Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by:

M.Sc. Megan Druce

Born in: Durban, South Africa

Oral examination: 26.01.21

The impact of ageing, replication and stress on genome stability in hematopoietic stem cells

Referees:

Prof. Dr. Benedikt Brors

Dr. Michael Milsom

The research presented in this dissertation was performed from July 2016 until November 2020 under the supervision of Dr. Michael Milsom in the Division of Experimental Hematology at the German Cancer Research Center (DKFZ) and the Heidelberg Institute for Stem Cell Technology and Experimental Medicine (HI-STEM), as well as Prof. Dr. Benedikt Brors in the Division of Applied Bioinformatics at DKFZ in Heidelberg, Germany.

……

Megan Druce

Acknowledgments

First, I would like to thank my supervisor **Dr. Michael Milsom**. Thank you for taking a big leap of faith and giving me the opportunity to join your lab for my PhD. The time and effort you put into my training went above and beyond what I could have hoped for. This project had a lot of initial hurdles to overcome but you were always there ready to help in any way possible- even when it meant sending me to another country to get the resources I needed to ultimately bring this project forward. Beyond the academic supervision, thank you for your guidance and advice regarding my future career steps.

A special thanks to my second supervisor, **Prof. Dr. Benedikt Brors**. I have been very lucky having not one, but two great supervisors. Thank you for your support and valid input when it came to the Bioinformatics analysis. This was a very steep learning curve for me but you always had great ideas for how to solve issues I was having.

Thank you to **Jun.-Prof. Dr. Lazaro Centanin** and **Prof. Dr. Michaela Frye** for taking the time out of your schedules to be a part of my PhD defense committee. Additionally, thank you to my thesis committee members **Prof. Dr. med. Carsten Müller-Tidow** and **Dr. Jan Korbel** for all of your advice (scientific and non-scientific related) and the many academic discussions.

A very big thank you to **Dr. Inigo Martincorena** and **Dr. Alex Cagan**, as well as the rest of the Martincorena and Campbell labs, and a special shout out to **Raul Alcantara**. Your help during and following my visit to Sanger was vital to this project, as well as the development of my analysis skills. Thank you for making me feel so welcome during my short time there and taking the time to teach me.

A huge thank you to **Marleen Büchler-Schäff**- I cannot thank you enough for the time you spent proof-reading my thesis, not to mention all of your help with the clone assays. Late sorter nights don't seem so bad when you're there (and of course, pizza). You have been a great lab mate but also a great friend these past four years.

A big thank you to **Julia Knoch** and **Dr. Susanne Lux** for helping me with the library preps. I can honestly say there is no way I would have finished that many libraries in that short space of time without the help of you two! An additional thank you to Susi for proof-reading my results.

An additional big thank you to **Dr. Ruzhica Bogeska**. Thank you Ruzh for spending all of those years meticulously collecting a huge stock-pile of colonies. Even with all of the sequencing we have done between us- I think we have only touched the surface of all the samples we have. We probably will need to start busking to sequence all of them.

Thank you to **Dr. Charles Imbusch**. You started this project with Ruzhi and taught me so much when I started. Beyond that, thank you for being a good friend- we really do need to have another beer with Roma sometime!

To all of the past and present members of the Milsom lab- **Natasha Anstee, Melanie Ball, Ruzhica Bogeska, Marleen Büchler-Schaff, Fenia Fotopoulou, Julius Gräsel, Jeyan Jayarajan, Paul Kaschutnig, Julia Knoch, Susanne Lux, Ana-Matea Mikecin, Esther Rodriguez Correa, Sina Stäbl**e, and our honorary lab member **Jens Langstein** (wow, that was a lot). You guys are great, I could not have asked for a better work environment and TEAM. The Milsom lab is special, I've never known a lab to help each other out so much and actually enjoy the time we spend together. Who knew early mornings prepping mice could be filled with so much laughter (and screams). I will cherish every happy memory, especially those that involved cake.

Thank you to my **Applied Bioinformatics lab**- I wish I could have spent more time with you guys but the time I did spend there was always a happy one. Special thanks to **Lina Sieverling** for helping me with TelomereHunter, and **Hans Bartemeß** and **Frank Thommen** for helping me with my many questions and requests.

Thank you to the Genomics and Proteomics, Omics IT and Data Management, and Light Microscopy DKFZ Core Facility, especially Ursula Fürst, Ivo Buchhalter, Damir Krunic for all of your help and advice. A special thanks to all members of the Flow Cytometry DKFZ Core Facility for helping with the sorts.

Thank you to **Dr. Andreas Trumpp** for creating such a great working environment at HI-STEM. And to all past and present members of HI-STEM- I have enjoyed the many conversations, celebrations and laughs over the years. A special thanks to **Dagmar Wolf** and **Erika Krückel** for the support of all things non-science related.

And now to thank those who have supported me outside of the lab.

To the bros: **Mattia "Boo" Falcone, Manuel "Dingo" Reitberger** and **Stewart "Macklemore" Mein**. You made life in Heidelberg so much better. I don't think I can remember a time out in Untere without at least one of you there with me. Thank you for the many laughs, I'm looking forward to many more soon (and Samurai Sword).

To **Natasha Anstee, Jens Langstein, Nicco Arecco, Sina Stäble** and **Paula Werner**- thank you for the Italian trips, cheesecake and coffee, Christmas parties, stuzzico pizza and many other good times.

To **Marie Groth** and **Julius Gräsel**- thank you for being my honorary German parents (sorry Marie, I had to). Julius, I am so happy that I lucked out with you as a lab mate. Not only did you take me under your wing when I started, but you brought Marie into my life too. You two have been two of the best friends I could have asked for and probably the only couple I would ever dare go on vacation alone with. Thank you for everything and I look forward to many more years of memories together.

To **Traudel Berkes, Sandra** and **Florian Strasse** and the rest of the Strasse gang, especially little Clara and Timmy. Thank you for welcoming me into your homes and family.

To **Max**. You may have only joined me on my PhD path towards the end of it but your support has helped me during the time it mattered most. I cannot imagine how I would have gotten through this last year without you by my side. Thank you for keeping me calm and making me laugh when the stress got the better of me. I have enjoyed every moment with you and look forward to making many more memories. Oh, and thanks for proof-reading the zusammenfassung- I'm really happy you understood it!

Last but definitely not least, my biggest thanks goes to my family in South Africa. Leaving SA and moving so far away was one of the hardest things I have ever done. Thank you for your support and believing in me and my future. I have missed many big moments but knowing you are always there cheering me on from a distance has made this difficult journey worth it. Thank you to the rest of my family spread all over the world- I know we had planned a big reunion to celebrate my defense but even having your support from a distance means the world to me. A special thanks to my Prague family: Derek, Renee, Jody and Daniel. Thank you for the visits to Heidelberg and for letting me have some much-needed family time in Prague. To my sister- **Kerryn Druce** aka Dude, thank you for being my inspiration (see, I made it official). To my dad, **Stephen Druce**, thank you for your support. And to my mom, **June Druce**, thank you for believing in my dream even though it meant me being so far away. I hope I made you all proud.

Table of contents

I. Abstract

The acquisition of mutations within the genome of hematopoietic stem cells (HSCs) is of particular importance as this is a likely driver of malignant transformation for many leukemias, as well as a hallmark of ageing. Importantly, the study of normal physiologic mediators of HSC mutation acquisition is a largely underdeveloped area. The advent of next generation sequencing (NGS) technologies provides a route to interrogate this phenomenon but is complicated by the fact that a genome-wide analysis at the clonal level is necessary to determine the mutational signature of individual HSCs.

We have developed an *in vitro* model that allows for the sensitive assessment of clonal mutations occurring within single HSCs using whole genome sequencing. In order to optimize our clonal mutation analysis, we performed a benchmarking exercise where we deeply sequenced an individual HSC colony to ~90X coverage and performed somatic nucleotide variant (SNV) analysis at various down-sampled coverages. Importantly, we established that the number of mutations called increases in an almost linear fashion with increasing coverage, until a plateau is reached at around 30X coverage. We additionally developed optimal filtering parameters, which demonstrated a much-improved capacity for discerning true positive and false positive mutations at low coverage, compared to previously published methodologies.

Using this optimized sequencing pipeline, we collected and sequenced HSC clones from young and old mice, as well as those exposed to stress agonists known to induce HSC cycling. We additionally employed a genetic label-retention system to segregate dormant and actively cycling HSCs in order to assess whether mutations are predominantly acquired during replication. Genomic coverage of the majority of these HSC colonies ranged from 30-40X. As seen in humans, we found a progressive increase in mutation burden with age within the murine HSC compartment, corresponding to a rate of ~40 SNVs per year. Furthermore, these HSCs had mutational signatures corresponding to that observed in aged human tissues. In contrast to previous reports, data from the label-retention model demonstrated that this age-associated increase in mutation burden correlated with HSC replication, as dormant aged HSCs had similar

mutation burdens to young HSCs. Seemingly contradictory to this finding, we observed no difference in mutation burden upon stress-induced cycling of HSCs. However, it appears we have unintentionally introduced a large selection bias with regards to the agonist-treated clones and future work will focus on rectifying this caveat.

In summary, we have developed a sensitive and specific analysis to accurately detect mutations within individual HSCs. From our results we have clearly demonstrated that mutation acquisition within HSCs accumulates with age and that this increase correlates with an increase in replication history. We envisage that these findings will be an important step towards interrogating whether replication stress is a biologically relevant driver of genome instability in HSCs.

