davies, D.R., David, S.S., Sampson, J.R., Cheadle, J.P., 2002. Inherited variants of MYH associated with somatic G:C– > T:A mutations in colorectal tumors. Nat. Genet. 30 (2), 227–232.

AlDubayan, S.H., Giannakis, M., Moore, N.D., Han, G.C., Reardon, B., Hamada, T., Mu, X.J., Nishihara, R., Qian, Z., Liu, L., Yurgelun, M.B., Syngal, S., Garraway, L.A., Ogino, S., Fuchs, C.S., Van Allen, E.M., 2018. Inherited DNA-repair defects in colorectal cancer. Am. J. Hum. Genet. 102 (3), 401–414.

Alexandrov, L., Kim, J., Haradhvala, N.J., Huang, M.N., Ng, A.W.T., Boot, A., Covington, K.R., Gordenin, D.A., Bergstrom, E., Lopez-Bigas, N., Klimczak, L.J., McPherson, J.R., Morganella, S., Sabarinathan, R., Wheeler, D.A., Mustonen, V., Getz, G., Rozen, S.G., Stratton, M.R., 2018. The repertoire of mutational signatures in human cancer. bioRxiv.

Alexandrov, L.B., Jones, P.H., Wedge, D.C., Sale, J.E., Campbell, P.J., Nik-Zainal, S., Stratton, M.R., 2015a. Clock-like mutational processes in human somatic cells. Nat. Genet. 47 (12), 1402–1407.

Alexandrov, L.B., Nik-Zainal, S., Siu, H.C., Leung, S.Y., Stratton, M.R., 2015b. A mutational signature in gastric cancer suggests therapeutic strategies. Nat. Commun. 6, 8683.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.L., Boyault, S., Burkhardt, B., Butler, A.P., Caldas, C., Davies, H.R., Desmedt, C., Eils, R., Eyfjord, J.E., Foekens, J.A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinski, M., Jager, N., Jones, D.T., Jones, D., Knappskog, S., Kool, M., Lakhani, S.R., Lopez-Otin, C., Martin, S., Munshi, N.C., Nakamura, H., Northcott, P.A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J.V., Puente, X.S., Raine, K., Ramakrishna, M., Richardson, A.L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T.N., Span, P.N., Teague, J.W., Totoki, Y., Tutt, A.N., Valdes-Mas, R., van Buuren, M.M., van 't Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L.R., Zucman-Rossi, J., Futreal, P.A., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S.M., Siebert, R., Campo, E., Shibata, T., Pfister, S.M., Campbell, P.J., Stratton, M.R., 2013a. Signatures of mutational processes in human cancer. Nature 500 (7463), 415–421.

Alexandrov, Ludmil B., Nik-Zainal, S., Wedge, David C., Campbell, Peter J., Stratton, Michael R., 2013b. Deciphering signatures of mutational processes operative in human cancer. Cell Rep. 3 (1), 246–259.

Ardin, M., Cahais, V., Castells, X., Bouaoun, L., Byrnes, G., Herceg, Z., Zavadil, J., Olivier, M., 2016. MutSpec: a Galaxy toolbox for streamlined analyses of somatic mutation spectra in human and mouse cancer genomes. BMC Bioinf. 17 (1) 170–170.

Baez-Ortega, A., Gori, K., 2019. Computational approaches for discovery of mutational signatures in cancer. Brief. Bioinform. 20 (1), 77–88.

Berry, M.W., Browne, M., Langville, A.N., Pauca, V.P., Plemmons, R.J., 2007. Algorithms and applications for approximate nonnegative matrix factorization. Comput. Stat. Data Anal. 52 (1), 155–173.

Blokzijl, F., de Ligt, J., Jager, M., Sasselli, V., Roerink, S., Sasaki, N., Huch, M., Boymans, S., Kuijk, E., Prins, P., Nijman, I.J., Martincorena, I., Mokry, M., Wiegerinck, C.L., Middendorp, S., Sato, T., Schwank, G., Nieuwenhuis, E.E., Verstegen, M.M., van der Laan, L.J., de Jonge, J., JN, I.J., Vries, R.G., van de Wetering, M., Stratton, M.R., Clevers, H., Cuppen, E., van Boxtel, R., 2016. Tissue-specific mutation accumulation in human adult stem cells during life. Nature 538 (7624), 260–264.

Blokzijl, F., Janssen, R., van Boxtel, R., Cuppen, E., 2018. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. Genome Med. 10 (1), 33.

Boot, A., Huang, M.N., Ng, A.W.T., Ho, S.C., Lim, J.Q., Kawakami, Y., Chayama, K., Teh, B.T., Nakagawa, H., Rozen, S.G., 2018. In-depth characterization of the cisplatin mutational signature in human cell lines and in esophageal and liver tumors. Genome Res. 28 (5), 654–665.

Bruna, A., Rueda, O.M., Greenwood, W., Batra, A.S., Callari, M., Batra, R.N., Pogrebniak,

K., Sandoval, J., Cassidy, J.W., Tufegdzic-Vidakovic, A., Sammut, S.-J., Jones, L., Provenzano, E., Baird, R., Eirew, P., Hadfield, J., Eldridge, M., McLaren-Douglas, A., Barthorpe, A., Lightfoot, H., O'Connor, M.J., Gray, J., Cortes, J., Baselga, J., Marangoni, E., Welm, A.L., Aparicio, S., Serra, V., Garnett, M.J., Caldas, C., 2016. A biobank of breast cancer explants with preserved intra-tumor heterogeneity to screen anticancer compounds. Cell 167 (1), 260–274 e222.

Campbell, B.B., Light, N., Fabrizio, D., Zatzman, M., Fuligni, F., de Borja, R., Davidson, S., Edwards, M., Elvin, J.A., Hodel, K.P., Zahurancik, W.J., Suo, Z., Lipman, T., Wimmer, K., Kratz, C.P., Bowers, D.C., Laetsch, T.W., Dunn, G.P., Johanns, T.M., Grimmer, M.R., Smirnov, I.V., Larouche, V., Samuel, D., Bronsema, A., Osborn, M., Stearns, D., Raman, P., Cole, K.A., Storm, P.B., Yalon, M., Opocher, E., Mason, G., Thomas, G.A., Sabel, M., George, B., Ziegler, D.S., Lindhorst, S., Issai, V.M., Constantini, S., Toledano, H., Elhasid, R., Farah, R., Dvir, R., Dirks, P., Huang, A., Galati, M.A., Chung, J., Ramaswamy, V., Irwin, M.S., Aronson, M., Durno, C., Taylor, M.D., Rechavi, G., Maris, J.M., Bouffet, E., Hawkins, C., Costello, J.F., Meyn, M.S., Pursell, Z.F., Malkin, D., Tabori, U., Shlien, A., 2017. Comprehensive analysis of hypermutation in human cancer. Cell 171 (5), 1042–1056 e1010.

