Clonal evolution in healthy and premalignant tissues: implications for early cancer interception strategies

Running title: Clonal evolution in histologically normal tissues

Jayant K. Rane[1,2,3], Alexander M. Frankell[3,4], Clare E. Weeden[3], Charles Swanton[3,4,5,*]

[1]University College London Cancer Institute, London, UK
[2]Department of Clinical Oncology, University College London Hospitals, London, UK
[3]Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK
[4]Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK
[5]Department of Medical Oncology, University College London Hospitals, London, UK

[*] Corresponding author:
Prof. Charles Swanton
Cancer Evolution and Genome Instability Laboratory
The Francis Crick Institute
London, NW1 1AT
Charles.Swanton@crick.ac.uk
Tel: 00 44 20 3796 2047

Conflict of interest:

neoantigens (PCT/EP2016/059401), identifying patient responses to immune checkpoint blockade (PCT/EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival rates of patients with cancer (PCT/GB2020/050221), identifying patients who respond to cancer treatment (PCT/GB2018/051912), a US patent relating to detecting tumour mutations (PCT/US2017/28013), methods for lung cancer detection (US20190106751A1) and both a European and US patent related to identifying indel mutation targets (PCT/GB2018/051892) and is a co-inventor on a patent application to determine methods and systems for tumour monitoring (PCT/EP2022/077987). C.S. is a named inventor on a provisional patent related to a ctDNA detection algorithm. A.M.F. is a named inventor on a patent application to determine methods and systems for tumour monitoring (PCT/EP2022/077987). J.K.R. and C.E.W declares no competing interests.

Abstract

Histologically normal human tissues accumulate significant mutational burden with age. The extent and spectra of mutagenesis are comparable both in rapidly proliferating and post-mitotic tissues and in stem cells compared to their differentiated progeny. Some of these mutations provide increased fitness, giving rise to clones which, at times, can replace the entire surface area of tissues. Compared to cancer, somatic mutations in histologically normal tissues are primarily single nucleotide variations. Interestingly though, the presence of these mutations and positive clonal selection in isolation remains a poor indicator of potential future cancer transformation in solid tissues. Common clonally expanded mutations in histologically normal tissues also do not always represent the most frequent early mutations in cancers of corresponding tissues, indicating differences in selection pressures. Preliminary evidence implies that stroma and immune system co-evolve with age, which may impact selection dynamics. In this review, we will explore the mutational landscape of histologically normal and pre-malignant human somatic tissues in detail and discuss cell-intrinsic and environmental factors that can determine the fate of positively selected mutations within them. Precisely pinpointing these determinants of cancer transformation would aid development of early cancer interventional and prevention strategies.

Introduction

In 1914, Theodor Boveri hypothesised that irreparable defect in chromosomes leads to the malignant transformation (1). Experimental evidence in 1970s-80s suggested that the acquisition of stepwise mutational events in a normal cell can cause transformation into an indefinitely dividing cell (2-4). Subsequent genomic studies found that all tumour cells share a set of genomic mutations, implying a single, most recent common ancestor (MRCA) (5). The genome of MRCA progeny continues to evolve as it expands in response to rapidly changing cell-intrinsic and cell-extrinsic factors and as a result of neutral genetic drift, resulting in considerable intra-tumour heterogeneity (ITH) (6). ITH is associated with tumour aggressiveness, response to treatment and ultimately, patient outcome (7-14). Such somatic mutational diversity and clonal heterogeneity is now considered as one of the fundamental hallmarks of cancer (15). However in addition, clonally expanded cancer-associated gene mutations have now been identified in multiple histologically normal human tissues (16). Investigating the fate of mutant clones and their interactions with surrounding tissue could inform key issues in cancer medicine including; whether positively selected mutations in histologically normal tissues preferentially transform into cancer, what the barriers are for malignant transformation of 'initiated' cells, and whether environmental exposures and germline variations co-operate with somatic mutagenesis during cancer initiation. We will examine pertinent literature regarding the mutational landscape of clones within histologically normal and pre-malignant human tissues and discuss how these insights may in the future be applied to clinical decision making around early cancer detection and prevention.

Somatic mutations and clonality in histologically normal tissue

Genomic sequencing-based analyses of clonal evolution in cancer and pre-invasive disease have suggested the possibility of clonal expansions even prior to malignant transformation (17,18). The possibility of positive mutational selection leading to clonal expansion in histologically normal tissue was supported by observations of regions of clonal expansions carrying skewed X-inactivation, specific mitochondrial enzyme inactivation, and shared *TP53* inactivation patterns (19). The common challenges associated with detailed clonal analysis within histologically normal tissues such as extremely small clonal size, low

DNA input during library preparation, and presence of low frequency alterations have been overcome in the last decade using inventive sample collection and sequencing strategies (20-37). These strategies include the use of micro biopsies, sequencing of single clonal units (e.g., crypts), *in-vitro* expansion of single cells (e.g., organoids), single cell DNA sequencing, and NanoSeq (20-22,26,33). DNA-sequencing of histologically normal tissue using these strategies is providing unequivocal evidence for the presence of widespread somatic mutations and the positive selection of some of these mutations leading to clonal expansion employing methods such as dNdScv (38). This method employs trinucleotide context-dependent substitution matrices to identify positively selected mutations within genes over and above the background mutational changes in the genome.

Origins of somatic mutations in histologically normal tissue

Age-driven somatic mutations were detected in all tested human tissues, including in germ cells (24). The source of these mutations could be divided into three categories (Figure 1). The first category includes cell-intrinsic mutagenic insults. These could be cell cycle-dependent, such as those resulting from replication errors (39) or cell cycle-independent, such as those caused by spontaneous deamination of methyl cytosine and oxidative damage from endogenous sources, such as the mitochondria (33). Secondly, mutations can arise from endogenous environmental factors such as the microbiome. The microbiome is an integral part of several tissues and can contribute to somatic mutagenesis (40-42). Specific example includes colibactin produced by *Escherichia Coli* from the gut microbiome, which induces DNA-alkylation driven mutations in colon (43). Even hormones such as oestrogen can directly stimulate activation-induced deaminase function to generate somatic mutations (44). The last, and the most preventable source of somatic mutations is external environmental processes. These include smoking, ultra-violet (UV) light exposure, dietary components, and iatrogenic causes such as radiotherapy and chemotherapy (45,46). Irrespective of the source of mutagenic insult, they primarily cause single nucleotide variations (SNVs) and small insertions and deletions (INDELs) in genomes. (46-48). They can rarely also result in more complex alterations including structural chromosomal changes. Normal somatic cells accumulate roughly 9-56 SNVs/cell/year depending on the tissue-type (49). Many mutagenic processes have different probabilities to alter specific nucleotides in certain nucleotide sequence contexts, which allows deconvolution into distinct 'mutational

5

signatures', each of which may represent a specific mutagenic exposure (45). The most prevalent mutational signatures found in human histologically normal somatic tissues were driven by spontaneous or enzymatic deamination of 5-methylcytosine (SBS1) and aging/oxidative damage (SBS5/40) (19). Temporal sequence can also be inferred for these mutational processes either where multiple clonally related samples are available or by timing relative to copy number gains, where early mutations increase in number according to the extent of chromosomal copy number gain upon which they are encoded, allowing identification of stepwise clonal evolutionary patterns (5,8,20-23,26,27,29,50-53). Evidence is emerging that some mutations are clonally expanded in histologically normal tissues, allowing either widespread clonal sweeps or prolonged restricted mutated clonal dominance within a single clonal unit based on local anatomy.