II. Zusammenfassung

Die Aneignung von Mutationen im Genom hämatopoetischer Stammzellen (HSZ) ist von besonderer Bedeutung, da dies wahrscheinlich ein Treiber der malignen Transformation bei vielen Leukämien und ein Markenzeichen des Alterns ist. Wichtig ist, dass die Untersuchung der normalen physiologischen Mediatoren der HSC-Mutationsakquisition ein äußerst wenig untersuchtes Gebiet ist. Das Aufkommen von Sequenzierungstechnologien der nächsten Generation (Next Generation Sequencing, NGS) bietet einen Weg zur Untersuchung dieses Phänomens, wird jedoch durch die Tatsache erschwert, dass eine genomweite Analyse auf klonaler Ebene notwendig ist, um die Mutationssignatur einzelner HSZ zu bestimmen.

Wir haben ein In-vitro-Modell entwickelt, das die empfindliche Untersuchung klonaler Mutationen innerhalb einzelner HSZ unter Verwendung der Sequenzierung des gesamten Genoms ermöglicht. Um unsere klonale Mutationsanalyse zu optimieren, haben wir ein Benchmarking durchgeführt, bei dem wir eine individuelle HSZ-Kolonie bis zu einer ~90-fachen Abdeckung tiefgehend sequenziert und somatische Nukleotid-Variante (SNV)-Analysen bei verschiedenen heruntergesampelten Abdeckungen durchgeführt haben. Wichtig war, dass wir feststellten, dass die Anzahl der entdeckten Mutationen mit zunehmender Abdeckung fast linear zunimmt, bis ein Plateau bei etwa 30-facher Abdeckung erreicht ist. Darüber hinaus entwickelten wir optimale Filterparameter, die im Vergleich zu zuvor veröffentlichten Methoden eine stark verbesserte Fähigkeit zur Erkennung echt positiver und falsch positiver Mutationen bei geringer Abdeckung zeigten.

Mit Hilfe dieser optimierten Sequenzierungspipeline sammelten und sequenzierten wir HSC-Klone von jungen und alten Mäusen sowie von Mäusen, die Stress-Agonisten ausgesetzt waren, von denen bekannt ist, dass sie HSZ-Zyklen induzieren. Zusätzlich setzten wir ein genetisches Markierungs-Retentionssystem ein, um ruhende und aktiv zyklisierende HSZ zu segregieren, und um zu beurteilen, ob Mutationen vorwiegend während der Replikation erworben werden. Die genomische Erfassung der Mehrheit dieser HSZ-Kolonien reichte von einer 30- bis 40-fachen Abdeckung. Wie beim Menschen fanden wir innerhalb des HSZ-Kompartiments der Maus eine

progressive Zunahme der Mutationslast mit zunehmendem Alter, was einer Rate von ~40 SNVs pro Jahr entspricht. Darüber hinaus wiesen diese HSZ Mutationssignaturen auf, die den in gealterten menschlichen Geweben beobachteten Merkmalen entsprechen. Im Gegensatz zu früheren Berichten zeigten die Daten aus dem Label-Retention-Modell, dass diese altersbedingte Zunahme der Mutationslast mit der HSZ-Replikation korrelierte, da schlafende HSZ im Alter ähnliche Mutationslasten wie junge HSZ aufwiesen. Scheinbar widersprüchlich zu diesem Befund beobachteten wir keinen Unterschied in der Mutationslast beim stressbedingten Zyklus der HSZ. Es scheint jedoch, dass wir unbeabsichtigt eine große Selektionsverzerrung in Bezug auf die Agonisten-behandelten Klone eingeführt haben. Die zukünftige Arbeit wird sich darauf konzentrieren, diesen Fehler zu korrigieren.

Zusammenfassend lässt sich sagen, dass wir eine empfindliche und spezifische Analyse entwickelt haben, um Mutationen innerhalb einzelner HSZ genau nachzuweisen. Aus unseren Ergebnissen haben wir eindeutig nachgewiesen, dass sich der Mutationserwerb innerhalb der HSZ mit dem Alter akkumuliert und dass dieser Anstieg mit einer Zunahme der Replikationsgeschichte korreliert. Wir gehen davon aus, dass diese Ergebnisse ein wichtiger Schritt sein werden, um zu untersuchen, ob der Replikationsstress ein biologisch relevanter Treiber der Genominstabilität bei HSZ ist.

III. Introduction

1. Hematopoiesis

The hematopoietic system is highly versatile in its functions; mediating immunity, nutrient transport, hemostasis and wound healing[1]. This requires many different blood cell types working together in a highly regulated and complex system. These include erythrocytes which transport oxygen throughout the body; megakaryocytes which are the cellular precursors to platelets, the cells responsible for blood clotting; and numerous cells which are involved the innate and adaptive immune systems. All of these cells are produced within the hematopoietic system of mammals, which consists of the bone marrow, spleen and liver, during a process called hematopoiesis.

1.1. Hematopoietic hierarchy

Over the last decade, the advancement of next generation sequencing (NGS) technologies has provided researchers with novel methods to investigate hematopoiesis. Assessing the hematopoietic system at single-cell resolution has revealed a heterogeneity in blood cell compartments which were previously thought to be phenotypically and functionally homogeneous. As a result, the classical model of the hematopoietic hierarchy has evolved[1].

Briefly, the classic model of hematopoiesis shows cells of equal potential within discrete groups and hematopoietic stem cells (HSCs) at the top of the hierarchy (Figure 1.A.). This model relies on the dogma that HSCs possess a self-renewal capacity and can give rise to all blood cell lineages. As we "move down" the hierarchy, cells become increasingly more lineage-restricted until eventually producing mature blood cells[2]. This model assumes that the HSCs at the top of the hierarchy, known as long-term (LT) HSCs have the highest potency and remain quiescent for the majority of their lifetime, rather relying on more differentiated short-term (ST) HSCs and progenitors to proliferate for the maintenance of the hematopoietic system.

Conversely, single-cell RNA sequencing (scRNA-seq) has highlighted the heterogeneity in blood cells previously thought to be of uniform potential[3–5]. These results have rather argued for a continuum model of the hematopoietic system, instead of distinct punctuated fate decisions as cells differentiate (Figure 1.B.)[1]. From scRNA-seq data, researchers observed a continuity in the distribution of the different blood cell types, from HSCs to more committed cells[3,5]. As such, this continuum model depicts a constant differentiation process from HSCs to progenitors to more committed mature blood cell types. However, some researchers argued that although the continuum model tackles the oversimplified view of the classical model, using gene expression data alone may not be sufficient for understanding the hematopoietic system hierarchy[6]. In fact, key findings have demonstrated that there are distinct punctuated changes which occur across this continuous gene expression landscape, resulting in distinct cell groups[7]. Therefore, the continuum model has been redefined to include these distinct changes and has been appointed as the punctuated continuum model (Figure 1.C.).

1.2. The role of HSCs in hematopoiesis

At the top of all of the hematopoietic system hierarchy models are HSCs. HSCs are characterized by their self-renewal capacity and are responsible for the lifelong replenishment of blood cells[2]. In general, the majority of HSCs remain dormant throughout an individuals' life and only exit out of dormancy when required for blood production and maintenance. However, it has been observed that during the lifetime of an individual these cells are forced into cycle multiple times, leading to their gradual depletion. This time-associated depletion in HSC function is known as ageing and can result in the impaired regeneration of the blood compartment.