Campbell, J.D., Alexandrov, A., Kim, J., Wala, J., Berger, A.H., Pedamallu, C.S., Shukla, S.A., Guo, G., Brooks, A.N., Murray, B.A., Imielinski, M., Hu, X., Ling, S., Akbani, R., Rosenberg, M., Cibulskis, C., Ramachandran, A., Collisson, E.A., Kwiatkowski, D.J., Lawrence, M.S., Weinstein, J.N., Verhaak, R.G.W., Wu, C.J., Hammerman, P.S., Cherniack, A.D., Getz, G., Network, C.G.A.R., Zenklusen, J.C., Zhang, J., Felau, I., Demchok, J.A., Yang, L., Wang, Z., Ferguson, M.L., Tarnuzzer, R., Hutter, C.M., Sofia, H.J., Pihl, T., Wan, Y., Chudamani, S., Liu, J., Sun, C., Naresh, R., Lolla, L., Wu, Y., Creighton, C.J., Rathmell, W.K., Auman, J.T., Balu, S., Bodenheimer, T., Hayes, D.N., Hoadley, K.A., Hoyle, A.P., Jones, C.D., Jefferys, S.R., Meng, S., Mieczkowski, P.A., Mose, L.E., Perou, C.M., Roach, J., Shi, Y., Simons, J.V., Skelly, T., Soloway, M.G., Tan, D., Wu, J., Veluvolu, U., Parker, J.S., Wilkerson, M.D., Boice, L., Huang, M., Thorne, L.B., Getz, G., Noble, M.S., Zhang, H., Heiman, D.I., Cho, J., Gehlenborg, N., Saksena, G., Voet, D., Lin, P., Frazer, S., Kim, J., Lawrence, M.S., Chin, L., Tsao, M.-S., Allison, F., Chadwick, D., Muley, T., Meister, M., Dienemann, H., Kucherlapati, R., Park, P., Bowen, J., Gastier-Foster, J.M., Gerken, M., Leraas, K.M., Lichtenberg, T.M., Ramirez, N.C., Wise, L., Zmuda, E., Stuart, J., Collisson, E., Peifer, M., Kwiatkowski, D., Campbell, J.D., Murray, B.A., Cherniack, A.D., Berger, A.H., Sougnez, C., Schumacher, S.E., Shih, J., Beroukhim, R., Zack, T.I., Gabriel, S.B., Meyerson, M., Byers, L.A., Davidsen, T., Laird, P.W., Weisenberger, D.J., Van Den Berg, D.J., Bootwalla, M.S., Lai, P.H., Maglinte, D.T., Baylin, S.B., Herman, J.G., Danilova, L., Cope, L., Crain, D.J., Curley, E., Gardner, J., Lau, K., Mallery, D., Morris, S., Paulauskis, J., Penny, R., Shelton, C., Shelton, T., Sherman, M., Yena, P., Mills, G.B., Artyomov, M.N., Schreiber, R., Govindan, R., Meyerson, M., 2016. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat. Genet. 48 607–607.

Castellsague, E., Li, R., Aligue, R., Gonzalez, S., Sanz, J., Martin, E., Velasco, A., Capella, G., Stewart, C.J.R., Vidal, A., Majewski, J., Rivera, B., Polak, P., Matias-Guiu, X., Brunet, J., Foulkes, W.D., 2018. Novel POLE pathogenic germline variant in a family with multiple primary tumors results in distinct mutational signatures. Human mutation.

Chang, J., Tan, W., Ling, Z., Xi, R., Shao, M., Chen, M., Luo, Y., Zhao, Y., Liu, Y., Huang, X., Xia, Y., Hu, J., Parker, J.S., Marron, D., Cui, Q., Peng, L., Chu, J., Li, H., Du, Z., Han, Y., Tan, W., Liu, Z., Zhan, Q., Li, Y., Mao, W., Wu, C., Lin, D., 2017. Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. Nat. Commun. 8 15290–15290.

Church, D.N., Briggs, S.E., Palles, C., Domingo, E., Kearsey, S.J., Grimes, J.M., Gorman, M., Martin, L., Howarth, K.M., Hodgson, S.V., Kaur, K., Taylor, J., Tomlinson, I.P., 2013. DNA polymerase epsilon and delta exonuclease domain mutations in endometrial cancer. Hum. Mol. Genet. 22 (14), 2820–2828.

Connor, A.A., Denroche, R.E., Jang, G.H., Timms, L., Kalimuthu, S.N., Selander, I., McPherson, T., Wilson, G.W., Chan-Seng-Yue, M.A., Borozan, I., Ferretti, V., Grant, R.C., Lungu, I.M., Costello, E., Greenhalf, W., Palmer, D., Ghaneh, P., Neoptolemos, J.P., Buchler, M., Petersen, G., Thayer, S., Hollingsworth, M.A., Sherker, A., Durocher, D., Dhani, N., Hedley, D., Serra, S., Pollett, A., Roehrl, M.H.A., Bavi, P., Bartlett, J.M.S., Cleary, S., Wilson, J.M., Alexandrov, L.B., Moore, M., Wouters, B.G., McPherson, J.D., Notta, F., Stein, L.D., Gallinger, S., 2017. Association of distinct mutational signatures with correlates of increased immune activity in pancreatic ductal adenocarcinoma. JAMA Oncology 3 (6) 774–774.

Davies, H., Glodzik, D., Morganella, S., Yates, L.R., Staaf, J., Zou, X., Ramakrishna, M., Martin, S., Boyault, S., Sieuwerts, A.M., Simpson, P.T., King, T.A., Raine, K., Eyfjord, J.E., Kong, G., Borg, Å., Birney, E., Stunnenberg, H.G., van de Vijver, M.J., Børresen-Dale, A.-L., Martens, J.W.M., Span, P.N., Lakhani, S.R., Vincent-Salomon, A., Sotiriou, C., Tutt, A., Thompson, A.M., Van Laere, S., Richardson, A.L., Viari, A., Campbell, P.J., Stratton, M.R., Nik-Zainal, S., 2017. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nat. Med. 23 (4), 517–525.

Devarajan, K., 2008. Nonnegative matrix factorization: an analytical and interpretive tool in computational biology. PLoS Comput. Biol. 4 (7) e1000029–e1000029.