Somatic mutations resulting in large clonal sweeps in histologically normal tissues

In tissues with no clear spatially restricted anatomical units, such as squamous epithelia or haematopoietic system, it is possible for a single clone with survival advantage over its neighbours to expand across a relatively large surface. This is indeed noted in the haematopoietic system, where just 12-18 clones contribute to about 30-60% of haematopoietic output in people over the age of 75 compared to more than 20,000 active clones in younger people (37). Similarly, the histologically normal urothelium in people exposed to aristolochic acid (AA) had clones measuring several square centimetres each (28). Moderately large clones measuring more than 5 square millilitres have also been identified in histologically normal oesophageal and epidermal epithelia (22,23). It is unclear what limits total clonal sweep by the dominant clone over the entire organ surface. Potentially, the timing and nature of the founder mutation, founder mutation cell-of-origin, the microenvironmental cellular and metabolic feedback loops, and the clonal competition between diverse clones could determine eventual clonal expansion (54,55). Exposure to known carcinogens also influences clonal size and diversity (23,26,28). The analysis of sun-exposed skin, urothelium in people with AA exposure history, and bronchial epithelium of smokers demonstrated that the exposure to carcinogen resulted in a higher mutational burden and a higher frequency of more than one driver mutation/clone (23,26,28). Interestingly, colonic and haematopoietic stem cells were not protected against these

mutations compared to their differentiated progeny (20). The commonest of these mutations were SNVs with mutational signatures predominantly derived from spontaneous or enzymatic deamination of 5-methylcytosine (SBS1), aging (SBS5/40), and *APOBEC* activity (SBS2/13) (19). Rarely, loss of heterozygosity (LOH), copy number alterations (CNAs), and structural abnormalities were also detected, but widespread chromosomal instability (CIN) was not detected (20-37). Positively selected genes often had distinct mutations in different clones, indicating parallel evolution of clones in normal tissue. Mutations in NOTCH1 were the most frequent event in normal bronchial, oesophageal and skin epithelium (squamous epithelia), whereas mutations in epigenetic modifiers DNMT3A and KMT2D were most common in haematopoietic tissue and urothelium respectively (22,23,26,28,35,53,56-63). It is interesting to note that none of these genes is the mostly frequently mutated gene in cancers of corresponding tissues, suggesting that selection pressures in histologically normal tissues are different to those in cancer (52,53).

Somatic mutations in histologically normal tissues where large clonal sweeps are rare

Clonal expansions are limited in size in several tissues, likely due to histological constraints imposed by small functional units, often containing restricted stem cell pools. These include colonic crypts and endometrial glands, and post-mitotic tissues, such as neurons and smooth muscles (20,21,24,25,27,33,34,64). Positive selection of driver mutations was identified within single functional units but there were considerable variations in driver mutation frequency, for example, evidence of positive selection was noted in only 1-5% of colonic crypts and liver micro biopsies, whereas almost 90% endometrial glands in post-menopausal women were replaced by clones with positively selected driver genes (21,25,27). Interestingly, post-mitotic tissue such as neurons and smooth muscles had quantitatively similar mutations per genome per year and comparable mutational spectra with rapidly proliferating tissues such as colon, suggesting that many mutations in histologically normal tissues are replication-independent (20,33,34). Similar to tissues exhibiting large clonal sweeps most mutations were SNVs, the most frequent mutational signatures observed were SBS1 and SBS5/40, and there were differences in the relative frequency of driver gene mutations in these histologically normal tissues compared to their corresponding cancer types (20,21,24,25,33). Thus, the frequency of clones

containing positively selected driver genes was variable as was their size in histologically normal tissues, but irrespectively, they shared comparable mutational signatures (Figure 2).

Understanding the clonal and mutational overlap between histologically normal tissues and cancer

Selective pressures for alterations in driver genes in histologically normal tissues are different to those in the corresponding cancers. In the haematopoietic system, clonal expansion in haematopoietic tissue is associated with a significantly higher risk of haematological malignancies, yet no such obvious link has been noted for solid tissue malignancies (58). In fact, the likely risk of transforming a histologically normal crypt containing a positively selected driver gene into a carcinoma remains extremely low at less than 1 in 3 million over next few decades (21). There is also no direct evidence to suggest that age-matched overall mutational burden increases cancer risk. This is exemplified by the mutational burden in the small intestine, which is comparable with that of the colon at ~50 SNVs/cell/year, yet lifetime cancer risk is significantly lower in small intestine (65). This indicates the insufficiency of positively selected driver mutations and overall mutational burden in histologically normal solid tissues to significantly increase cancer risk. Precisely delineating specific bottlenecks limiting transformation of positively selected histologically normal clones and the vulnerable cell types in which these mutations occur could help differentiate between the earliest genomic aberrations required for cancer initiation from benign age-related changes which do not contribute to tumour development.

How mutated histologically normal tissues differs from cancer

The mutational spectra and clonal dominance observed in histologically normal tissues exhibit three distinct differences compared to genomes and clonal architecture seen in cancers. Firstly, extensive CIN and whole genome duplications are rarely seen in genomes of histologically normal tissues and even LOH and single CNAs are rare (20-37). Secondly, there is a lower rate of tumour suppressor gene (TSG) inactivation in histologically normal tissues. *TP53* mutations are found in over 50% of cancers but are comparatively less frequent in histologically normal tissues, for example *TP53* mutational frequency in histologically normal oesophagus is about 25% as opposed to around 80% in oesophageal

cancer (19). *PTEN* is mutated in about 80% of endometroid cancers but is mutated only in 2% of the normal endometrial glands (25,66), and *APC* mutations, an early and highly frequent event in colon cancer, are exceptionally rare in normal colonic crypts (21). Decreased selection for TSGs may be caused by a lack of CIN which in turn decreases the probability of a second hit via LOH, potentially due to lower frequency of *TP53* mutations. It is important to note that other mechanisms resulting in TSG dysfunction, such as epigenetic inactivation, within histologically normal tissues are still under explored. Thirdly, surrounding microenvironment and wild-type tissue maintain a degree of control over the extent of clonal expansion in mutated histologically normal tissues. For instance, the mouse model evidence suggests that *Trp53* mutated epidermal clones expand exponentially when exposed to UV light whereas histologically normal human epidermal clones with *TP53* mutation in UV-exposed skin remain relatively small (67). Interestingly, once UV exposure is stopped in mouse models, the survival advantage of *Trp53* mutant clones is curtailed by surrounding wild-type epidermal clones (55,67). Evidence for recolonisation of human bronchial epithelium in ex-smokers by clones carrying few tobacco-induced mutations further support the hypothesis that surrounding wild-type clones and local microenvironmental pressures can reverse the dominance of established mutant clones (26). In summary, mutated clones in histologically normal tissues have less complex genomic abnormalities, infrequently mutated tumour suppressor genes, and surrounding wild-type clones with the local microenvironment can impose a degree of control on clonal expansions.