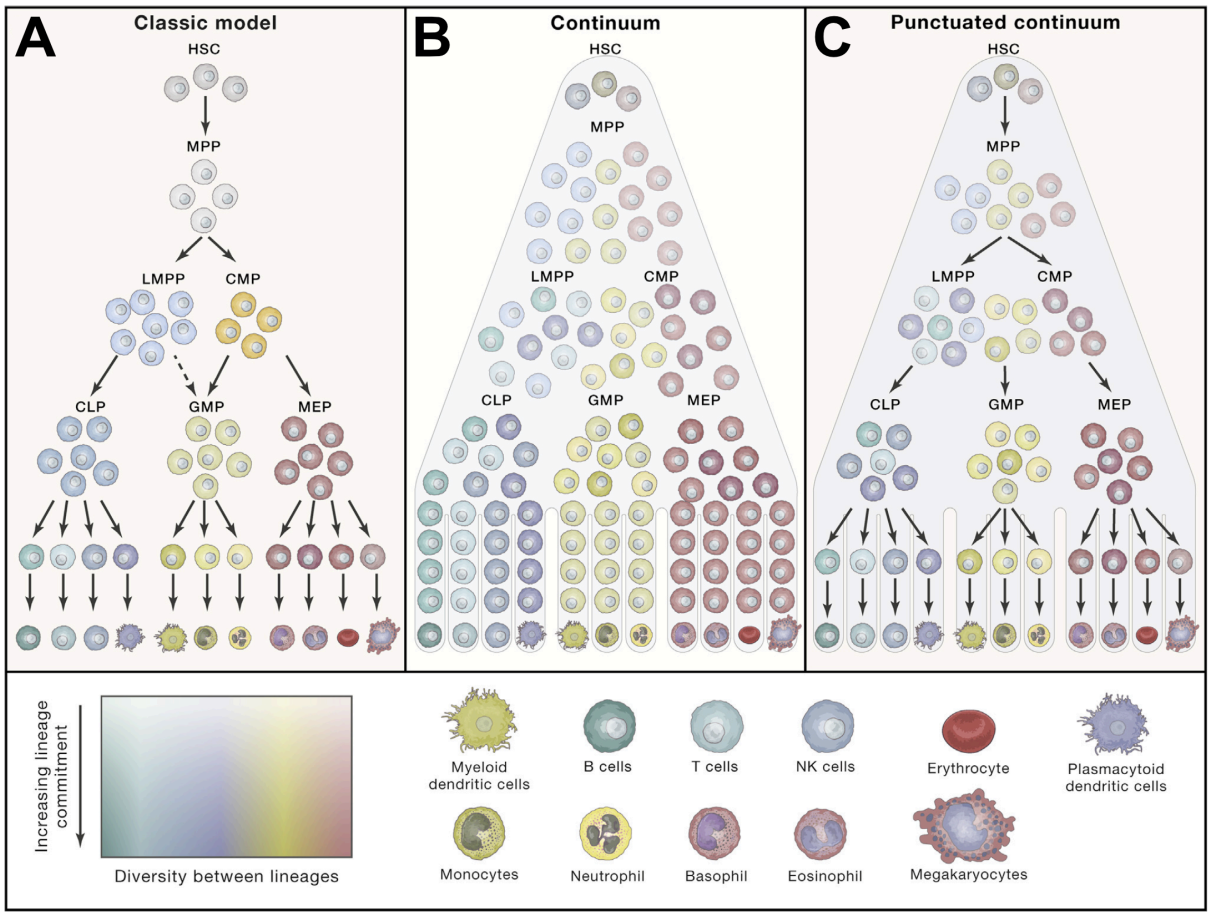


Figure 1. The evolving models of the mammalian hematopoietic hierarchy. (A) The classic model of hematopoiesis. **(B)** The continuum model of hematopoiesis. **(C)** The punctuated model of hematopoiesis. Mature blood cell types are depicted in the legend at the bottom of the figure. Different lineages are depicted in different colors and lineage commitment is depicted as a decrease in transparency (see legend). HSC = hematopoietic stem cell; MPP = multipotent progenitor; LMPP = lymphoid-primed multipotent progenitor; CMP = common myeloid progenitor; CLP = common lymphoid progenitor; GMP = granulocyte/macrophage progenitor; MEP = megakaryocytic/erythroid progenitor. (Figure from Liggett *et al.*, 2020, reprinted with permission from Elsevier).

2. Ageing

2.1. Hallmarks of ageing

Ageing is characterized by the deterioration in normal cells and tissue function over time, resulting in an increased risk for the development of various pathologies including cancer, cardiovascular diseases and many others. With regards to the mammalian system, factors which

affect ageing have been broadly classified into nine hallmarks[8]. These are: genomic instability, telomere attrition, mitochondrial dysfunction, stem cell exhaustion, cellular senescence, epigenetic alterations, loss of proteostasis, deregulated nutrient-sensing and altered intercellular communication (Figure 2). Of importance to this study is the genomic instability and types of DNA damage which have been observed with ageing.

Figure 2. The Hallmarks of Ageing. Diagram from López-Otín *et al.* (2013) illustrating the nine hallmarks of ageing (reprinted with permission from Elsevier).

2.2. Genomic instability and the types of DNA damage viewed with age

It has long been hypothesized that the age-associated accumulation of DNA damage is a probable mechanism of stem cell ageing and attrition[8,9]. Likewise, several advanced ageing, progeroid disorders have been linked to an increased accumulation of DNA damage due to defective or deficient DNA repair mechanisms[10]. There are various exogenous and endogenous threats which

affect the stability and integrity of DNA and can result in numerous types of genetic lesions[8]. Such mutations include single nucleotide variants (SNVs), insertions and deletions (INDELs), and translocations, as well as telomere shortening and damage to mitochondrial DNA. To repair such lesions, organisms contain highly complex and specific DNA repair mechanisms, as well as mechanisms to ensure the integrity of telomeres and mitochondrial DNA. Even so, certain mutations persist and accumulate within the genomes of organisms, some of which have been associated with ageing.

Accumulation of somatic mutations

It has been observed that somatic mutations accumulate within the genomes of organisms with age[8]. These variants include an increase SNVs[11], as well as the presence of copy-number variations[12]. Such genomic DNA alterations could disrupt transcriptional pathways and gene-coding regions, leading to the production of mutated cells with altered cellular functions. If these mutated cells are not eliminated, homeostasis can be disrupted. This is particularly relevant to mutations present within adult stem cells (ASCs). Age-associated mutation acquisition within the genomes of ASCs likely has a large impact on tissue health as these variants are passed on to progeny cells during stem cell self-renewal and replication[13]. Furthermore, it has been hypothesized that the accumulation of mutations within ASCs underlies the initiation of age-related diseases like organ failure and cancer[13–15].

Mitochondrial DNA aberrations

Similarly to SNVs, an increase in mutations and deletions within mitochondrial DNA (mtDNA) has been shown to contribute to ageing and age-associated diseases[16,17]. This was assumed to be caused by oxidative damage considering the oxidative environment of mitochondria. However studies revealed that most mutations present within mtDNA were caused by replication errors and variant accumulation is likely a result of the inefficient mtDNA repair mechanisms. Conversely, an age-associated reduction in mtDNA-copy number (CN) has been observed and correlates with the general health of aged individuals[17,18] Together, aberrations within the

mtDNA genomes of organisms have a prominent effect on ageing and the development of age-associated diseases.

Telomere shortening

A sequential and cumulative shortening of telomeric sequences has been observed in mammalian organisms with age[19]. Such shortening occurs during cellular division when telomeres are replicated[8], a process mediated by a specialized DNA polymerase called telomerase. Unfortunately, the majority of mammalian somatic cells are deficient of telomerase and so telomeric DNA is progressively lost during cellular replication and ageing of these cells. It has been observed that various exogenous and physiological factors can also affect telomere length and that shorter telomeres have been associated with a decrease in health and survival, and an increase in disease incidence. Furthermore, the rate at which telomeres shorten can be used as a potential indicator of the rate of individual ageing.

2.3. DNA damage and repair mechanisms

In order to understand how mutations arise and accumulate with age, it is not only important to understand what types of DNA damage there are but more so, their corresponding repair mechanisms and downstream effects. When the DNA of a cell has been damaged, whether due to endogenous or exogenous mediators, a DNA damage response (DDR) signaling pathway is initiated. DDR effectors are subsequently activated and carry out the appropriate response to this damage, either initiating senescence, DNA repair or initiating apoptosis[20]. There are many types of DNA damage repair pathways including, but not limited to: base excision repair (BER), homologous recombination (HR), nucleotide excision repair (NER), non-homologous end joining (NHEJ), mismatch repair (MMR) and translesion synthesis (TLS)[21], and the deregulation of each of these results in specific mutation types and signatures. Depending on the cell type, responses to DNA damage will be different. This is true for HSCs, where it has been observed that following certain types of DNA damage, these cells preferentially undergo apoptosis rather than potentially performing defective DNA damage repair[22]. This suggests that HSCs contain some form of a protective mechanism against the acquisition mutations following DNA damage. That being said,

various studies have still observed an increase in mutation burden with age within HSCs[11,23,24]. Considering that with age the efficiency of DDR and repair pathways decreases, this is the likely explanation for this observed age-associated increase in genome instability and mutation acquisition[8].

Understanding the mechanisms which impact on the age-associated accumulation of DNA damage is of great importance in order to prevent the incidence of age-related diseases and malignancies. The advancement of NGS technologies has allowed researchers to delve deeper into the genetic and epigenetic changes which occur with age. This has highlighted that, to an extent, the rate of ageing is controlled by DNA damage, as well as other biochemical processes[8]. As such, repair and prevention of DNA damage could potentially decrease disease incidence and slow the rate of ageing in individuals. Therefore, ageing research needs to be extended to focus on the physiologic sources and possible mechanisms involved in the accumulation of DNA damage with age.

3. DNA damage and ageing of the hematopoietic system

Regarding the hematopoietic system, issues which arise during ageing broadly fall into two seemingly contradictory categories: the loss of cellular output and the massive increase in certain cells. Stem cell exhaustion results in decreased cell cycle activity[9] and the subsequent failure of HSCs to produce daughter cells, leading to a decrease in certain blood cell types[8]. This is the case for hematologic phenotypes like anemia and myeloid malignancies. Conversely, an excessive increase in the cellular output of certain blood cells is another detrimental age-associated phenomenon which has been observed. This occurs when a cell acquires a gain of function mutation resulting in aberrant cellular function and replication, which is the case with certain leukemias. Both of these phenomena have been linked to the acquisition and accumulation of DNA damage with age.

3.1. Age-associated hematologic phenotypes caused by mutation acquisition

As people age, an increase in clonal expansion of mutated HSCs is commonly viewed[11,25,26] (Figure 3.A.). This clonal expansion is known as clonal hematopoiesis and has been associated with an increased risk of hematologic malignancy development. Although the mechanistic processes of how this occurs still remains largely unclear, it has been hypothesized that certain aberrant mutations could promote the development of this clonal expansion[14,25,27], and other age-associated hematologic malignancies. This theory has been strengthened by the identification of pre-leukemic HSCs which exhibit aberrant mutations in the *DNMT3a* gene[28], as well as other leukemic driver mutations[29]. Furthermore, clonal hematopoiesis of indeterminate potential (CHIP), an age-associated phenotype of clonal hematopoiesis in the absence of hematologic neoplasia, has been characterized by the presence of certain somatic mutations which drive clonal expansion and accumulate with age during normal hematopoieisis[30,31].