Diaz-Gay, M., Vila-Casadesus, M., Franch-Exposito, S., Hernandez-Illan, E., Lozano, J.J., Castellvi-Bel, S., 2018. Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples. BMC Bioinf. 19 (1), 224.

Dow, M., Pyke, R.M., Tsui, B.Y., Alexandrov, L.B., Nakagawa, H., Taniguchi, K., Seki, E., Harismendy, O., Shalapour, S., Karin, M., Carter, H., Font-Burgada, J., 2018. Integrative genomic analysis of mouse and human hepatocellular carcinoma. Proc. Natl. Acad. Sci. Unit. States Am. 115 (42), E9879 LP-E9888.

Drost, J., van Boxtel, R., Blokzijl, F., Mizutani, T., Sasaki, N., Sasselli, V., de Ligt, J., Behjati, S., Grolleman, J.E., van Wezel, T., Nik-Zainal, S., Kuiper, R.P., Cuppen, E., Clevers, H., 2017. Use of CRISPR-modified human stem cell organoids to study the origin of mutational signatures in cancer. Science 358 (6360), 234–238.

Fantini, D., Glaser, A.P., Rimar, K.J., Wang, Y., Schipma, M., Varghese, N., Rademaker, A., Behdad, A., Yellapa, A., Yu, Y., Sze, C.C.-L., Wang, L., Zhao, Z., Crawford, S.E., Hu, D., Licht, J.D., Collings, C.K., Bartom, E., Theodorescu, D., Shilatifard, A., Meeks, J.J., 2018. A Carcinogen-induced mouse model recapitulates the molecular alterations of human muscle invasive bladder cancer. Oncogene 37 (14), 1911–1925.

Fischer, A., Illingworth, C.J.R., Campbell, P.J., Mustonen, V., 2013. EMu: probabilistic inference of mutational processes and their localization in the cancer genome. Genome Biol. 14 (4) R39–R39.

Fong, P.C., Boss, D.S., Yap, T.A., Tutt, A., Wu, P., Mergui-Roelvink, M., Mortimer, P., Swaisland, H., Lau, A., O'Connor, M.J., Ashworth, A., Carmichael, J., Kaye, S.B., Schellens, J.H., de Bono, J.S., 2009. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N. Engl. J. Med. 361 (2), 123–134.

Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., Ponting, L., Stefancsik, R., Harsha, B., Kok, C.Y., Jia, M., Jubb, H., Sondka, Z., Thompson, S., De, T., Campbell, P.J., 2017. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res. 45 (D1), D777–D783.

Gehring, J.S., Fischer, B., Lawrence, M., Huber, W., 2015. SomaticSignatures: inferring mutational signatures from single-nucleotide variants. Bioinformatics 31 (22), 3673–3675.

George, J., Walter, V., Peifer, M., Alexandrov, L.B., Seidel, D., Leenders, F., Maas, L., Müller, C., Dahmen, I., Delhomme, T.M., Ardin, M., Leblay, N., Byrnes, G., Sun, R., De Reynies, A., McLeer-Florin, A., Bosco, G., Malchers, F., Menon, R., Altmüller, J., Becker, C., Nürnberg, P., Achter, V., Lang, U., Schneider, P.M., Bogus, M., Soloway, M.G., Wilkerson, M.D., Cun, Y., McKay, J.D., Moro-Sibilot, D., Brambilla, C.G., Lantuejoul, S., Lemaitre, N., Soltermann, A., Weder, W., Tischler, V., Brustugun, O.T., Lund-Iversen, M., Helland, Å., Solberg, S., Ansén, S., Wright, G., Solomon, B., Roz, L., Pastorino, U., Petersen, I., Clement, J.H., Sänger, J., Wolf, J., Vingron, M., Zander, T., Perner, S., Travis, W.D., Haas, S.A., Olivier, M., Foll, M., Büttner, R., Hayes, D.N., Brambilla, E., Fernandez-Cuesta, L., Thomas, R.K., 2018. Integrative genomic profiling of large-cell neuroendocrine carcinomas reveals distinct subtypes of high-grade neuroendocrine lung tumors. Nat. Commun. 9 (1) 1048–1048.

Glaser, A.P., Fantini, D., Wang, Y., Yu, Y., Rimar, K.J., Podojil, J.R., Miller, S.D., Meeks, J.J., 2018. APOBEC-mediated mutagenesis in urothelial carcinoma is associated with improved survival, mutations in DNA damage response genes, and immune response. Oncotarget 9 (4), 4537–4548.

Goh, G., Schmid, R., Guiver, K., Arpornwirat, W., Chitapanarux, I., Ganju, V., Im, S.-A., Kim, S.-B., Dechaphunkul, A., Maneechavakajorn, J., Spector, N., Yau, T., Afrit, M., Ahmed, S.B., Johnston, S.R., Gibson, N., Uttenreuther-Fischer, M., Herrero, J., Swanton, C., 2016. Clonal evolutionary analysis during HER2 blockade in HER2-positive inflammatory breast cancer: a phase II open-label clinical trial of afatinib +/- vinorelbine. PLoS Med. 13 (12) e1002136–e1002136.

Goncearenco, A., Rager, S.L., Li, M., Sang, Q.-X., Rogozin, I.B., Panchenko, A.R., 2017. Exploring background mutational processes to decipher cancer genetic heterogeneity. Nucleic Acids Res. 45 (W1), W514–W522.

Grolleman, J.E., de Voer, R.M., Elsayed, F.A., Nielsen, M., Weren, R.D.A., Palles, C., Ligtenberg, M.J.L., Vos, J.R., ten Broeke, S.W., de Miranda, N.F.C.C., Kuiper, R.A., Kamping, E.J., Jansen, E.A.M., Vink-Börger, M.E., Popp, I., Lang, A., Spier, I., Hüneburg, R., James, P.A., Li, N., Staninova, M., Lindsay, H., Cockburn, D., Spasic-Boskovic, O., Clendenning, M., Sweet, K., Capellá, G., Sjursen, W., Høberg-Vetti, H., Jongmans, M.C., Neveling, K., Geurts van Kessel, A., Morreau, H., Hes, F.J., Sijmons, R.H., Schackert, H.K., Ruiz-Ponte, C., Dymerska, D., Lubinski, J., Rivera, B., Foulkes, W.D., Tomlinson, I.P., Valle, L., Buchanan, D.D., Kenwrick, S., Adlard, J., Dimovski, A.J., Campbell, I.G., Aretz, S., Schindler, D., van Wezel, T., Hoogerbrugge, N., Kuiper, R.P., 2019. Mutational signature analysis reveals NTHL1 deficiency to cause a multi-tumor phenotype. Cancer Cell 35 (2) 256-266 e255.