How mutated histologically normal tissues may facilitate cancer transformation

Clones harbouring selected driver gene mutations in histologically normal tissues could promote cancer initiation via non-cell autonomous and cell-autonomous co-operation with pre-existing/subsequent insults. Dominance by a few clones can reduce overall fitness and adaptability in histologically normal tissues, increasing its susceptibility to age-related diseases such as cancer (33,59,68-71). Moreover, some of these mutated clones may also co-operate to promote cancer initiation with neighbouring cells. At least some minor clones can facilitate proliferation of a major clone during cancer progression and treatment response in cell culture and xenotransplantation assays (72,73). Even interaction between mutated and wild-type tissue can also promote cancer initiation as suggested by mouse

9

experiments where Notch1-mutated cells promoted cancer initiation within surrounding Notch1-wild-type cells by promoting inflammatory microenvironment akin to wound healing and by disrupting the skin barrier (74). Somatic mutations in histologically normal tissues can also potentially provide a second hit to pre-existing germline cancer susceptibility variants or could facilitate cancer initiation in conjunction with non-genetic second hits. Large scale tumour sequencing methods estimate that germline variants facilitate cancer initiation in about 2% of all cancer cases within 20 tumour types investigated (75-77). Non-genetic tumour promotor can also induce cancer initiation, as shown classically by Berenblum and Shubik in 1947 (78). They showed that the mouse skin exposed to known skin carcinogens 3:4-benzpyrene or 9:10-dimethyl-1:2-benizanthracene could only transform into frank cancer if the exposure is succeeded by croton oil application, which acted as a tumour promotor. Given that only 3/20 known or suspected common human carcinogens demonstrated discernible mutational signatures, it is possible that a wide variety of carcinogens act as a cancer promoter similar to croton oil (79). Lastly, it is conceivable that histologically normal tissues carrying mutated clones can attract an immune response, overwhelming and perhaps exhausting the immune surveillance network, and allowing cancer initiating clones to face limited scrutiny in some cases. Recent evidence suggests that pre-malignant lesions can attract such an immune response (80-82), however this possibility is unexplored in histologically normal tissues exhibiting widespread mutations and clonal expansions.

Thus, clonality in histologically normal tissue can in principle facilitate as well as inhibit cancer initiation, very likely in a context dependent manner. Assessing clonal architecture in pre-malignant lesions, understanding the cell types in which they occur and longitudinal clonal analyses from histologically normal to pre-malignant lesions could assist in identifying specific contexts in which certain clones in histologically normal tissue potentiate or abrogate cancer initiation.

Insights from the mutational landscape of pre-malignant lesions and inferred early events during cancer initiation

Risk factors for pre-malignant lesions are heterogeneous such as acid reflux for Barret's oesophagus, repeated scarring and regeneration for liver cirrhosis, and long-

10

standing inflammation for inflammatory bowel diseases. Such heterogeneity in pathogenesis and variability in lead time to eventual carcinogenesis results in diverse mutational spectra and clonal evolutionary patterns. However, there are still some commonalities in their evolution. These alterations are studied in detail in the context of evolution of Barrett's oesophagus and its further transformation into oesophageal adenocarcinoma (Figure 3) (83-86). Analysis of DNA methylation signature patterns in breast and lung *carcinoma-in-situ* suggested that epigenetic clonal mosaicism is highest just before cancer initiation (87). Specific epigenetic patterns may therefore guide fates of pre-malignant clones and their interaction with stromal/immune cells allowing cancer initiation (88,89). For example, *DNMTA3A* clonal mutation in haematopoietic cells preferentially hypomethylates polycomb repressive complex 2 targets and specific CpG motifs, favouring a myeloid fate, leading to dysregulated expression of leukaemia stem cell markers (90). It is also possible that these specific DNA methylation changes promote biallelic functional loss of tumour suppressor genes which could then drive clonal evolution in histologically normal tissue towards cancer initiation (91). In addition to epigenetic changes, some pre-malignant clones exhibit short telomeres, suggesting that activating standard or alternate telomere lengthening pathways may be necessary for cancer transformation (92,93). Some pre-malignant clones can undergo catastrophic genomic changes such as chromothripsis and kataegis in the absence of telomere maintenance (94). There is compelling evidence to suggest that these genomic changes happen very early during cancer initiation and could be the primary driver for malignant transformation (9,17,95). Complex genomic changes such as structural abnormalities, CNAs, and HLA LOH are also seen at higher frequency in pre-malignant clones (84,92,96-99). Indeed, the presence of CIN in pre-malignant lesions such as clonal haematopoiesis and Barrett's oesophagus is associated with significantly increased risk of transformation (100-102). Further investigations suggested that cell-intrinsic factors such as HLA LOH and cell-extrinsic stromal and immune system re-education could also facilitate immune evasion permitting carcer initiation (81,97,103,104). Hence, pre-malignant lesions demonstrate specific epigenetic reprograming applying cell fate restrictions, short telomeres, more complex genetic changes, and re-educated stroma and immune system compared to mutated histologically normal tissues.

Investigating how the stroma and immune system dynamically co-evolve with age in mutated histologically normal clonal tissues and neighbouring pre-malignant lesions will be

11

essential for a comprehensive understanding of the cancer initiation process. Preparing a 'precancer atlas' is one such endeavour, which aims to systematically analyse longitudinal clonal architecture of pre-malignant lesions within all major tissues at cell-intrinsic (genomic and epigenetic) and extrinsic (stromal and immune) levels (105,106).

Implication for cancer prevention and early cancer diagnosis

Focussing on unique changes in cancer over and above that of the mutational landscape in histologically normal tissue could be exploited for early cancer intervention. Most cancers diagnosed at an early stage (stage 1-2) have an average 1-year survival of 90-100% with some exceptions, however only about 50% of new cancers are diagnosed early in England (107) (Figure 4). One of the promising avenues to increase early cancer diagnosis by screening for circulating tumour DNA (ctDNA) (108,109). The validity of ctDNA for cancer monitoring is well-established, however using ctDNA for early cancer diagnosis remains challenging due to extremely limited ctDNA shedding at this stage, lack of tumour tissue to inform assays and requirements for ultra-high specificity (110). Exponential advances in DNA sequencing techniques and combination of mutational analysis with copy number and DNA-methylation pattern recognition in commercial tests allow detection of tumour fractions below 10^{-4} in blood with a very high specificity of 99.5% but poor sensitivity for stage 1 cancers (16.5% for DNA-sequencing approach and 39% for methylation-based analysis) in a pan-cancer setting (109-112). DNA from non-malignant clonal haemopoietic expansions can confound several ctDNA assays, however sampling of white cells in blood alongside plasma provides a convenient negative control (113). Recently, circulating cell free DNA derived from pre-malignant colonic adenomas was detected in plasma using a tumour-informed ctDNA assay suggesting future ctDNA early detection tests may be confounded by mutated DNA released from other pre-malignant or even histologically normal tissues across the body (bioRxiv 2022.01.17.476508). In such a case, targeting events specific to malignant tissues such as CNAs or specific epigenetic events may be required. There are currently limited studies focused upon DNA methylation and microenvironmental changes at the earliest stages of cancer development which may assist in refining future ctDNA assays. Our recent work has revealed that pollution activated macrophages can drive non-small cell lung cancer initiation in an *EGFR* mutant mouse model, primarily via the pro-inflammatory

12

cytokine IL1β (Research Square 10.21203/rs.3.rs-1770054/v1). Only the AT2 cells carrying mutant EGFR could transform in the presence of IL1β in these experiments, suggesting that the fate of mutant cells in histologically normal tissue carrying positively selected mutation could be dependent on its subtype. Detection of such intermediary molecules within exosomes purified from serum/other body fluids or even via novel metabolic imaging techniques could also potentially assist in early detection and screening for cancer in specific contexts (114,115). Available data of histologically normal tissue also suggest that genomic, epigenomic, and microenvironmental changes have significant tissue-specific components which likely relate to different mechanisms of malignant progression in these tissues. Cancer-type specific early detection assays may therefore have an advantage in comparison to pan-cancer assays and high-risk populations can also usually be more readily identified for individual cancer types. Indeed, ctDNA fragmentation-based and methylation-based assays for stage A/1 liver and colon cancer detection respectively have demonstrated sensitivity of about 85% (116,117). Lastly, cancer prevention strategies could also be realised when functional determinants and promoters of clonal progression and regression during cancer initiation from histologically normal tissue would be identified. These avenues imply that somatic mutational pattern in histologically normal tissue could readily inform future cancer prevention and intervention strategies (Figure 5).