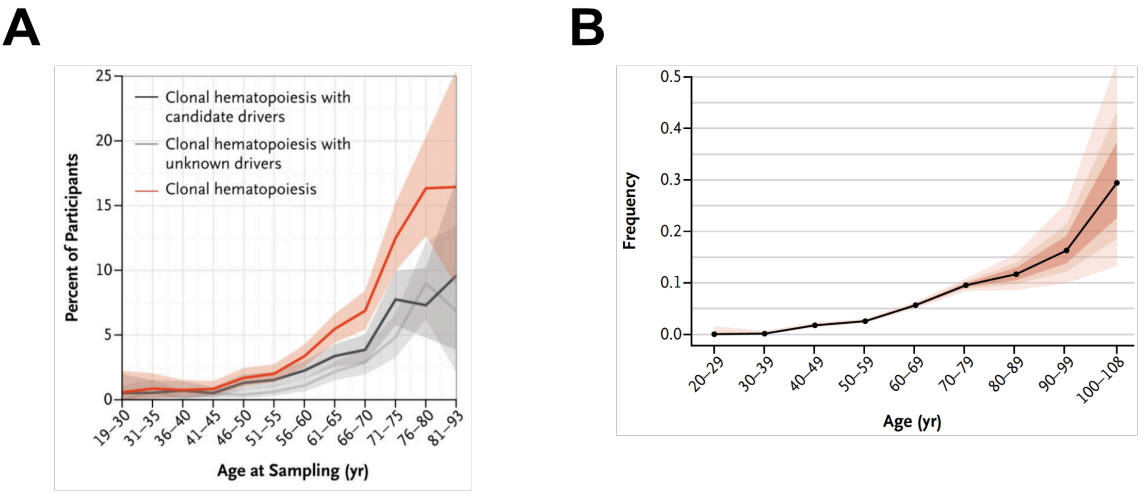


Figure 3. Increase in somatic mutation burden and clonal hematopoiesis in peripheral blood from humans with age. (A) Prevalence of clonal hematopoiesis identified using whole-exome sequencing of DNA from peripheral-blood cells in donors of various ages. Colored bands represent 95% confidence intervals. Yr = year. (Reproduced with permission from Genovese *et al.*, 2014, Copyright Massachusetts Medical Society). **(B)** Prevalence of somatic mutations with age identified using whole-exome sequencing of DNA from peripheral-blood cells. Colored bands represent the 50th, 75th and 95th percentiles. Yr = year. (Reproduced with permission from Jaiswal *et al.*, 2014, Copyright Massachusetts Medical Society).

3.2. Increase in somatic single nucleotide variant burden with age

Early research of DNA damage within HSCs largely focused on the characterization of the different types of DNA damage, however there was a large gap in knowledge regarding the

downstream consequences of this damage. An early study using lacZ-plasmid transgenic mice (small blue mouse model) tried to quantify the type of mutations present within HSCs as a result of DNA damage[22,32]. However until the advent of next generation sequencing (NGS), assessing mutation burden of HSCs using native DNA was not possible.

As previously mentioned, there is a gradual accumulation of genetic damage within various tissues, as well as ASCs with age[33]. Likewise, this phenomenon has also been observed within the hematopoietic system of humans. When looking at whole-exome sequencing data from peripheral-blood from various aged donors with no record of cancer or hematologic malignancies, researchers observed a steady increase in somatic mutation burden with age[11] (Figures 3.B.). Furthermore, it was later observed that SNVs accumulate within hematopoietic stem and progenitor cells (HSPCs) in a linear fashion with age and that this increase occurs gradually during an individual's lifetime[23,24], at approximately 14 SNVs per year (Figure 4.A. and B.). This is much lower than the mutation rates viewed from more active ASCs where mutation accumulation occurs steadily at a rate of approximately 40 mutations per year[33], highlighting the predominantly quiescent nature of HSCs.

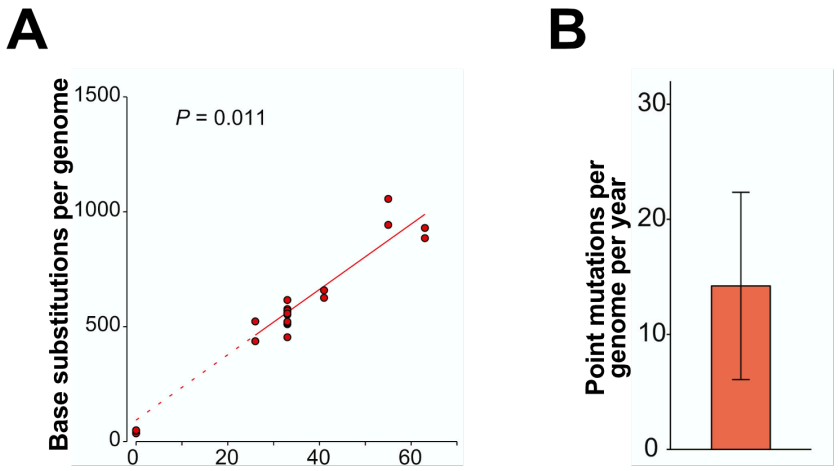


Figure 4. Increase in somatic SNV burden in human HSPC colonies. (A) Correlation of the number of SNVs within HSPC colonies derived from donors of different ages. **(B)** Annual SNV rate estimated from **(A)**. (Figure reprinted with permission from Osorio *et al.*, 2018 [DOI: 10.1016/j.celrep.2018.11.014] through the Creative Commons user license: https://creativecommons.org/licenses/by-nc-nd/4.0/).

In order to understand the mechanisms which drive the accumulation of mutations in and clonal expansion of HSCs under normal hematopoiesis, a retrospective analysis of DNA damage and mutation accumulation at various ages is required. Currently, this is not feasible in human studies as sample acquisition is an invasive process and is only performed when absolutely necessary. Furthermore, human bone marrow aspirations are usually performed for health reasons and it is not often that healthy samples, with no pre-existing hematologic phenotype, are acquired. The murine model provides us with an attractive alternative to assess the types of DNA damage which occur within HSCs during ageing and the physiological mechanisms which drive this.

4. Mediators of DNA damage and ageing

As previously mentioned, the quiescent nature of HSCs results in a much lower mutation burden compared to more active ASCs, such as those found in the small intestine, colon and liver[33]. Even so, mutation accumulation within the genomes of HSCs can contribute to the dysfunction of normal hematopoiesis[13,23]. However the underlying physiological sources and mechanisms of DNA damage and ageing of HSCs still remains largely unclear.

Previous studies have studied DNA damage in HSCs exposed to ionizing radiation and alkylating agents[22], however this is irrelevant to DNA damage acquired during normal ageing. To date, the physiological sources which mediate DNA damage in HSCs during ageing remain largely unclear.

4.1. Stem cell replication

Although reduced HSC proliferation has been noted as a harmful side-effect of ageing, excessive replication can result in the decreased potency and exhaustion of HSC niches[8]. As such, replication stress has been noted as a potential source of DNA damage within HSCs[34]. Replication stress results in the stalling of the replication fork during DNA replication, leaving the genome exposed to stressors, and can result in DNA strand breaks[35]. This of course highlights some issues, as replication stress is understandably a concern for actively proliferating cells, which is not the case for HSCs as they are believed to remain predominantly quiescent throughout their lifetime.

That being said, many studies have observed an increase in DNA damage foci, namely γH2AX, within old murine HSCs compared to young[9,34,36] (Figure 5). While some believe that this increase in DNA damage is due to the impairment of DNA damage repair mechanisms with age[9], others have argued that such DNA damage foci are actually representing replication stress[34]. Further complicating the matter, another study argued that DNA damage accumulates in HSCs because repair mechanisms are reduced during their normal quiescent state, and that HSC replication results in this DNA damage being repaired due to the reactivation of DNA damage repair mechanisms[36]. Of course, these studies assessed DNA damage accumulation in HSCs by only using DNA damage response foci. Since then, sequencing studies have shown that mutation acquisition increases with age so the idea that HSC replication repairs DNA damage does not seem feasible. Furthermore, using a label-retaining mouse model, researchers showed that age-related phenotypic changes of HSCs was dependent on their divisional history[37]. That is, HSCs which cycled more were less functionally potent and lost their self-renewal capacity, while those which remained dormant throughout their lifetime closer resembled those with LT-HSC capacity.

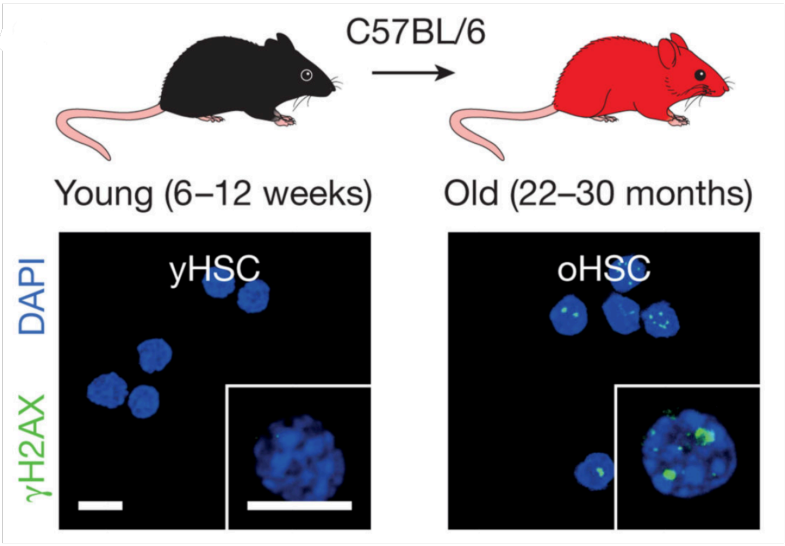


Figure 5. Accumulation of γH2AX foci in old murine HSCs. (Figure from Flach *et al.*, 2014, reprinted with permission from Springer Nature).