Hao, J.-J., Lin, D.-C., Dinh, H.Q., Mayakonda, A., Jiang, Y.-Y., Chang, C., Jiang, Y., Lu, C.-C., Shi, Z.-Z., Xu, X., Zhang, Y., Cai, Y., Wang, J.-W., Zhan, Q.-M., Wei, W.-Q., Berman, B.P., Wang, M.-R., Koeffler, H.P., 2016. Spatial intratumoral heterogeneity and temporal clonal evolution in esophageal squamous cell carcinoma. Nat. Genet. 48 1500–1500.

Haradhvala, N.J., Kim, J., Maruvka, Y.E., Polak, P., Rosebrock, D., Livitz, D., Hess, J.M., Leshchiner, I., Kamburov, A., Mouw, K.W., Lawrence, M.S., Getz, G., 2018. Distinct mutational signatures characterize concurrent loss of polymerase proofreading and mismatch repair. Nat. Commun. 9 (1), 1746.

Hatakeyama, K., Ohshima, K., Nagashima, T., Ohnami, S., Ohnami, S., Serizawa, M., Shimoda, Y., Maruyama, K., Akiyama, Y., Urakami, K., Kusuhara, M., Mochizuki, T., Yamaguchi, K., 2018. Molecular profiling and sequential somatic mutation shift in hypermutator tumours harbouring POLE mutations. Sci. Rep. 8 (1), 8700.

Hayward, N.K., Wilmott, J.S., Waddell, N., Johansson, P.A., Field, M.A., Nones, K., Patch, A.-M., Kakavand, H., Alexandrov, L.B., Burke, H., Jakrot, V., Kazakoff, S., Holmes, O., Leonard, C., Sabarinathan, R., Mularoni, L., Wood, S., Xu, Q., Waddell, N., Tembe, V., Pupo, G.M., De Paoli-Iseppi, R., Vilain, R.E., Shang, P., Lau, L.M.S., Dagg, R.A., Schramm, S.-J., Pritchard, A., Dutton-Regester, K., Newell, F., Fitzgerald, A., Shang, C.A., Grimmond, S.M., Pickett, H.A., Yang, J.Y., Stretch, J.R., Behren, A., Kefford, R.F., Hersey, P., Long, G.V., Cebon, J., Shackleton, M., Spillane, A.J., Saw, R.P.M., López-Bigas, N., Pearson, J.V., Thompson, J.F., Scolyer, R.A., Mann, G.J., 2017. Whole-genome landscapes of major melanoma subtypes. Nature 545 175–175.

Huang, P.J., Chiu, L.Y., Lee, C.C., Yeh, Y.M., Huang, K.Y., Chiu, C.H., Tang, P., 2018a. mSignatureDB: a database for deciphering mutational signatures in human cancers. Nucleic Acids Res. 46 (D1), D964–d970.

Huang, X., Wojtowicz, D., Przytycka, T.M., 2018b. Detecting presence of mutational signatures in cancer with confidence. Bioinformatics 34 (2), 330–337.

Inman, G.J., Wang, J., Nagano, A., Alexandrov, L.B., Purdie, K.J., Taylor, R.G., Sherwood, V., Thomson, J., Hogan, S., Spender, L.C., South, A.P., Stratton, M., Chelala, C., Harwood, C.A., Proby, C.M., Leigh, I.M., 2018. The genomic landscape of cutaneous SCC reveals drivers and a novel azathioprine associated mutational signature. Nat.