Conclusions

The mutational spectra of histologically normal tissues are surprisingly complex with clear evidence for positive selection of cancer driver mutations. Clonal diversity decreases with age due to accumulation of driver mutations in clones that outcompete surrounding unmutated clones in proliferative tissues. Tumour initiation may begin by mutations in one such permissive clone acquiring additional genomic or epigenetic alterations with/without support from microenvironment and non-genetic promoter to acquire a pre-malignant phenotype, further increasing in clonal dominance before acquiring the next genetic or epigenetic event to become a founder cancer clone. In tissues where positively selected mutations is restricted to single functional units, increase in clonal size may not be observed until cancer initiation. Somatic mutational landscapes have now been described in more than 25 histologically normal human tissues, however outstanding questions remain about

whether specific germlines SNPs promote differences in clonal architecture and clonal fate, whether similar changes are also present at epigenetic level, and whether stroma and immune system evolve differentially around each clonal expansion. Answering these questions will assist in understanding how fate of each clone is regulated and what determines cancer transformation allowing better delineation of early cancer intervention and prevention approaches.

Acknowledgements

References

1. Boveri T. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. J Cell Sci **2008**;121 Suppl 1:1-84 doi 10.1242/jcs.025742.
2. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A **1971**;68(4):820-3 doi 10.1073/pnas.68.4.820.
3. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell **1990**;61(5):759-67 doi 10.1016/0092-8674(90)90186-i.
4. Ashley DJ. The two "hit" and multiple "hit" theories of carcinogenesis. Br J Cancer **1969**;23(2):313-28 doi 10.1038/bjc.1969.41.
5. Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D*, et al.* The evolutionary history of 2,658 cancers. Nature **2020**;578(7793):122-8 doi 10.1038/s41586-019-1907-7.
6. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature **2013**;501(7467):338-45 doi 10.1038/nature12625.
7. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D*, et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med **2012**;366(10):883-92 doi 10.1056/NEJMoa1113205.
8. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S*, et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. N Engl J Med **2017**;376(22):2109-21 doi 10.1056/NEJMoa1616288.

9.   Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O'Brien T*, et al.* Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell **2018**;173(3):611-23 e17 doi 10.1016/j.cell.2018.02.020.

10.   Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, Marioni JC*, et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci U S A **2013**;110(10):4009-14 doi 10.1073/pnas.1219747110.

11.   Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P*, et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med **2015**;21(7):751-9 doi 10.1038/nm.3886.

12.   Murugaesu N, Wilson GA, Birkbak NJ, Watkins T, McGranahan N, Kumar S*, et al.* Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. Cancer Discov **2015**;5(8):821-31 doi 10.1158/2159-8290.CD-15-0412.

13.   Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J*, et al.* Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science **2014**;346(6206):256-9 doi 10.1126/science.1256930.

14.   McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell **2017**;168(4):613-28 doi 10.1016/j.cell.2017.01.018.

15.   Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov **2022**;12(1):31-46 doi 10.1158/2159-8290.CD-21-1059.

16.   Martincorena I. Somatic mutation and clonal expansions in human tissues. Genome Med **2019**;11(1):35 doi 10.1186/s13073-019-0648-4.

17.   Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X*, et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature **2014**;512(7513):155-60 doi 10.1038/nature13600.

18.   Abbosh C, Venkatesan S, Janes SM, Fitzgerald RC, Swanton C. Evolutionary dynamics in pre-invasive neoplasia. Curr Opin Syst Biol **2017**;2:1-8 doi 10.1016/j.coisb.2017.02.009.

19.   Kakiuchi N, Ogawa S. Clonal expansion in non-cancer tissues. Nat Rev Cancer **2021**;21(4):239-56 doi 10.1038/s41568-021-00335-3.

20.   Abascal F, Harvey LMR, Mitchell E, Lawson ARJ, Lensing SV, Ellis P*, et al.* Somatic mutation landscapes at single-molecule resolution. Nature **2021**;593(7859):405-10 doi 10.1038/s41586-021-03477-4.

21.   Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L*, et al.* The landscape of somatic mutation in normal colorectal epithelial cells. Nature **2019**;574(7779):532-7 doi 10.1038/s41586-019-1672-7.

22.   Martincorena I, Fowler JC, Wabik A, Lawson ARJ, Abascal F, Hall MWJ*, et al.* Somatic mutant clones colonize the human esophagus with age. Science **2018**;362(6417):911-7 doi 10.1126/science.aau3879.

23.   Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S*, et al.* Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science **2015**;348(6237):880-6 doi 10.1126/science.aaa6806.

24.   Moore L, Cagan A, Coorens THH, Neville MDC, Sanghvi R, Sanders MA*, et al.* The mutational landscape of human somatic and germline cells. Nature **2021**;597(7876):381-6 doi 10.1038/s41586-021-03822-7.

25. Moore L, Leongamornlert D, Coorens THH, Sanders MA, Ellis P, Dentro SC*, et al.* The mutational landscape of normal human endometrial epithelium. Nature **2020**;580(7805):640-6 doi 10.1038/s41586-020-2214-z.

26. Yoshida K, Gowers KHC, Lee-Six H, Chandrasekharan DP, Coorens T, Maughan EF*, et al.* Tobacco smoking and somatic mutations in human bronchial epithelium. Nature **2020**;578(7794):266-72 doi 10.1038/s41586-020-1961-1.

27. Brunner SF, Roberts ND, Wylie LA, Moore L, Aitken SJ, Davies SE*, et al.* Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. Nature **2019**;574(7779):538-42 doi 10.1038/s41586-019-1670-9.

28. Li R, Du Y, Chen Z, Xu D, Lin T, Jin S*, et al.* Macroscopic somatic clonal expansion in morphologically normal human urothelium. Science **2020**;370(6512):82-9 doi 10.1126/science.aba7300.

29. Blokzijl F, de Ligt J, Jager M, Sasselli V, Roerink S, Sasaki N*, et al.* Tissue-specific mutation accumulation in human adult stem cells during life. Nature **2016**;538(7624):260-4 doi 10.1038/nature19768.

30. Franco I, Johansson A, Olsson K, Vrtacnik P, Lundin P, Helgadottir HT*, et al.* Somatic mutagenesis in satellite cells associates with human skeletal muscle aging. Nat Commun **2018**;9(1):800 doi 10.1038/s41467-018-03244-6.

31. Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC*, et al.* The origin and evolution of mutations in acute myeloid leukemia. Cell **2012**;150(2):264-78 doi 10.1016/j.cell.2012.06.023.

32. Brazhnik K, Sun S, Alani O, Kinkhabwala M, Wolkoff AW, Maslov AY*, et al.* Single-cell analysis reveals different age-related somatic mutation profiles between stem and differentiated cells in human liver. Sci Adv **2020**;6(5):eaax2659 doi 10.1126/sciadv.aax2659.

33. Lodato MA, Rodin RE, Bohrson CL, Coulter ME, Barton AR, Kwon M*, et al.* Aging and neurodegeneration are associated with increased mutations in single human neurons. Science **2018**;359(6375):555-9 doi 10.1126/science.aao4426.