To summarize, it has been shown that there is an increase in DNA damage foci within older murine HSCs compared to young, and there is an age-associated increase in mutation burden

within aged human HSCs. Furthermore, HSCs which have replicated more have exhibited decreased functionality and self-renewal potential; phenotypes which have also been observed in aged HSCs. Taken together, it could be hypothesized that this age-associated increase in DNA damage results in an increase in mutation burden within HSCs and that this could be a result of replication history. However, it still remains unclear whether this age-associated increase in DNA damage foci results in an increase in mutation burden within HSCs.

4.2. Stress-induced replication

Besides the natural cycling of HSCs for normal hematopoiesis, stress-induced hematopoiesis by factors like infection, blood loss, and endogenous and exogenous toxins can occur multiple times throughout an individuals' lifetime[38]. Exposure to these stresses promotes the exit of HSCs from quiescence, into proliferation[38–40]. It has been shown that stress-induced cycling severely affects the functionality of HSCs, leading to their diminished self-renewal and proliferation potential, functional exhaustion and an increase in apoptosis[9]. As such, chronic exposure to such stress factors can result in hematopoietic system failure[38–41].

Numerous studies have demonstrated that HSCs exposed to infection exhibit an inflammatory response which subsequently forces them into cycle. Separate studies where mice were treated with pIpC[38], an agonist which is structurally similar to and mimics a virus infection, and a *Mycobacterium avium* infection[39] both showed an activation of inflammatory response via interferon-alpha (IFNα) and interferon-gamma (IFNγ), respectively. In addition, activating dormant HSCs via IFNα treatment rendered these newly cycling HSCs sensitive to the chemotherapeutic agent 5-fluoro-uracil (5FU), contrary to their previously 5FU-resistant dormant state[38]. Furthermore, chronic exposure to pIpC and IFNα resulted in an increase in DNA damage foci within HSCs, and loss of functionality[40,41]. Taken together these observations highlighted the fact that replicating HSCs are more sensitive to chemotherapeutics and DNA damage, and dormancy seems to protect HSCs from damage.

This stress-induced HSC activation and subsequent increase in DNA damage foci was observed for various physiological challenges (Figure 6. A. and B., respectively)[40]. As previously mentioned, the increase in DNA damage foci does not necessarily represent an increase in mutation burden within HSCs. However, a recent study using whole genome sequencing (WGS) demonstrated that in the absence of functional DNA damage repair mechanisms, stress-induced DNA damage via treatment with reactive aldehydes resulted in an increase in mutation burden within these HSCs[42]. Although this highlights a different route of potential DNA damage accumulation within HSCs, it goes to show that stress-induced DNA damage, whatever the cause, can result in an increase in mutations within HSCs, and that NGS technologies provide an accurate method for interrogating this.

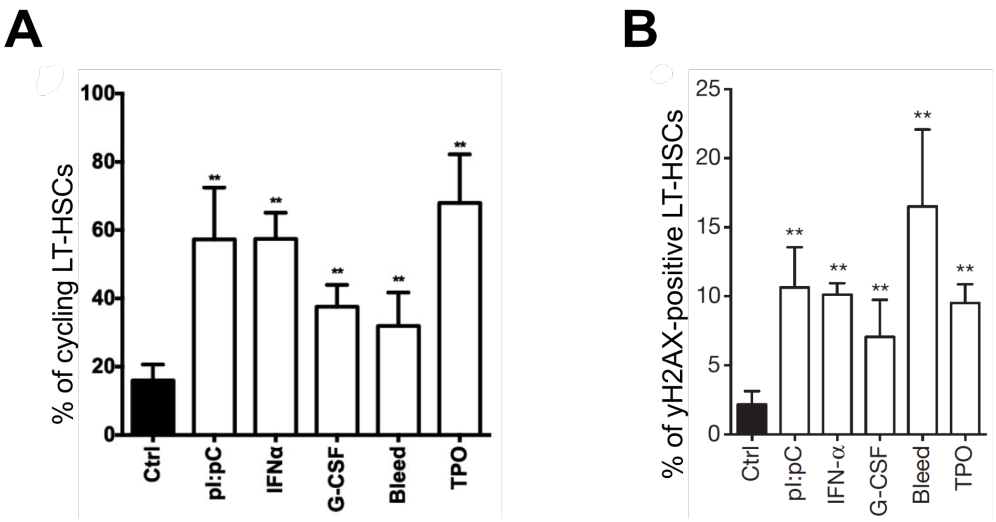


Figure 6. Stress-induced replication and DNA damage within murine HSCs. (A) Treatment with different agonists results in the exit of LT-HSCs from dormancy. **(B)** Stress-induced cycling of HSCs results in the increase in DNA damage response γH2AX foci within LT-HSCs. Percentage of cycling and γH2AX-positive HSCs after treatment is represented in graphs A. and B., respectively. Error bars indicate SD. * indicates a statistical difference between treatment groups compared to control (ctrl) group (unpaired two-sided t test). (Figures A. and B. from Walter *et al.*, 2015, reprinted with permission from Springer Nature).

5. Mutational processes associated with ageing

Besides the presence of CHIP- and leukemic driver mutations, and the overall increase in mutation burden within blood cells, the presence of certain mutation signatures have also been correlated to ageing in humans[43,44]. Mutational signatures take into account the type of DNA damage and subsequent DNA damage repair capacity of a particular cell. Thus allowing researchers the opportunity to identify which mutational processes were present or absent within their samples.

Mutational signatures are characterized by combinations of mutations. Initially these mutation signatures were identified using SNV substitutions and classed according to their base change, resulting in 96 different trinucleotide base change combination categories[43]. More recently, mutation signatures have also been identified using INDEL calls from WGS data of large human sequencing cancer cohorts[44]. Currently the updated list of human mutational signatures which have been identified encompass numerous signatures based on SNVs, doublet base substitutions (DBS) and INDELs (https://cancer.sanger.ac.uk/cosmic/signatures).

Of particular interest to this study are the signatures which have been correlated with ageing, signatures 1 and 5. The mutation profile of signature 1 is dominated by C>T transitions, in particular ACG, CCG, GCG and TCG, and is a result of enzymatic or spontaneous deamination of 5-methylcytosine to thymine. While signature 5 mutation profile is dominated by an increase in C>T and T>C transitions, and has an unknown cause. These signatures were initially observed to be correlated to the age of patients from tumor sequencing cohorts[43], but were also later observed in non-cancerous samples derived from HSPCs from various aged humans[24]. Specifically, both of these signatures are reported to act in a "clock-like" manner and the number of SNVs attributed to these signatures increased in a linear fashion with age within HSPCs (Figure 7.A.) Furthermore, when simply looking at the 96-trinucleotide mutation profile from blood colonies of various aged donors with no history of hematologic malignancies, it was observed that the mutation spectra was dominated by C>T and T>C transitions[24,45] (Figure 7.B.; top mutation profile). This mutation spectra is similar to that observed in patients with myeloid

cancer[23,45] (Figure 7.B.; bottom mutation profile) and those identified with age-related clonal hematopoiesis[27].

Although it is possible to determine mutational signatures using whole-exome and targeted sequencing, these methods are biased towards gene-coding regions[46]. This could result in an overview which is not representative of the mutational processes occurring over the whole genome. Therefore, in order to assess mutational signatures which are representative of the whole genome of a cell, accurate WGS analysis is required[46].