Commun. 9 (1), 3667.
Jia, P., Pao, W., Zhao, Z., 2014. Patterns and processes of somatic mutations in nine major cancers. BMC Med. Genomics 7, 11.
Kanu, N., Cerone, M.A., Goh, G., Zalmas, L.-P., Bartkova, J., Dietzen, M., McGranahan, N., Rogers, R., Law, E.K., Gromova, I., Kschischo, M., Walton, M.I., Rossanese, O.W., Bartek, J., Harris, R.S., Venkatesan, S., Swanton, C., 2016. DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. Genome Biol. 17 (1) 185–185.
Kasar, S., Kim, J., Improgo, R., Tiao, G., Polak, P., Haradhvala, N., Lawrence, M.S., Kiezun, A., Fernandes, S.M., Bahl, S., Sougnez, C., Gabriel, S., Lander, E.S., Kim, H.T., Getz, G., Brown, J.R., 2015. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. Nat. Commun. 6 8866–8866.
Kim, J., Mouw, K.W., Polak, P., Braunstein, L.Z., Kamburov, A., Tiao, G., Kwiatkowski, D.J., Rosenberg, J.E., Van Allen, E.M., D'Andrea, A.D., Getz, G., 2016. Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. Nat. Genet. 48 (6), 600–606.
Krüger, S., Piro, R.M., 2019. decompTumor2Sig: identification of mutational signatures active in individual tumors. BMC Bioinf. 20, 152.
Kucab, J.E., Zou, X., Morganella, S., Joel, M., Nanda, A.S., Nagy, E., Gomez, C., Degasperi, A., Harris, R., Jackson, S.P., Arlt, V.M., Phillips, D.H., Nik-Zainal, S., 2019. A compendium of mutational signatures of environmental agents. Cell 177 (4), 821–836 e16.
Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., et al., 1993. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 75 (6), 1215–1225.
Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. Nature 401 788–788.
Lee, J., Lee, Andy J., Lee, J.-K., Park, J., Kwon, Y., Park, S., Chun, H., Ju, Y.S., Hong, D., 2018. Mutalisk: a web-based somatic MUTation AnaLyIS toolKit for genomic, transcriptional and epigenomic signatures. Nucleic Acids Res. 46 (W1), W102–W108.
Letouzé, E., Shinde, J., Renault, V., Couchy, G., Blanc, J.-F., Tubacher, E., Bayard, Q., Bacq, D., Meyer, V., Semhoun, J., Bioulac-Sage, P., Prévôt, S., Azoulay, D., Paradis, V., Imbeaud, S., Deleuze, J.-F., Zucman-Rossi, J., 2017. Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis. Nat. Commun. 8, 1315.
Li, X.C., Wang, M.Y., Yang, M., Dai, H.J., Zhang, B.F., Wang, W., Chu, X.L., Wang, X., Zheng, H., Niu, R.F., Zhang, W., Chen, K.X., 2018. A mutational signature associated with alcohol consumption and prognostically significantly mutated driver genes in esophageal squamous cell carcinoma. Annals Oncol : Off. J. Europ. Soc. Med. Oncol. 29 (4), 938–944.
Ligtenberg, M.J., Kuiper, R.P., Chan, T.L., Goossens, M., Hebeda, K.M., Voorendt, M., Lee, T.Y., Bodmer, D., Hoenselaar, E., Hendriks-Cornelissen, S.J., Tsui, W.Y., Kong, C.K., Brunner, H.G., van Kessel, A.G., Yuen, S.T., van Krieken, J.H., Leung, S.Y., Hoogerbrugge, N., 2009. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nat. Genet. 41 (1), 112–117.
Liu, D., Abbosh, P., Keliher, D., Reardon, B., Miao, D., Mouw, K., Weiner-Taylor, A., Wankowicz, S., Han, G., Teo, M.Y., Cipolla, C., Kim, J., Iyer, G., Al-Ahmadie, H., Dulaimi, E., Chen, D.Y.T., Alpaugh, R.K., Hoffman-Censits, J., Garraway, L.A., Getz, G., Carter, S.L., Bellmunt, J., Plimack, E.R., Rosenberg, J.E., Van Allen, E.M., 2017. Mutational patterns in chemotherapy resistant muscle-invasive bladder cancer. Nat. Commun. 8 (1), 2193.
Llosa, N.J., Cruise, M., Tam, A., Wicks, E.C., Hechenbleikner, E.M., Taube, J.M., Blosser, R.L., Fan, H., Wang, H., Luber, B.S., Zhang, M., Papadopoulos, N., Kinzler, K.W., Vogelstein, B., Sears, C.L., Anders, R.A., Pardoll, D.M., Housseau, F., 2015. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. Cancer Discov. 5 (1), 43–51.
Macintyre, G., Goranova, T.E., De Silva, D., Ennis, D., Piskorz, A.M., Eldridge, M., Sie, D., Lewsley, L.A., Hanif, A., Wilson, C., Dowson, S., Glasspool, R.M., Lockley, M., Brockbank, E., Montes, A., Walther, A., Sundar, S., Edmondson, R., Hall, G.D., Clamp, A., Gourley, C., Hall, M., Fotopoulou, C., Gabra, H., Paul, J., Supernat, A., Millan, D., Hoyle, A., Bryson, G., Nourse, C., Mincarelli, L., Sanchez, L.N., Ylstra, B., Jimenez-Linan, M., Moore, L., Hofmann, O., Markowetz, F., McNeish, I.A., Brenton, J.D., 2018. Copy number signatures and mutational processes in ovarian carcinoma. Nat. Genet. 50 (9), 1262–1270.
Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., Koeffler, H.P., 2018. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 28 (11), 1747–1756.
Meier, B., Volkova, N.V., Hong, Y., Schofield, P., Campbell, P.J., Gerstung, M., Gartner, A., 2018. Mutational signatures of DNA mismatch repair deficiency in C. elegans and human cancers. Genome Res. 28 (5), 666–675.
Mensenkamp, A.R., Vogelaar, I.P., van Zelst-Stams, W.A., Goossens, M., Ouchene, H., Hendriks-Cornelissen, S.J., Kwint, M.P., Hoogerbrugge, N., Nagtegaal, I.D., Ligtenberg, M.J., 2014. Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in Lynch syndrome-like tumors. Gastroenterology 146 (3), 643–646 e648.
Miyaki, M., Konishi, M., Tanaka, K., Kikuchi-Yanoshita, R., Muraoka, M., Yasuno, M., Igari, T., Koike, M., Chiba, M., Mori, T., 1997. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat. Genet. 17 (3), 271–272.
Morganella, S., Alexandrov, L.B., Glodzik, D., Zou, X., Davies, H., Staaf, J., Sieuwerts, A.M., Brinkman, A.B., Martin, S., Ramakrishna, M., Butler, A., Kim, H.-Y., Borg, Å., Sotiriou, C., Futreal, P.A., Campbell, P.J., Span, P.N., Van Laere, S., Lakhani, S.R., Eyfjord, J.E., Thompson, A.M., Stunnenberg, H.G., van de Vijver, M.J., Martens, J.W.M., Børresen-Dale, A.-L., Richardson, A.L., Kong, G., Thomas, G., Sale, J., Rada, C., Stratton, M.R., Birney, E., Nik-Zainal, S., 2016. The topography of mutational processes in breast cancer genomes. Nat. Commun. 7 11383–11383.
Nagahashi, M., Wakai, T., Shimada, Y., Ichikawa, H., Kameyama, H., Kobayashi, T., Sakata, J., Yagi, R., Sato, N., Kitagawa, Y., Uetake, H., Yoshida, K., Oki, E., Kudo, S.-e., Izutsu, H., Kodama, K., Nakada, M., Tse, J., Russell, M., Heyer, J., Powers, W., Sun, R., Ring, J.E., Takabe, K., Protopopov, A., Ling, Y., Okuda, S., Lyle, S., 2016. Genomic landscape of colorectal cancer in Japan: clinical implications of comprehensive genomic sequencing for precision medicine. Genome Med. 8 (1) 136–136.