34. Lodato MA, Woodworth MB, Lee S, Evrony GD, Mehta BK, Karger A*, et al.* Somatic mutation in single human neurons tracks developmental and transcriptional history. Science **2015**;350(6256):94-8 doi 10.1126/science.aab1785.

35. Yokoyama A, Kakiuchi N, Yoshizato T, Nannya Y, Suzuki H, Takeuchi Y*, et al.* Age-related remodelling of oesophageal epithelia by mutated cancer drivers. Nature **2019**;565(7739):312-7 doi 10.1038/s41586-018-0811-x.

36. Cooper CS, Eeles R, Wedge DC, Van Loo P, Gundem G, Alexandrov LB*, et al.* Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. Nat Genet **2015**;47(4):367-72 doi 10.1038/ng.3221.

37. Mitchell E, Spencer Chapman M, Williams N, Dawson KJ, Mende N, Calderbank EF*, et al.* Clonal dynamics of haematopoiesis across the human lifespan. Nature **2022**;606(7913):343-50 doi 10.1038/s41586-022-04786-y.

38. Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P*, et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. Cell **2017**;171(5):1029-41 e21 doi 10.1016/j.cell.2017.09.042.

39. Robinson PS, Coorens THH, Palles C, Mitchell E, Abascal F, Olafsson S*, et al.* Increased somatic mutation burdens in normal human cells due to defective DNA polymerases. Nat Genet **2021**;53(10):1434-42 doi 10.1038/s41588-021-00930-y.

16

40. Koeppel M, Garcia-Alcalde F, Glowinski F, Schlaermann P, Meyer TF. Helicobacter pylori Infection Causes Characteristic DNA Damage Patterns in Human Cells. Cell Rep **2015**;11(11):1703-13 doi 10.1016/j.celrep.2015.05.030.

41. Lennard KS, Goosen RW, Blackburn JM. Bacterially-Associated Transcriptional Remodelling in a Distinct Genomic Subtype of Colorectal Cancer Provides a Plausible Molecular Basis for Disease Development. PLoS One **2016**;11(11):e0166282 doi 10.1371/journal.pone.0166282.

42. Gall A, Fero J, McCoy C, Claywell BC, Sanchez CA, Blount PL*, et al.* Bacterial Composition of the Human Upper Gastrointestinal Tract Microbiome Is Dynamic and Associated with Genomic Instability in a Barrett's Esophagus Cohort. PLoS One **2015**;10(6):e0129055 doi 10.1371/journal.pone.0129055.

43. Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, van Hoeck A, Wood HM, Nomburg J*, et al.* Mutational signature in colorectal cancer caused by genotoxic pks(+) E. coli. Nature **2020**;580(7802):269-73 doi 10.1038/s41586-020-2080-8.

44. Pauklin S, Sernandez IV, Bachmann G, Ramiro AR, Petersen-Mahrt SK. Estrogen directly activates AID transcription and function. J Exp Med **2009**;206(1):99-111 doi 10.1084/jem.20080521.

45. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y*, et al.* The repertoire of mutational signatures in human cancer. Nature **2020**;578(7793):94-101 doi 10.1038/s41586-020-1943-3.

46. Kucab JE, Zou X, Morganella S, Joel M, Nanda AS, Nagy E*, et al.* A Compendium of Mutational Signatures of Environmental Agents. Cell **2019**;177(4):821-36 e16 doi 10.1016/j.cell.2019.03.001.

47. Steele CD, Abbasi A, Islam SMA, Bowes AL, Khandekar A, Haase K*, et al.* Signatures of copy number alterations in human cancer. Nature **2022**;606(7916):984-91 doi 10.1038/s41586-022-04738-6.

48. Drews RM, Hernando B, Tarabichi M, Haase K, Lesluyes T, Smith PS*, et al.* A pan-cancer compendium of chromosomal instability. Nature **2022**;606(7916):976-83 doi 10.1038/s41586-022-04789-9.

49. Ren P, Dong X, Vijg J. Age-related somatic mutation burden in human tissues. Front Aging **2022**;3:1018119 doi 10.3389/fragi.2022.1018119.

50. Kader T, Zethoven M, Gorringe KL. Evaluating statistical approaches to define clonal origin of tumours using bulk DNA sequencing: context is everything. Genome Biol **2022**;23(1):43 doi 10.1186/s13059-022-02600-6.

51. Black JRM, McGranahan N. Genetic and non-genetic clonal diversity in cancer evolution. Nat Rev Cancer **2021**;21(6):379-92 doi 10.1038/s41568-021-00336-2.

52. Buhigas C, Warren AY, Leung WK, Whitaker HC, Luxton HJ, Hawkins S*, et al.* The architecture of clonal expansions in morphologically normal tissue from cancerous and non-cancerous prostates. Mol Cancer **2022**;21(1):183 doi 10.1186/s12943-022-01644-3.

53. Lawson ARJ, Abascal F, Coorens THH, Hooks Y, O'Neill L, Latimer C*, et al.* Extensive heterogeneity in somatic mutation and selection in the human bladder. Science **2020**;370(6512):75-82 doi 10.1126/science.aba8347.

54. Colom B, Alcolea MP, Piedrafita G, Hall MWJ, Wabik A, Dentro SC*, et al.* Spatial competition shapes the dynamic mutational landscape of normal esophageal epithelium. Nat Genet **2020**;52(6):604-14 doi 10.1038/s41588-020-0624-3.

55.   Murai K, Skrupskelyte G, Piedrafita G, Hall M, Kostiou V, Ong SH, *et al.* Epidermal Tissue Adapts to Restrain Progenitors Carrying Clonal p53 Mutations. Cell Stem Cell **2018**;23(5):687-99 e8 doi 10.1016/j.stem.2018.08.017.

56.   Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nat Med **2014**;20(12):1472-8 doi 10.1038/nm.3733.

57.   Yoshizato T, Dumitriu B, Hosokawa K, Makishima H, Yoshida K, Townsley D, *et al.* Somatic Mutations and Clonal Hematopoiesis in Aplastic Anemia. N Engl J Med **2015**;373(1):35-47 doi 10.1056/NEJMoa1414799.

58.   Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med **2014**;371(26):2477-87 doi 10.1056/NEJMoa1409405.

59.   Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med **2014**;371(26):2488-98 doi 10.1056/NEJMoa1408617.

60.   Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL, *et al.* Functional genomic landscape of acute myeloid leukaemia. Nature **2018**;562(7728):526-31 doi 10.1038/s41586-018-0623-z.

61.   Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med **2016**;374(23):2209-21 doi 10.1056/NEJMoa1516192.

62.   Grinfeld J, Nangalia J, Baxter EJ, Wedge DC, Angelopoulos N, Cantrill R, *et al.* Classification and Personalized Prognosis in Myeloproliferative Neoplasms. N Engl J Med **2018**;379(15):1416-30 doi 10.1056/NEJMoa1716614.

63.   Desai P, Mencia-Trinchant N, Savenkov O, Simon MS, Cheang G, Lee S, *et al.* Somatic mutations precede acute myeloid leukemia years before diagnosis. Nat Med **2018**;24(7):1015-23 doi 10.1038/s41591-018-0081-z.

64.   Kakiuchi N, Yoshida K, Uchino M, Kihara T, Akaki K, Inoue Y, *et al.* Frequent mutations that converge on the NFKBIZ pathway in ulcerative colitis. Nature **2020**;577(7789):260-5 doi 10.1038/s41586-019-1856-1.

65.   Li R, Di L, Li J, Fan W, Liu Y, Guo W, *et al.* A body map of somatic mutagenesis in morphologically normal human tissues. Nature **2021**;597(7876):398-403 doi 10.1038/s41586-021-03836-1.