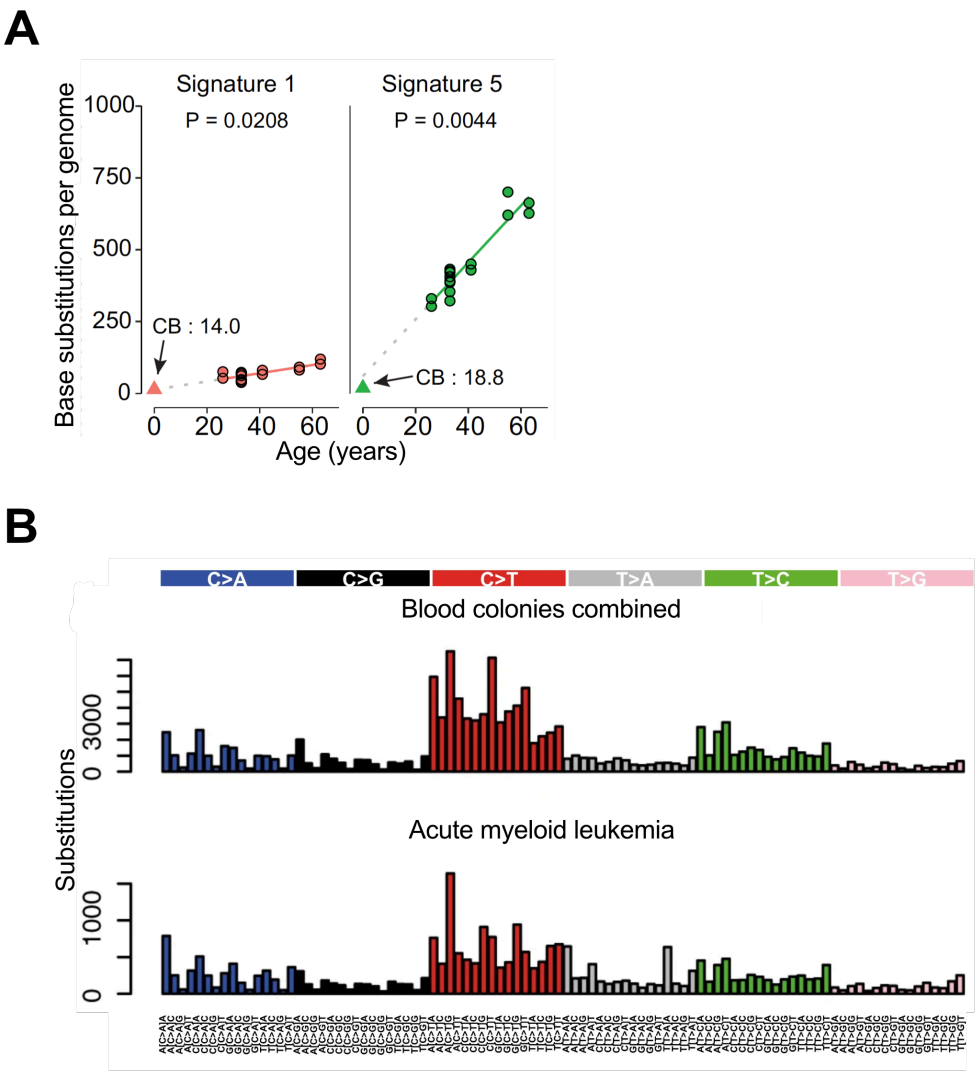


Figure 7. Mutation profile and signatures identified in human HSPC colonies. (A) Contribution of SNV mutation signatures 1 and 5 in the genomes of HSPCs, plotted against the age of donors. P values indicate a significant increase

in these signatures with age (two-tailed t test). (Figure reprinted with permission from Osorio *et al.*, 2018 [DOI: 10.1016/j.celrep.2018.11.014] through the Creative Commons user license: https://creativecommons.org/licenses/by-nc-nd/4.0/). **(B)** The trinucleotide context of SNVs present within multiple colonies derived from HSPCs from a 59-year old male, compared to those identified in pooled acute myeloid leukemia (AML) genomes. Any AML genomes with >1500 mutations were excluded. (Figure from Lee-Six *et al.*, 2018, reprinted with permission from Springer Nature).

6. Whole genome sequencing of single HSC-derived colonies as a model to study DNA damage and ageing

In order to understand how ageing affects the functionality of stem cells, it is important to understand the mutational processes within these cells[46]. Single-cell derived colonies provide us with an excellent method to assess the mutation burden within individual HSCs. By assessing the quantitative and qualitative mutations within stem cells derived from donors of different ages, researchers are able to evaluate the relationship of mutation accumulation with age and subsequent stem cell attrition[33]. Furthermore, this technique provides us with a unique opportunity to investigate the roles that normal and stress-induced replication have on normal and aberrant mutational processes, which could ultimately lead to the development of age-related malignancies and cancer.

6.1. Rationale for using WGS of *in* vitro expanded colonies to assess DNA damage within individual HSCs

Due to the polyclonal nature of a healthy hematopoietic system, each somatic mutation which occurs within an HSC is only present within a small subset of cells within the entire blood cell population. Therefore, it is not possible to detect these somatic events at the stem cell level using bulk sequencing technologies[46]. As such, the mutations of single HSCs must be assessed in order to determine which mutational processes are at play during HSC ageing. Previous studies have shown that the mutation burden within ASCs is relatively low and variants are distributed across the entire genome[24,33,46,47]. Hence, it is important to utilize a sequencing method which can accurately sequence the whole genomes of single cells without introducing many sequencing errors, which could potentially be falsely called as true somatic events. Therefore single-cell

sequencing methods, although promising, are currently not suitable for such an analysis as the necessity for whole-genome amplification introduces too many sequencing errors, making it difficult to differentiate these errors from the low number of true somatic events within the stem cell[48]. Additionally, ultra-deep sequencing of tissues which naturally display a degree of clonality has successfully been used to detect somatic mutations within clone-initiating cells[25,26]. However, extremely deep sequencing is required (~500X) and this method relies on the assumption that the sample in question contains clonal cell populations.

6.2. Model for the identification of somatic mutations at the stem cell level in single-cell derived colonies

In order to assess DNA damage across the whole genome of an individual stem cell, it is vitally important that you obtain sufficient DNA for the accurate assessment of variants using WGS. Previous human studies have utilized *in vitro* expansion techniques in order to overcome this obstacle [24,45,46]. In short, the necessary cells are isolated from a patient, whether from tissue biopsies or sorted stem cells, and subsequently cultured in suitable expansion medium and conditions (Figure 8). Depending on the cell type, cells are allowed to expand in culture for an appropriate amount of time after which DNA is isolated from these cell colonies. WGS libraries are prepared for each cell colony and sent for sequencing, along with an appropriate control. Whether using organoid or cell colony cultures, the principle behind cataloguing mutations at the individual stem cell level is similar. Simply put, if a heterozygous mutation is present within the original stem cell, it will be present within all progeny cells and thus, will be observed at a variant allele frequency (VAF) of 0.5 (Figure 9.A.). Any variants which have been acquired during *in vitro* expansion will be viewed at lower VAFs and can be filtered out using a suitable VAF cut-off (Figure 9.B.). As such, somatic variants within the original cell can only be identified if the sample is purely clonal[46].

The issue that arises with this model is when there is a selection for the outgrowth of a daughter cell, resulting in an increased VAF for *in vitro* expanded mutations. Coupled with the possible

presence of sequencing artefacts and repetitive/homologous regions within the genome, it is vitally important to develop strict filtering parameters which can control for such phenomena.

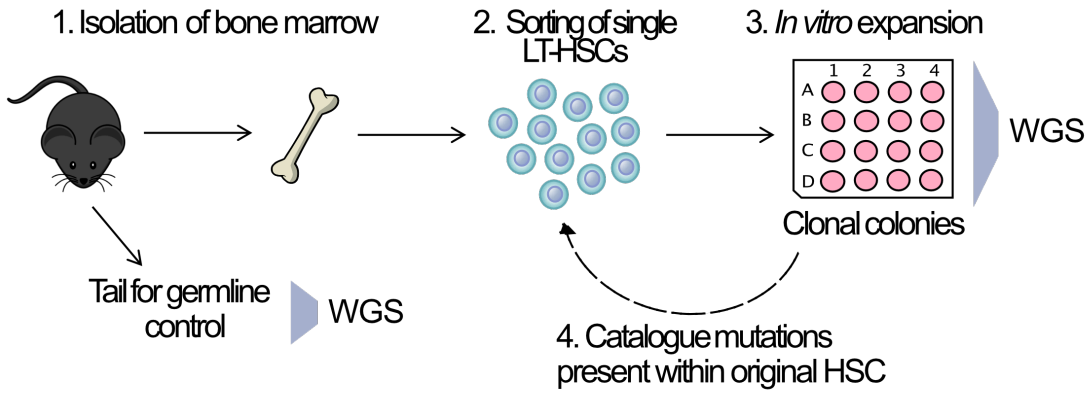


Figure 8. Clonal expansion of single HSCs to determine somatic mutation burden. Schematic overview of experimental pipeline to identify mutations within single murine HSC colonies. LT-HSCs = long-term hematopoietic stem cells; WGS = whole-genome sequencing. (Figure adapted with permission from Osorio *et al.*, 2018 [DOI: 10.1016/j.celrep.2018.11.014] through the Creative Commons user license: https://creativecommons.org/licenses/by-nc-nd/4.0/)

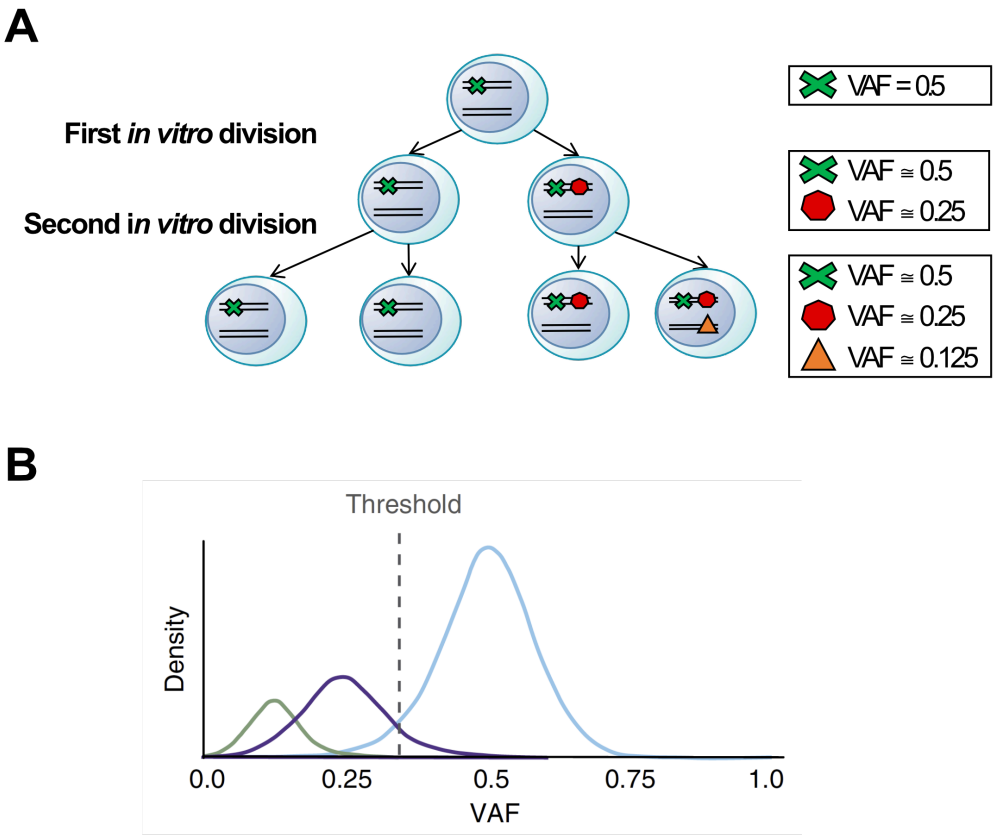


Figure 9. Model of variant allele frequency (VAF) distribution of heterozygous variants within single-cell derived cultures. **(A)** *In vitro* expansion of a single-cell derived colony. Heterozygous variants present within the original cell will be passed on to all progeny cells and will therefore have a VAF of 0.5 (green variant). Any variants which are acquired during *in vitro* expansion will be viewed at a lower VAF and can be filtered out (red and orange variants). **(B)** Theoretical VAF density plot of scenario depicted in **(A)**. A VAF cutoff of 0.3 is applied to filter out variants acquired *in vitro* (Figure from Jager *et al.*, 2017, reprinted with permission from Springer Nature).