Nesic, K., Wakefield, M., Kondrashova, O., Scott, C.L., McNeish, I.A., 2018. Targeting DNA repair: the genome as a potential biomarker. J. Pathol. 244 (5), 586–597.
Ng, A.W.T., Poon, S.L., Huang, M.N., Lim, J.Q., Boot, A., Yu, W., Suzuki, Y., Thangaraju, S., Ng, C.C.Y., Tan, P., Pang, S.-T., Huang, H.-Y., Yu, M.-C., Lee, P.-H., Hsieh, S.-Y., Chang, A.Y., Teh, B.T., Rozen, S.G., 2017. Aristolochic acids and their derivatives are widely implicated in liver cancers in Taiwan and throughout Asia. Science Translational Medicine 9(412).
Nicolaides, N.C., Papadopoulos, N., Liu, B., Wei, Y.F., Carter, K.C., Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., et al., 1994. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 371 (6492), 75–80.
Nik-Zainal, S., Alexandrov, Ludmil B., Wedge, David C., van loo, P., Greenman, Christopher D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, Lucy A., Menzies, A., Martin, S., Leung, K., Chen, L., Leroy, C., Ramakrishna, M., Rance, R., Lau, King W., Mudie, Laura J., Varela, I., McBride, David J., Bignell, Graham R., Cooke, Susanna L., Shlien, A., Gamble, J., Whitmore, I., Maddison, M., Tarpey, Patrick S., Davies, Helen R., Papaemmanuil, E., Stephens, Philip J., McLaren, S., Butler, Adam P., Teague, Jon W., Jönsson, G., Garber, Judy E., Silver, D., Miron, P., Fatima, A., Boyault, S., Langerød, A., Tutt, A., Martens, John W.M., Aparicio, Samuel A.J.R., Borg, Å., Salomon, Anne V., Thomas, G., Børresen-Dale, A.-L., Richardson, Andrea L., Neuberger, Michael S., Futreal, P.A., Campbell, Peter J., Stratton, Michael R., 2012. Mutational processes molding the genomes of 21 breast cancers. Cell, vol. 149 (5), 979–993.
Nik-Zainal, S., Davies, H., Staaf, J., Ramakrishna, M., Glodzik, D., Zou, X., Martincorena, I., Alexandrov, L.B., Martin, S., Wedge, D.C., Van Loo, P., Ju, Y.S., Smid, M., Brinkman, A.B., Morganella, S., Aure, M.R., Lingjaerde, O.C., Langerod, A., Ringner, M., Ahn, S.M., Boyault, S., Brock, J.E., Broeks, A., Butler, A., Desmedt, C., Dirix, L., Dronov, S., Fatima, A., Foekens, J.A., Gerstung, M., Hooijer, G.K., Jang, S.J., Jones, D.R., Kim, H.Y., King, T.A., Krishnamurthy, S., Lee, H.J., Lee, J.Y., Li, Y., McLaren, S., Menzies, A., Mustonen, V., O'Meara, S., Pauporte, I., Pivot, X., Purdie, C.A., Raine, K., Ramakrishnan, K., Rodriguez-Gonzalez, F.G., Romieu, G., Sieuwerts, A.M., Simpson, P.T., Shepherd, R., Stebbings, L., Stefansson, O.A., Teague, J., Tommasi, S., Treilleux, I., Van den Eynden, G.G., Vermeulen, P., Vincent-Salomon, A., Yates, L., Caldas, C., van't Veer, L., Tutt, A., Knappskog, S., Tan, B.K., Jonkers, J., Borg, A., Ueno, N.T., Sotiriou, C., Viari, A., Futreal, P.A., Campbell, P.J., Span, P.N., Van Laere, S., Lakhani, S.R., Eyfjord, J.E., Thompson, A.M., Birney, E., Stunnenberg, H.G., van de Vijver, M.J., Martens, J.W., Borresen-Dale, A.L., Richardson, A.L., Kong, G., Thomas, G., Stratton, M.R., 2016. Landscape of Somatic Mutations in 560 Breast Cancer Whole-Genome Sequences. Nature, vol. 534 (7605), 47–54.
Nik-Zainal, S., Kucab, J.E., Morganella, S., Glodzik, D., Alexandrov, L.B., Arlt, V.M., Weninger, A., Hollstein, M., Stratton, M.R., Phillips, D.H., 2015. The genome as a record of environmental exposure. Mutagenesis 30 (6), 763–770.
Nik-Zainal, S., Wedge, D.C., Alexandrov, L.B., Petljak, M., Butler, A.P., Bolli, N., Davies, H.R., Knappskog, S., Martin, S., Papaemmanuil, E., Ramakrishna, M., Shlien, A., Simonic, I., Xue, Y., Tyler-Smith, C., Campbell, P.J., Stratton, M.R., 2014. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. Nat. Genet. 46 (5), 487–491.
Palles, C., Cazier, J.B., Howarth, K.M., Domingo, E., Jones, A.M., Broderick, P., Kemp, Z., Spain, S.L., Guarino, E., Salguero, I., Sherborne, A., Chubb, D., Carvajal-Carmona, L.G., Ma, Y., Kaur, K., Dobbins, S., Barclay, E., Gorman, M., Martin, L., Kovac, M.B., Humphray, S., Consortium, C., Consortium, W.G.S., Lucassen, A., Holmes, C.C., Bentley, D., Donnelly, P., Taylor, J., Petridis, C., Roylance, R., Sawyer, E.J., Kerr, D.J., Clark, S., Grimes, J., Kearsey, S.E., Thomas, H.J., McVean, G., Houlston, R.S., Tomlinson, I., 2013. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat. Genet. 45 (2), 136–144.
Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., et al., 1994. Mutation of a mutL homolog in hereditary colon cancer. Science 263 (5153), 1625–1629.
Petljak, M., Alexandrov, L.B., Brammeld, J.S., Price, S., Wedge, D.C., Grossmann, S., Dawson, K.J., Ju, Y.S., Iorio, F., Tubio, J.M.C., Koh, C.C., Georgakopoulos-Soares, I., Rodriguez-Martin, B., Otlu, B., O'Meara, S., Butler, A.P., Menzies, A., Bhosle, S.G., Raine, K., Jones, D.R., Teague, J.W., Beal, K., Latimer, C., O'Neill, L., Zamora, J., Anderson, E., Patel, N., Maddison, M., Ng, B.L., Graham, J., Garnett, M.J., McDermott, U., Nik-Zainal, S., Campbell, P.J., Stratton, M.R., 2019. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. Cell 176 (6), 1282–1294 e1220.
Pfeifer, G.P., 2006. Mutagenesis at methylated CpG sequences. Curr. Top. Microbiol. Immunol. 301, 259–281.
Pilati, C., Shinde, J., Alexandrov, L.B., Assie, G., Andre, T., Helias-Rodzewicz, Z., Ducoudray, R., Le Corre, D., Zucman-Rossi, J., Emile, J.F., Bertheral, J., Letouze, E., Laurent-Puig, P., 2017. Mutational signature analysis identifies MUTYH deficiency in colorectal cancers and adrenocortical carcinomas. J. Pathol. 242 (1), 10–15.
Pleasance, E.D., Cheetham, R.K., Stephens, P.J., McBride, D.J., Humphray, S.J., Greenman, C.D., Varela, I., Lin, M.L., Ordonez, G.R., Bignell, G.R., Ye, K., Alipaz, J., Bauer, M.J., Beare, D., Butler, A., Carter, R.J., Chen, L., Cox, A.J., Edkins, S., Kokko-Gonzales, P.I., Gormley, N.A., Grocock, R.J., Haudenschild, C.D., Hims, M.M., James, T., Jia, M., Kingsbury, Z., Leroy, C., Marshall, J., Menzies, A., Mudie, L.J., Ning, Z.,