66.   Suda K, Nakaoka H, Yoshihara K, Ishiguro T, Tamura R, Mori Y, *et al.* Clonal Expansion and Diversification of Cancer-Associated Mutations in Endometriosis and Normal Endometrium. Cell Rep **2018**;24(7):1777-89 doi 10.1016/j.celrep.2018.07.037.

67.   Klein AM, Brash DE, Jones PH, Simons BD. Stochastic fate of p53-mutant epidermal progenitor cells is tilted toward proliferation by UV B during preneoplasia. Proc Natl Acad Sci U S A **2010**;107(1):270-5 doi 10.1073/pnas.0909738107.

68.   Markert JA, Champlin DM, Gutjahr-Gobell R, Grear JS, Kuhn A, McGreevy TJ, Jr*., et al.* Population genetic diversity and fitness in multiple environments. BMC Evol Biol **2010**;10:205 doi 10.1186/1471-2148-10-205.

69.   Ellegren H, Galtier N. Determinants of genetic diversity. Nat Rev Genet **2016**;17(7):422-33 doi 10.1038/nrg.2016.58.

70.   Reed DH, Frankham R. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. Evolution **2001**;55(6):1095-103 doi 10.1111/j.0014-3820.2001.tb00629.x.

71. Jaiswal S, Natarajan P, Silver AJ, Gibson CJ, Bick AG, Shvartz E*, et al.* Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. N Engl J Med **2017**;377(2):111-21 doi 10.1056/NEJMoa1701719.

72. Marusyk A, Tabassum DP, Altrock PM, Almendro V, Michor F, Polyak K. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature **2014**;514(7520):54-8 doi 10.1038/nature13556.

73. Hobor S, Van Emburgh BO, Crowley E, Misale S, Di Nicolantonio F, Bardelli A. TGFalpha and amphiregulin paracrine network promotes resistance to EGFR blockade in colorectal cancer cells. Clin Cancer Res **2014**;20(24):6429-38 doi 10.1158/1078-0432.CCR-14-0774.

74. Demehri S, Turkoz A, Kopan R. Epidermal Notch1 loss promotes skin tumorigenesis by impacting the stromal microenvironment. Cancer Cell **2009**;16(1):55-66 doi 10.1016/j.ccr.2009.05.016.

75. Park S, Supek F, Lehner B. Systematic discovery of germline cancer predisposition genes through the identification of somatic second hits. Nat Commun **2018**;9(1):2601 doi 10.1038/s41467-018-04900-7.

76. Yngvadottir B, Andreou A, Bassaganyas L, Larionov A, Cornish AJ, Chubb D*, et al.* Frequency of pathogenic germline variants in cancer susceptibility genes in 1336 renal cell carcinoma cases. Hum Mol Genet **2022**;31(17):3001-11 doi 10.1093/hmg/ddac089.

77. Uson PLS, Jr., Riegert-Johnson D, Boardman L, Kisiel J, Mountjoy L, Patel N*, et al.* Germline Cancer Susceptibility Gene Testing in Unselected Patients With Colorectal Adenocarcinoma: A Multicenter Prospective Study. Clin Gastroenterol Hepatol **2022**;20(3):e508-e28 doi 10.1016/j.cgh.2021.04.013.

78. Berenblum I, Shubik P. The role of croton oil applications, associated with a single painting of a carcinogen, in tumour induction of the mouse's skin. Br J Cancer **1947**;1(4):379-82 doi 10.1038/bjc.1947.35.

79. Riva L, Pandiri AR, Li YR, Droop A, Hewinson J, Quail MA*, et al.* The mutational signature profile of known and suspected human carcinogens in mice. Nat Genet **2020**;52(11):1189-97 doi 10.1038/s41588-020-0692-4.

80. Pennycuick A, Teixeira VH, AbdulJabbar K, Raza SEA, Lund T, Akarca AU*, et al.* Immune Surveillance in Clinical Regression of Preinvasive Squamous Cell Lung Cancer. Cancer Discov **2020**;10(10):1489-99 doi 10.1158/2159-8290.CD-19-1366.

81. Mascaux C, Angelova M, Vasaturo A, Beane J, Hijazi K, Anthoine G*, et al.* Immune evasion before tumour invasion in early lung squamous carcinogenesis. Nature **2019**;571(7766):570-5 doi 10.1038/s41586-019-1330-0.

82. Beane JE, Mazzilli SA, Campbell JD, Duclos G, Krysan K, Moy C*, et al.* Molecular subtyping reveals immune alterations associated with progression of bronchial premalignant lesions. Nat Commun **2019**;10(1):1856 doi 10.1038/s41467-019-09834-2.

83. Paulson TG, Maley CC, Li X, Li H, Sanchez CA, Chao DL*, et al.* Chromosomal instability and copy number alterations in Barrett's esophagus and esophageal adenocarcinoma. Clin Cancer Res **2009**;15(10):3305-14 doi 10.1158/1078-0432.CCR-08-2494.

84. Stachler MD, Camarda ND, Deitrick C, Kim A, Agoston AT, Odze RD*, et al.* Detection of Mutations in Barrett's Esophagus Before Progression to High-Grade Dysplasia or

Adenocarcinoma. Gastroenterology **2018**;155(1):156-67 doi 10.1053/j.gastro.2018.03.047.

85. Frankell AM, Jammula S, Li X, Contino G, Killcoyne S, Abbas S*, et al.* The landscape of selection in 551 esophageal adenocarcinomas defines genomic biomarkers for the clinic. Nat Genet **2019**;51(3):506-16 doi 10.1038/s41588-018-0331-5.

86. Contino G, Vaughan TL, Whiteman D, Fitzgerald RC. The Evolving Genomic Landscape of Barrett's Esophagus and Esophageal Adenocarcinoma. Gastroenterology **2017**;153(3):657-73 e1 doi 10.1053/j.gastro.2017.07.007.

87. Yang Z, Wong A, Kuh D, Paul DS, Rakyan VK, Leslie RD*, et al.* Correlation of an epigenetic mitotic clock with cancer risk. Genome Biol **2016**;17(1):205 doi 10.1186/s13059-016-1064-3.

88. Vicente-Duenas C, Hauer J, Cobaleda C, Borkhardt A, Sanchez-Garcia I. Epigenetic Priming in Cancer Initiation. Trends Cancer **2018**;4(6):408-17 doi 10.1016/j.trecan.2018.04.007.

89. Yan J, Chen Y, Patel AJ, Warda S, Lee CJ, Nixon BG*, et al.* Tumor-intrinsic PRC2 inactivation drives a context-dependent immune-desert microenvironment and is sensitized by immunogenic viruses. J Clin Invest **2022**;132(17) doi 10.1172/JCI153437.

90. Nam AS, Dusaj N, Izzo F, Murali R, Myers RM, Mouhieddine TH*, et al.* Single-cell multi-omics of human clonal hematopoiesis reveals that DNMT3A R882 mutations perturb early progenitor states through selective hypomethylation. Nat Genet **2022**;54(10):1514-26 doi 10.1038/s41588-022-01179-9.

91. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L*, et al.* A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet **2007**;39(2):237-42 doi 10.1038/ng1972.

92. Huang KK, Ramnarayanan K, Zhu F, Srivastava S, Xu C, Tan ALK*, et al.* Genomic and Epigenomic Profiling of High-Risk Intestinal Metaplasia Reveals Molecular Determinants of Progression to Gastric Cancer. Cancer Cell **2018**;33(1):137-50 e5 doi 10.1016/j.ccell.2017.11.018.