This approach has successfully been used to assess the mutation burden within ASCs derived from liver, small intestine and colon samples from humans[33,46], as well as cells from the stomach, small intestine, colon and prostate from mice[47]. Furthermore, this single-cell derived colony approach has been utilized to determine the mutation burden within single hematopoietic stem and progenitor cells (HSPCs) from one 59-year old male[45], as well as various aged humans[24]. Similar to what was seen with bulk sequencing of human blood samples[11,23], there is a linear increase in mutation accumulation within the hematopoietic compartment with age[24] (Figure 4.A.).

Although various studies have used such a clonal culture method to identify somatic variants within the original stem cell using WGS, the majority of these studies have not performed an extensive analysis optimization prior to sequencing of their samples [24,33,45,46]. As such, there is a large variability in the sequencing techniques, coverage, variant analysis pipelines, filtering parameters and downstream analysis which has been used.

6.3. Benchmarking for WGS pipeline and analysis optimization

With the ever-increasing technical developments, WGS has become an affordable tool to assess variants across the whole genome of various organisms. That being said, little has been done to standardize the downstream analysis of WGS data. Although it is not feasible to broadly standardize WGS data analysis for all types of variant calling, it is important that individual studies optimize their analysis in order to accurately answer their question at hand.

It has been well-documented that prior polymerase chain reaction (PCR) amplification, sequencing library preparation and even the sequencing itself can introduce artefacts within sequencing data[49]. However, as new technologies are being developed, classification of associated sequencing artefacts is lagging behind. Such artefacts could have profound effects on the ability to accurately call true somatic mutations within a sample. Depending on the scope of a research project, it may not always be possible to optimize methods to reduce or eliminate these artefacts. However, it is possible to tailor one's downstream analysis to control for such artefacts and filter these out.

Unfortunately, the majority of studies using WGS have performed little to no analysis optimization, including several of the clonal culture studies mentioned before. A benchmarking study performed using high-coverage WGS data from cancer samples, highlighted the major discrepancies in WGS data analysis[50]. They emphasized the need for high coverage benchmarking datasets for determining the optimal sequencing coverage and downstream analysis parameters. Additionally they advise researchers to keep the coverage of their sample

and germline control within a close range (± 10%). This appears to greatly affect downstream analysis as systematic artefacts can be balanced out in the control sample and subsequently filtered out. Furthermore, they found little to no concordance among mutation calling pipelines which highlighted the lack of standards available for dealing with sequencing artefacts. As such, this benchmarking study clearly demonstrates the need for prior sequencing and analysis pipeline optimization in order to generate consistent, high-quality WGS data from which results can be accurately interpreted.

7. Aim of thesis

The hypothesis that we wished to address in this study was that age-associated mutation acquisition within HSCs is, in part, mediated by replication stress and stress-induced cycling of HSCs results in an advanced-ageing phenotype with regards to mutation accumulation. In order to interrogate this hypothesis, we performed whole genome sequencing of single HSC-derived colonies to address the following aims:

Aim 1: Perform an extensive WGS benchmarking analysis prior to sequencing HSCs from our various experimental groups. Using a high coverage benchmarking dataset, and matched coverage germline control, we aimed to identify the optimal parameters of our sequencing and analysis pipelines, and assess the effect of various confounding factors. Such factors include the differences in variant calling pipelines, sequence coverage, and DNA input, as well as sequence filtering parameters.

Aim 2: Perform a quantitative and qualitative analysis to assess the effect of ageing on murine HSC genome stability. Here we aimed to interrogate the differences in mutation burden between young and old HSCs, and additionally assess the differences in types of mutations present, mutation signatures, as well as interrogate differences in mtDNA aberrations and telomere length.

Aim 3: Assess the effect of possible mediators of DNA damage on HSC mutation acquisition; including replication and stress-induced activation of HSCs. Again, we aimed to assess the differences in the types of mutations present, mutation burden, mutational signatures, mtDNA aberrations and telomere length between the different study groups.

IV. Results

1. Optimization of whole genome sequencing and analysis parameters

To assess the mutation burden within single HSCs using whole genome sequencing (WGS), we utilized an *in vitro* culture method in order to obtain sufficient DNA for downstream WGS library preparation. In short, we sorted single long-term hematopoietic stem cells (LT-HSCs)[40] from our experimental mice into separate wells of a 96-well plate and allowed these to expand in appropriate culturing medium and conditions for ~2 weeks (Figure 10). DNA from these HSC colonies was later isolated and sequencing libraries were prepared, along with libraries generated from matched control tail DNA, and sent for sequencing (DNA isolation, library preparation and sequencing details can be found in Methods section 3 and 4).

Prior to the sequencing of HSC colonies from our different experimental groups, we sequenced one HSC colony, and matched germline control, to a high depth (~90X) and used this as a benchmarking dataset to optimize our sequencing and analysis parameters. Utilizing this dataset we performed an extensive comparative analysis of two single nucleotide variant (SNV) pipelines, MuTect[51] and CaVEMan[52], which have both been previously used in WGS studies of clonal samples. This benchmarking study not only allowed us to determine which SNV pipeline and filtering parameters would be optimal for our downstream analysis, but also to identify what sequencing depth was required for the accurate assessment of somatic SNVs within single HSCs.

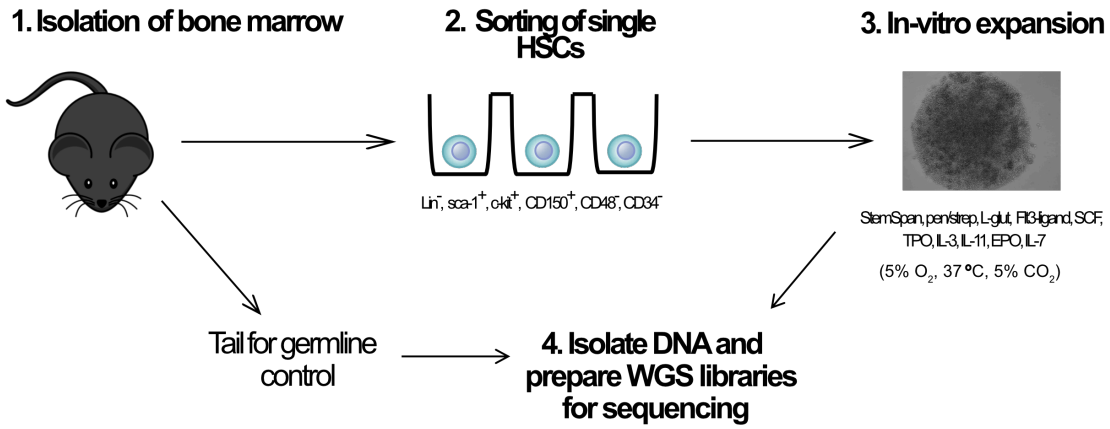


Figure 10. Schematic of experimental overview for WGS of single HSC colonies. HSCs were sorted for the antibodies shown in step 2. HSC expansion medium and incubation conditions shown in step 3. HSCs = hematopoietic stem cells; WGS = whole-genome sequencing; Lin^- = lineage negative; pen/strep = Penicillin-Streptomycin; L-glut = L Glutamine; EPO = Erythropoietin; TPO = Thrombopoietin; IL = interleukin; O_2 = oxygen; CO_2 = carbon dioxide.

1.1. Benchmarking analysis reveals CaVEMan outperforms MuTect for the detection of somatic SNVs within HSC colonies

1.1.1. Filtering parameters for calling SNVs

Prior to the sequencing of multiple murine HSC colonies, it was of high importance that we first perform a thorough benchmarking analysis in order to determine the optimal sequencing analysis and filtering parameters, as well as genome coverage for the accurate calling of somatic variants. Initial single nucleotide variant (SNV) analysis of a deeply sequenced HSC colony (~90X for both HSC colony and matched germline control) showed a distinct peak of high frequency, low variant allele frequency (VAF) mutations and an almost indiscernible peak at ~0.5 VAF (Figure 11. A. and C.). This was an unexpected phenomena based on the clonal nature of our samples, where we rather expected a higher frequency of somatic variants at ~0.5 VAF and a more binomial distribution. This presence of high frequency, low VAF mutations could be due to various reasons including a selection for the outgrowth of a daughter cell *in vitro*, sequencing artefacts, and/or repetitive regions within the genome. Thus, it was vitally important that we develop optimal sequencing parameters to filter out these non-somatic variants.