Royce, T., Schulz-Trieglaff, O.B., Spiridou, A., Stebbings, L.A., Szajkowski, L., Teague, J., Williamson, D., Chin, L., Ross, M.T., Campbell, P.J., Bentley, D.R., Futreal, P.A., Stratton, M.R., 2010a. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 463 (7278), 191–196.

Pleasance, E.D., Stephens, P.J., O'Meara, S., McBride, D.J., Meynert, A., Jones, D., Lin, M.L., Beare, D., Lau, K.W., Greenman, C., Varela, I., Nik-Zainal, S., Davies, H.R., Ordonez, G.R., Mudie, L.J., Latimer, C., Edkins, S., Stebbings, L., Chen, L., Jia, M., Leroy, C., Marshall, J., Menzies, A., Butler, A., Teague, J.W., Mangion, J., Sun, Y.A., McLaughlin, S.F., Peckham, H.E., Tsung, E.F., Costa, G.L., Lee, C.C., Minna, J.D., Gazdar, A., Birney, E., Rhodes, M.D., McKernan, K.J., Stratton, M.R., Futreal, P.A., Campbell, P.J., 2010b. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature 463 (7278), 184–190.

Polak, P., Kim, J., Braunstein, L.Z., Karlic, R., Haradhavala, N.J., Tiao, G., Rosebrock, D., Livitz, D., Kubler, K., Mouw, K.W., Kamburov, A., Maruvka, Y.E., Leshchiner, I., Lander, E.S., Golub, T.R., Zick, A., Orthwein, A., Lawrence, M.S., Batra, R.N., Caldas, C., Haber, D.A., Laird, P.W., Shen, H., Ellisen, L.W., D'Andrea, A.D., Chanock, S.J., Foulkes, W.D., Getz, G., 2017. A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. Nat. Genet. 49 (10), 1476–1486.

Poon, S.L., Huang, M.N., Choo, Y., McPherson, J.R., Yu, W., Heng, H.L., Gan, A., Myint, S.S., Siew, E.Y., Ler, L.D., Ng, L.G., Weng, W.-H., Chuang, C.-K., Yuen, J.S.P., Pang, S.-T., Tan, P., Teh, B.T., Rozen, S.G., 2015. Mutation signatures implicate aristolochic acid in bladder cancer development. Genome Med. 7 (1), 38.

Rahbari, R., Wuster, A., Lindsay, S.J., Hardwick, R.J., Alexandrov, L.B., Turki, S.A., Dominiczak, A., Morris, A., Porteous, D., Smith, B., Stratton, M.R., Consortium, U.K., Hurles, M.E., 2016. Timing, rates and spectra of human germline mutation. Nat. Genet. 48 (2), 126–133.

Robles-Espinoza, C.D., Roberts, N.D., Chen, S., Leacy, F.P., Alexandrov, L.B., Pornputtapong, N., Halaban, R., Krauthammer, M., Cui, R., Timothy Bishop, D., Adams, D.J., 2016. Germline MC1R status influences somatic mutation burden in melanoma. Nat. Commun. 7 12064–12064.

Rodrigues, M., Mobuchon, L., Houy, A., Fievet, A., Gardrat, S., Barnhill, R.L., Popova, T., Servois, V., Rampanou, A., Mouton, A., Dayot, S., Raynal, V., Galut, M., Putterman, M., Tick, S., Cassoux, N., Roman-Roman, S., Bidard, F.C., Lantz, O., Mariani, P., Piperno-Neumann, S., Stern, M.H., 2018. Outlier response to anti-PD1 in uveal melanoma reveals germline MBD4 mutations in hypermutated tumors. Nat. Commun. 9 (1), 1866.

Roerink, S.F., Sasaki, N., Lee-Six, H., Young, M.D., Alexandrov, L.B., Behjati, S., Mitchell, T.J., Grossmann, S., Lightfoot, H., Egan, D.A., Pronk, A., Smakman, N., van Gorp, J., Anderson, E., Gamble, S.J., Alder, C., van de Wetering, M., Campbell, P.J., Stratton, M.R., Clevers, H., 2018. Intra-tumour diversification in colorectal cancer at the single-cell level. Nature 556 (7702), 457–462.

Rosales, R.A., Drummond, R.D., Valieris, R., Dias-Neto, E., da Silva, I.T., 2017. signeR: an empirical Bayesian approach to mutational signature discovery. Bioinformatics 33 (1), 8–16.

Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S., Swanton, C., 2016. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol. 17, 31.

Sanders, M.A., Chew, E., Flensburg, C., Zeilemaker, A., Miller, S.E., Al Hinai, A.S., Bajel, A., Luiken, B., Rijken, M., McLennan, T., Hoogenboezem, R.M., Kavelaars, F.G., Frohling, S., Blewitt, M.E., Bindels, E.M., Alexander, W.S., Lowenberg, B., Roberts, A.W., Valk, P.J.M., Majewski, I.J., 2018. MBD4 guards against methylation damage and germline deficiency predisposes to clonal hematopoiesis and early-onset AML. Blood 132 (14), 1526–1534.

Schulze, K., Imbeaud, S., Letouze, E., Alexandrov, L.B., Calderaro, J., Rebouissou, S., Couchy, G., Meiller, C., Shinde, J., Soysouvanh, F., Calatayud, A.L., Pinyol, R., Pelletier, L., Balabaud, C., Laurent, A., Blanc, J.F., Mazzaferro, V., Calvo, F., Villanueva, A., Nault, J.C., Bioulac-Sage, P., Stratton, M.R., Llovet, J.M., Zucman-Rossi, J., 2015. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat. Genet. 47 (5), 505–511.

Secrier, M., Li, X., de Silva, N., Eldridge, M.D., Contino, G., Bornschein, J., MacRae, S., Grehan, N., O'Donovan, M., Miremadi, A., Yang, T.-P., Bower, L., Chettouh, H., Crawte, J., Galeano-Dalmau, N., Grabowska, A., Saunders, J., Underwood, T., Waddell, N., Barbour, A.P., Nutzinger, B., Achilleos, A., Edwards, P.A.W., Lynch, A.G., Tavaré, S., Fitzgerald, R.C., Consortium, t.O.C.C., Molecular, S., Noorani, A., Elliott, R.F., Weaver, J., Ross-Innes, C., Smith, L., Abdullahi, Z., de la Rue, R., Cluroe, A., Malhotra, S., Hardwick, R., Ford, H., Smith, M.L., Davies, J., Turkington, R., Hayes, S.J., Ang, Y., Preston, S.R., Oakes, S., Bagwan, I., Save, V., Skipworth, R.J.E., Hupp, T.R., O'Neill, J.R., Tucker, O., Taniere, P., Noble, F., Owsley, J., Lovat, L., Haidry, R., Eneh, V., Crichton, C., Barr, H., Shepherd, N., Old, O., Lagergren, J., Gossage, J., Davies, A., Chang, F., Zylstra, J., Sanders, G., Berrisford, R., Harden, C., Bunting, D., Lewis, M., Cheong, E., Kumar, B., Parsons, S.L., Soomro, I., Kaye, P., Collier, P., Igali, L., Welch, I., Scott, M., Sothi, S., Suortamo, S., Lishman, S., Beardsmore, D., Francies, H.E., Garnett, M.J., Pearson, J.V., Nones, K., Patch, A.-M., Grimmond, S.M., 2016. Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. Nat. Genet., vol. 48 1131–1131.