93. Raynaud CM, Jang SJ, Nuciforo P, Lantuejoul S, Brambilla E, Mounier N*, et al.* Telomere shortening is correlated with the DNA damage response and telomeric protein down-regulation in colorectal preneoplastic lesions. Ann Oncol **2008**;19(11):1875-81 doi 10.1093/annonc/mdn405.

94. Maciejowski J, Li Y, Bosco N, Campbell PJ, de Lange T. Chromothripsis and Kataegis Induced by Telomere Crisis. Cell **2015**;163(7):1641-54 doi 10.1016/j.cell.2015.11.054.

95. Bolkestein M, Wong JKL, Thewes V, Korber V, Hlevnjak M, Elgaafary S*, et al.* Chromothripsis in Human Breast Cancer. Cancer Res **2020**;80(22):4918-31 doi 10.1158/0008-5472.CAN-20-1920.

96. Teixeira VH, Pipinikas CP, Pennycuick A, Lee-Six H, Chandrasekharan D, Beane J*, et al.* Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. Nat Med **2019**;25(3):517-25 doi 10.1038/s41591-018-0323-0.

97. Chen H, Carrot-Zhang J, Zhao Y, Hu H, Freeman SS, Yu S*, et al.* Genomic and immune profiling of pre-invasive lung adenocarcinoma. Nat Commun **2019**;10(1):5472 doi 10.1038/s41467-019-13460-3.

98. Boutros PC, Fraser M, Harding NJ, de Borja R, Trudel D, Lalonde E, *et al.* Spatial genomic heterogeneity within localized, multifocal prostate cancer. Nat Genet **2015**;47(7):736-45 doi 10.1038/ng.3315.

99. Shain AH, Yeh I, Kovalyshyn I, Sriharan A, Talevich E, Gagnon A, *et al.* The Genetic Evolution of Melanoma from Precursor Lesions. N Engl J Med **2015**;373(20):1926-36 doi 10.1056/NEJMoa1502583.

100. Terao C, Suzuki A, Momozawa Y, Akiyama M, Ishigaki K, Yamamoto K, *et al.* Chromosomal alterations among age-related haematopoietic clones in Japan. Nature **2020**;584(7819):130-5 doi 10.1038/s41586-020-2426-2.

101. Loh PR, Genovese G, Handsaker RE, Finucane HK, Reshef YA, Palamara PF, *et al.* Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. Nature **2018**;559(7714):350-5 doi 10.1038/s41586-018-0321-x.

102. Killcoyne S, Gregson E, Wedge DC, Woodcock DJ, Eldridge MD, de la Rue R, *et al.* Genomic copy number predicts esophageal cancer years before transformation. Nat Med **2020**;26(11):1726-32 doi 10.1038/s41591-020-1033-y.

103. Gil Del Alcazar CR, Huh SJ, Ekram MB, Trinh A, Liu LL, Beca F, *et al.* Immune Escape in Breast Cancer During In Situ to Invasive Carcinoma Transition. Cancer Discov **2017**;7(10):1098-115 doi 10.1158/2159-8290.CD-17-0222.

104. Krtolica A, Campisi J. Cancer and aging: a model for the cancer promoting effects of the aging stroma. Int J Biochem Cell Biol **2002**;34(11):1401-14 doi 10.1016/s1357-2725(02)00053-5.

105. Spira A, Yurgelun MB, Alexandrov L, Rao A, Bejar R, Polyak K, *et al.* Precancer Atlas to Drive Precision Prevention Trials. Cancer Res **2017**;77(7):1510-41 doi 10.1158/0008-5472.CAN-16-2346.

106. Campbell JD, Mazzilli SA, Reid ME, Dhillon SS, Platero S, Beane J, *et al.* The Case for a Pre-Cancer Genome Atlas (PCGA). Cancer Prev Res (Phila) **2016**;9(2):119-24 doi 10.1158/1940-6207.CAPR-16-0024.

107. Public Health England. Cancer survival by stage at diagnosis for England (experimental statistics): Adults diagnosed 2012, 2013 and 2014 and followed up to 2015. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalbystageatdiagnosisforenglandexperimentalstatistics/adultsdiagnosed20122013and2014andfollowedupto2015 Accessed on 01/09/2023. 10/06/2016 ed2016.

108. Sanz-Garcia E, Zhao E, Bratman SV, Siu LL. Monitoring and adapting cancer treatment using circulating tumor DNA kinetics: Current research, opportunities, and challenges. Sci Adv **2022**;8(4):eabi8618 doi 10.1126/sciadv.abi8618.

109. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Consortium C. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol **2020**;31(6):745-59 doi 10.1016/j.annonc.2020.02.011.

110. Klein EA, Beer TM, Seiden M. The Promise of Multicancer Early Detection. Comment on Pons-Belda et al. Can Circulating Tumor DNA Support a Successful Screening Test for Early Cancer Detection? The Grail Paradigm. Diagnostics 2021, 11, 2171. Diagnostics (Basel) **2022**;12(5) doi 10.3390/diagnostics12051243.

111. Nadauld LD, McDonnell CH, 3rd, Beer TM, Liu MC, Klein EA, Hudnut A, *et al.* The PATHFINDER Study: Assessment of the Implementation of an Investigational Multi-

Cancer Early Detection Test into Clinical Practice. Cancers (Basel) **2021**;13(14) doi 10.3390/cancers13143501.

112.   Klein EA, Richards D, Cohn A, Tummala M, Lapham R, Cosgrove D, *et al.* Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. Ann Oncol **2021**;32(9):1167-77 doi 10.1016/j.annonc.2021.05.806.

113.   Abbosh C, Swanton C, Birkbak NJ. Clonal haematopoiesis: a source of biological noise in cell-free DNA analyses. Ann Oncol **2019**;30(3):358-9 doi 10.1093/annonc/mdy552.

114.   Pouli D, Thieu HT, Genega EM, Baecher-Lind L, House M, Bond B*, et al.* Label-free, High-Resolution Optical Metabolic Imaging of Human Cervical Precancers Reveals Potential for Intraepithelial Neoplasia Diagnosis. Cell Rep Med **2020**;1(2) doi 10.1016/j.xcrm.2020.100017.

115.   Wang X, Tian L, Lu J, Ng IO. Exosomes and cancer - Diagnostic and prognostic biomarkers and therapeutic vehicle. Oncogenesis **2022**;11(1):54 doi 10.1038/s41389-022-00431-5.

116.   Mo S, Dai W, Wang H, Lan X, Ma C, Su Z*, et al.* Early detection and prognosis prediction for colorectal cancer by circulating tumour DNA methylation haplotypes: A multicentre cohort study. EClinicalMedicine **2023**;55:101717 doi 10.1016/j.eclinm.2022.101717.

117.   Foda ZH, Annapragada AV, Boyapati K, Bruhm DC, Vulpescu NA, Medina JE*, et al.* Detecting Liver Cancer Using Cell-Free DNA Fragmentomes. Cancer Discov **2023**;13(3):616-31 doi 10.1158/2159-8290.CD-22-0659.

Figure legends:

Figure 1: Sources of mutations in histologically normal tissues. The mutations were assessed in epithelial cells of solid organs and a variety of haematopoietic cells while assessing blood or bone marrow. The samples were derived from individuals with no known cancer diagnosis for the assessed tissue on histological examination (Created with BioRender.com).

Figure 2: Changes in histologically normal tissues with age a. Large clones in rapidly dividing tissues without distinct anatomical units (e.g., oesophagus and skin) b. Restricted clonal size to anatomical unit in minority of tissue units (e.g., colon) or majority of tissue units (e.g., endometrium) (Created with BioRender.com).