Shinbrot, E., Henninger, E.E., Weinhold, N., Covington, K.R., Goksenin, A.Y., Schultz, N., Chao, H., Doddapaneni, H., Muzny, D.M., Gibbs, R.A., Sander, C., Pursell, Z.F., Wheeler, D.A., 2014. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. Genome Res. 24 (11), 1740–1750.

Shinde, J., Bayard, Q., Imbeaud, S., Hirsch, T.Z., Liu, F., Renault, V., Zucman-Rossi, J., Letouzé, E., 2018. Palimpsest: an R package for studying mutational and structural variant signatures along clonal evolution in cancer. Bioinformatics 34 (19), 3380–3381.

Shiraishi, Y., Tremmel, G., Miyano, S., Stephens, M., 2015. A simple model-based approach to inferring and visualizing cancer mutation signatures. PLoS Genet. 11 (12), e1005657.

Starrett, G.J., Luengas, E.M., McCann, J.L., Ebrahimi, D., Temiz, N.A., Love, R.P., Feng, Y., Adolph, M.B., Chelico, L., Law, E.K., Carpenter, M.A., Harris, R.S., 2016. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. Nat. Commun. 7, 12918.

Stratton, M.R., 2011. Exploring the genomes of cancer cells: progress and promise. Science 331 (6024), 1553–1558.

Stratton, M.R., Campbell, P.J., Futreal, P.A., 2009. The cancer genome. Nature 458 719–719.

Supek, F., Lehner, B., 2017. Clustered mutation signatures reveal that error-prone DNA repair targets mutations to active genes. Cell 170 (3), 534–547 e523.

Supek, F., Lehner, B., Hajkova, P., Warnecke, T., 2014. Hydroxymethylated cytosines are associated with elevated C to G transversion rates. PLoS Genet. 10 (9), e1004585.

Taylor, B.J., Nik-Zainal, S., Wu, Y.L., Stebbings, L.A., Raine, K., Campbell, P.J., Rada, C., Stratton, M.R., Neuberger, M.S., 2013. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. Elife 2 e00534.

Temko, D., Van Gool, I.C., Rayner, E., Glaire, M., Makino, S., Brown, M., Chegwidden, L., Palles, C., Depreeuw, J., Beggs, A., Stathopoulou, C., Mason, J., Baker, A.M., Williams, M., Cerundolo, V., Rei, M., Taylor, J.C., Schuh, A., Ahmed, A., Amant, F., Lambrechts, D., Smit, V.T., Bosse, T., Graham, T.A., Church, D.N., Tomlinson, I., 2018. Somatic POLE exonuclease domain mutations are early events in sporadic endometrial and colorectal carcinogenesis, determining driver mutational landscape, clonal neoantigen burden and immune response. J. Pathol. 245 (3), 283–296.

The Cancer Genome Atlas, N, 2012. Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330.

Viel, A., Bruselles, A., Meccia, E., Fornasarig, M., Quaia, M., Canzonieri, V., Policicchio, E., Urso, E.D., Agostini, M., Genuardi, M., Lucci-Cordisco, E., Venesio, T., Martayan, A., Diodoro, M.G., Sanchez-Mete, L., Stigliano, V., Mazzei, F., Grasso, F., Giuliani, A., Baiocchi, M., Maestro, R., Giannini, G., Tartaglia, M., Alexandrov, L.B., Bignami, M., 2017. A specific mutational signature associated with DNA 8-Oxoguanine persistence in MUTYH-defective colorectal cancer. EBioMedicine.

Waddell, N., Pajic, M., Patch, A.-M., Chang, D.K., Kassahn, K.S., Bailey, P., Johns, A.L., Miller, D., Nones, K., Quek, K., Quinn, M.C.J., Robertson, A.J., Fadlullah, M.Z.H., Bruxner, T.J.C., Christ, A.N., Harliwong, I., Idrisoglu, S., Manning, S., Nourse, C., Nourbakhsh, E., Wani, S., Wilson, P.J., Markham, E., Cloonan, N., Anderson, M.J., Fink, J.L., Holmes, O., Kazakoff, S.H., Leonard, C., Newell, F., Poudel, B., Song, S., Taylor, D., Waddell, N., Wood, S., Xu, Q., Wu, J., Pinese, M., Cowley, M.J., Lee, H.C., Jones, M.D., Nagrial, A.M., Humphris, J., Chantrill, L.A., Chin, V., Steinmann, A.M., Mawson, A., Humphrey, E.S., Colvin, E.K., Chou, A., Scarlett, C.J., Pinho, A.V., Giry-Laterriere, M., Rooman, I., Samra, J.S., Kench, J.G., Pettitt, J.A., Merrett, N.D., Toon, C., Epari, K., Nguyen, N.Q., Barbour, A., Zeps, N., Jamieson, N.B., Graham, J.S., Niclou, S.P., Bjerkvig, R., Grützmann, R., Aust, D., Hruban, R.H., Maitra, A., Iacobuzio-Donahue, C.A., Wolfgang, C.L., Morgan, R.A., Lawlor, R.T., Corbo, V., Bassi, C., Falconi, M., Zamboni, G., Tortora, G., Tempero, M.A., Australian pancreatic cancer genome, I., Gill, A.J., Eshleman, J.R., Pilarsky, C., Scarpa, A., Musgrove, E.A., Pearson, J.V., Biankin, A.V., Grimmond, S.M., 2015. Whole genomes redefine the mutational landscape of pancreatic cancer. Nature, vol. 518, 495.

Weren, R.D., Ligtenberg, M.J., Kets, C.M., de Voer, R.M., Verwiel, E.T., Spruijt, L., van Zelst-Stams, W.A., Jongmans, M.C., Gilissen, C., Hehir-Kwa, J.Y., Hoischen, A., Shendure, J., Boyle, E.A., Kamping, E.J., Nagtegaal, I.D., Tops, B.B., Nagengast, F.M., Geurts van Kessel, A., van Krieken, J.H., Kuiper, R.P., Hoogerbrugge, N., 2015. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. Nat. Genet. 47 (6), 668–671.

Zou, X., Owusu, M., Harris, R., Jackson, S.P., Loizou, J.I., Nik-Zainal, S., 2018. Validating the concept of mutational signatures with isogenic cell models. Nat. Commun. 9 (1), 1744.