Figure 3: Changes in the mutational spectra in the histologically normal oesophagus and Barrett's oesophagus compared with oesophageal dysplasia and adenocarcinoma.

Figure 4: Cancer diagnosis (a) and 1-year survival as per stage (b) in common cancer diagnosed across England between 2012-14 (public health England data).

Figure 5: Potential determinants of carcinogenic transformation of histologically normal tissue carrying clonally expanded mutations in cancer driver genes and its application to design the early intervention strategies for cancer prevention and detection (Created with BioRender.com).
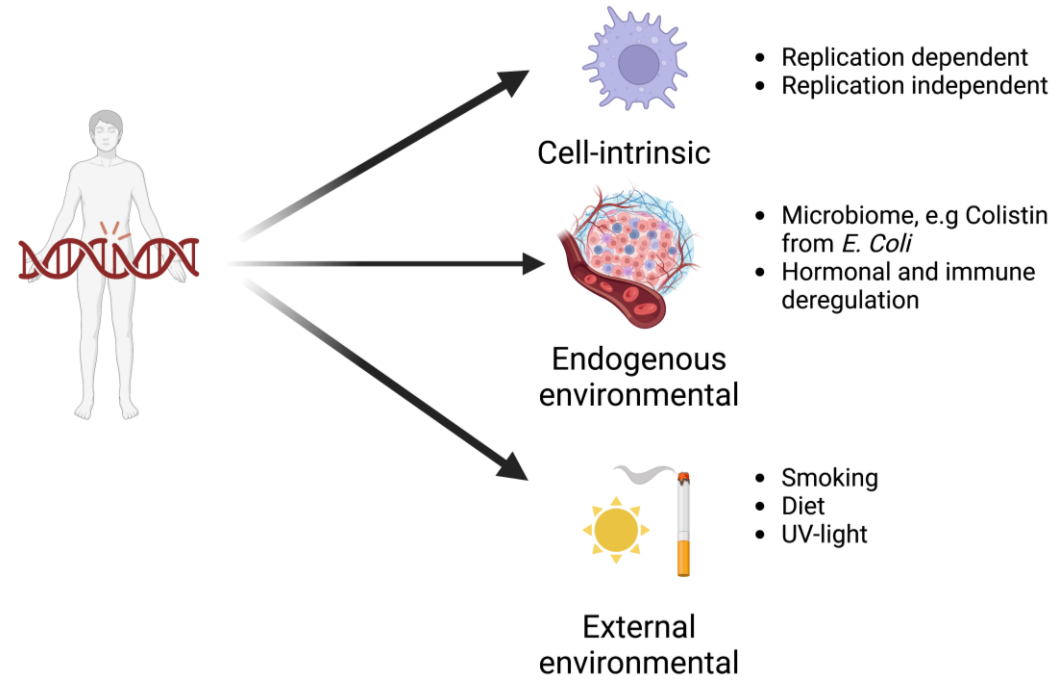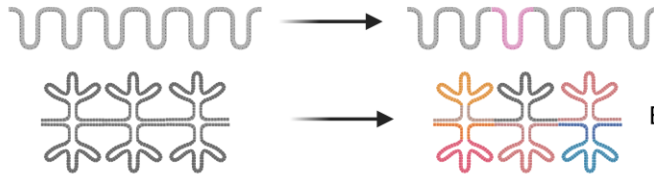
Figure 1

Figure 2

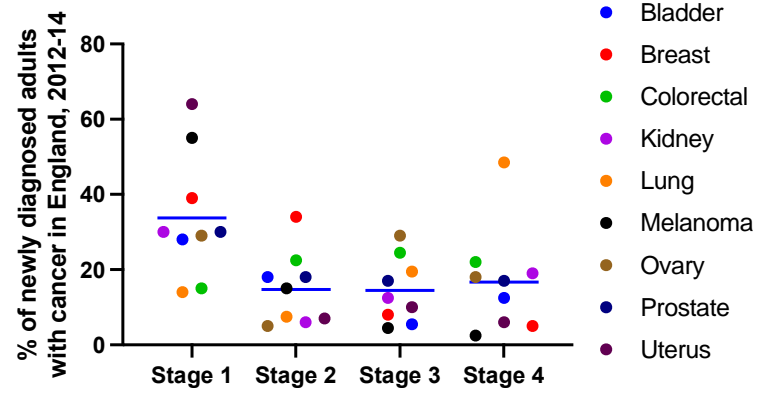|	Normal oesophagus	Barrett's oesophagus → Dysplasia → Oesophageal adenocarcinoma		
	Histologically normal tissue	Histologically abnormal tissues		
• SNV	++	+++	+++	++++
• Significant chromosomal instability (CIN)	-	Uncommon	+	+++
• GD	-	Uncommon	+	+++
• Commonest positively selected/mutated genes	*NOTCH1/2/3, TP53, FAT1*	*TP53, CDKN2A, ARID1A, ARID1B, APC*	*TP53, CDKN2A, ERBB2, FGFR, MET*	*TP53, CDKN2A, KRAS, MYC, ERBB2*
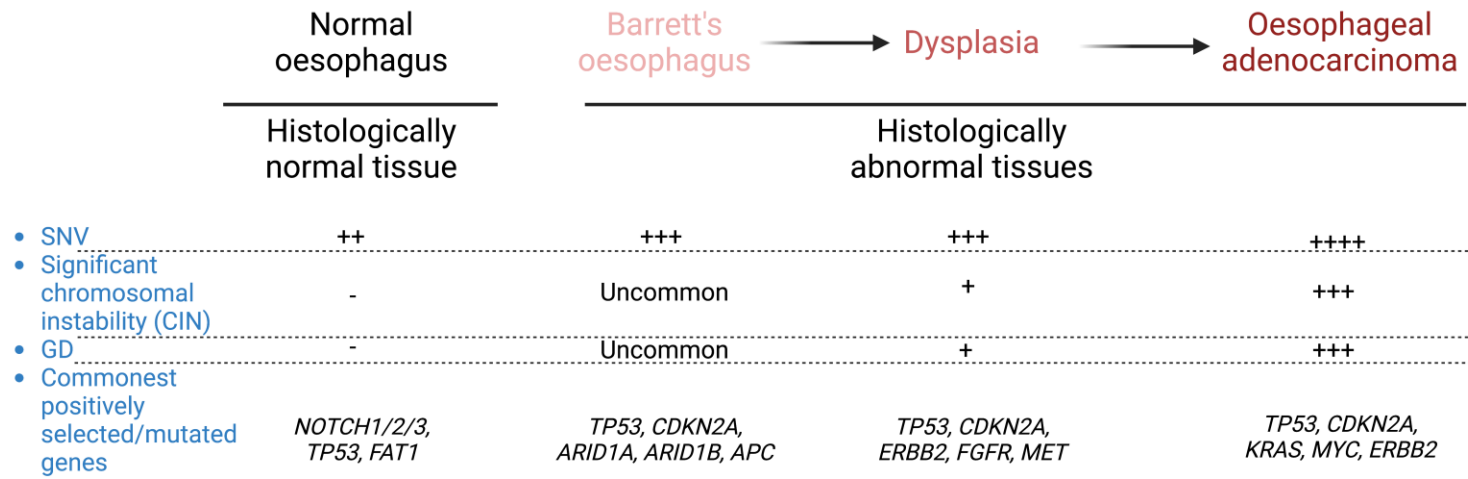
Figure 3

a. Stage at diagnosis

b. 1-year survival as per stage at diagnosis

Figure 4

Figure 5