Initially, we utilized the SNV caller MuTect[51] as this pipeline has been used to call SNVs in studies using clonal samples. As seen in Figure 11. A., when calling SNVs on the same deeply sequenced HSC colony as mentioned above, we viewed a high frequency of low VAF SNVs and a small peak around ~0.5 VAF. Following the application of prescribed filters[47], as well as a blacklist of repetitive regions (see Methods section 5.2), we were left with ~80% less SNVs spread over a more binomial distribution around ~0.5 VAF (initial SNV count at 89X was 2249 and 459 SNVs post-filtering) (Figure 11. B.). Later, we used the SNV calling pipeline CaVEMan[52] to call SNVs on the same HSC colony. The reason we decided to test this pipeline was because of a collaboration we began with the group of Inigo Marticorena lab at the Wellcome Trust Sanger Institute in Cambridge once we noticed discrepancies in our analysis using MuTect, which will be described at length in the following section. Like with MuTect, the distribution of SNVs and their corresponding VAFs showed a high frequency of low VAF variants, and small peak around ~0.5 VAF (Figure 11. C.). Notably, CaVEMan called ~40% less variants than MuTect (initial SNV counts of 2249 and 1313 for MuTect and CaVEMan, respectively). Following the application of filters and a blacklist, implemented with the help from the Martincorena lab, a prominent peak around ~0.5VAF was observed (Figure 11. D.). This blacklist was generated at the Wellcome Trust Sanger Institute and was comprised of repetitive and homologous regions to ignore, which were detected using multiple "control/normal" mouse genomes. Furthermore, CaVEMan filtering parameters not only focused on read depth, but also mapping and sequence read quality of reads covering each variant (see Methods section 5.2. for further details). An additional VAF filter was added based on the VAF distribution of the remaining SNVs, whereby low VAF, sub-allelic variants were filtered out (Figure 11. E.). This VAF-cutoff was based on the rationale that these mutations were not in the original HSC clone, but rather were acquired during *in vitro* culture. This left us with 206 SNVs, 84% less than the pre-filtered SNVs called by CaVEMan and 55% less than filtered MuTect calls, again illustrating the large difference between the amount of SNVs called by each pipeline.

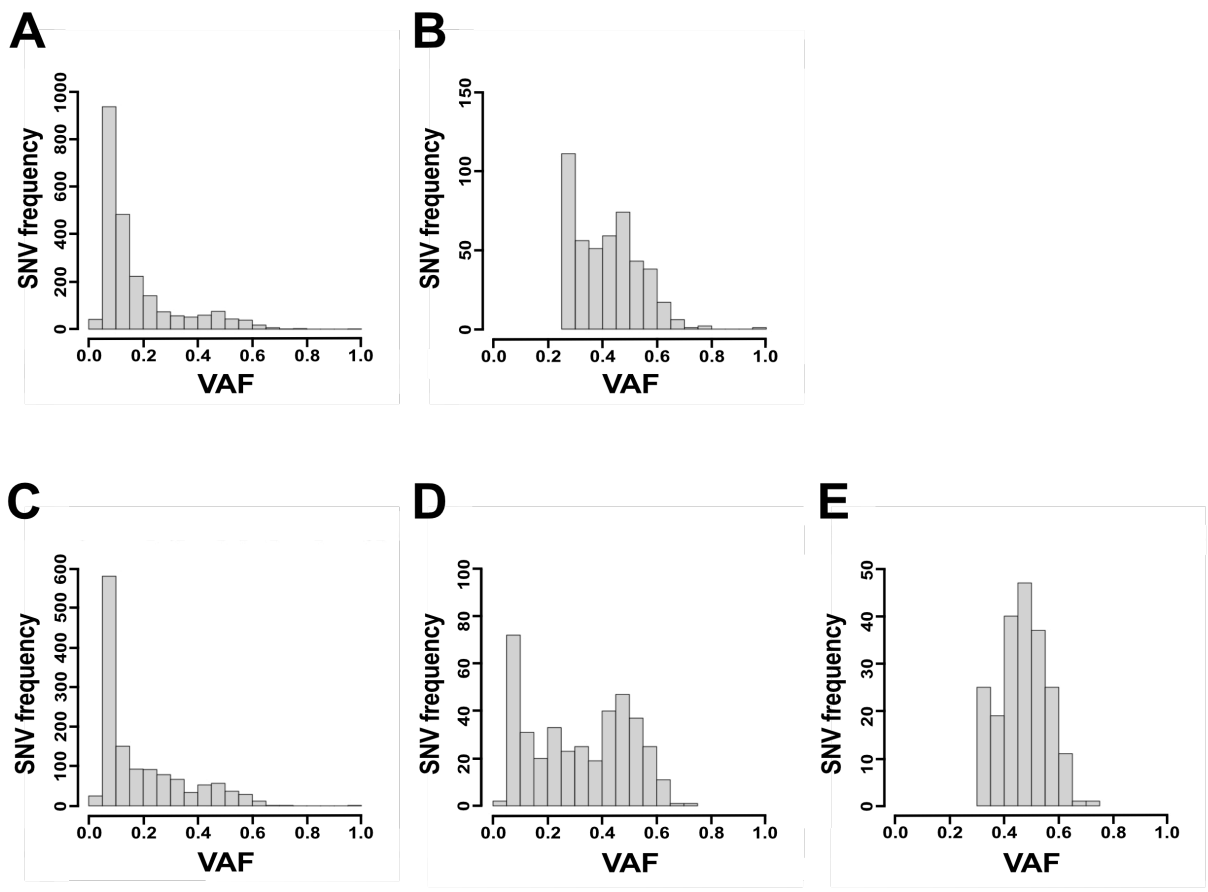


Figure 11. Variant allele frequency distributions for pre- and post-filtered MuTect and CaVEMan SNVs. Distribution of SNVs within a deeply sequenced (~90X, matched germline control), *in vitro* expanded colony derived from a single HSC, called by MuTect **(A)** pre- and **(B)** post-filtering. Distribution of SNVs called by CaVEMan on the same HSC, **(C)** pre- and **(D)** post-filtering, as well as **(E)** filtered by VAF.

1.1.2. Sensitivity assessment of SNV callers

In order to determine what coverage we required for the accurate analysis of somatic variants within HSC-derived colonies using Whole Genome Sequencing (WGS), we performed an initial sensitivity assessment of our SNV callers, MuTect and CaVEMan, as well as an INDEL caller, Pindel[53]. In brief, we downsampled our deeply sequenced HSC colony and matched germline control (both ~90X) in increments of 10X, using the Picard "DownsampleSAM" tool, and called variants on matched coverage between HSC colonies and DNA from tail samples (downsampling repeated in triplicate; see Methods section 5.3 and Table 4 for downsampling fractions). That is, bam files from the HSC colony and matched tail sample which had both been downsampled to

10X coverage, were used as input to call variants, and this was repeated for every coverage fraction. Initially this analysis was only performed using MuTect but after careful consideration of the results, we felt that MuTect was not the optimal pipeline for calling SNVs within our clonal samples. As such, we established a collaboration with the Martincorena lab at the Wellcome Sanger Institute, which has expertise in calling true somatic variants in clonal samples. The following results comparing the sensitivity of MuTect and CaVEMan will illustrate why MuTect did not meet our analysis requirements.

As mentioned before, MuTect calls a much higher number of SNVs than CaVEMan, however these increase exponentially with coverage and only begin to plateau at around 60X coverage (Figure 12. A., purple line). As previously mentioned, after applying the additional filters we see a dramatic decrease in SNVs called (Figure 12. A., orange line). However, an unusual result we observed was that there were more SNVs post-filtering at lower coverages, with the highest SNV count viewed at 20X, than at higher coverages (mean SNVs at 20X = 823; mean SNVs at 90X = 463). This seemed counterintuitive as we expected more SNVs to pass the filtering parameters at higher coverages. Not only because there is a higher read depth across the whole genome which would increase the confidence in SNV calls, but also there were more SNVs being called by MuTect at higher coverages pre-filtering. It is unclear why this phenomena occurred when analyzing our clonal samples and has not been noted in other studies where MuTect was used to call SNVs.

In comparison, although CaVEMan called fewer SNVs than MuTect at each matched coverage fraction, variant calls plateaued at a much lower coverage of ~30X (Figure 12. B., purple line). Furthermore, after applying additional filtering steps to these SNVs not only did we see a large decrease in variants calls, but variant calls plateaued again at ~30X (Figure 12. B., orange line).

1.1.3. Specificity assessment of SNV detection

In addition to assessing the sensitivity of each SNV caller for determining optimal WGS parameters, we also analyzed their specificity for calling true variants. The idea of this analysis

was adapted from Alioto, Buchhalter, Derdak *et al.*[50]. In short, for each SNV pipeline we have called the SNVs identified at the highest coverage and corresponding filtering parameters the "gold standard" of true SNVs. We used these gold standard SNVs for each pipeline and determined the overlap in SNVs (post-filtering) called at each downsampled coverage. Overlapping SNVs were called our "true positives", while any additional variants called that were not present within the gold standard were denoted as "false positives".

As seen in Figure 12. C., with MuTect the number of true positives keeps increasing with increasing coverage at an almost exponential rate (green line). Additionally, the number of false positives (red line) also increases with coverage until 20X, after which they decrease dramatically, in an almost inverse relationship to true positives. This high proportion of false positives at 20X is due to unexpected phenomena we viewed before, where there was an increased amount of filtered SNVs called at 20X compared to higher coverages. This high proportion of false positives at lower coverages is of concern as this means that until ~60X coverage, where the proportion of true and false positives is roughly equal at ~60%, the majority of filtered SNVs being called by MuTect are false positives. Therefore, in theory we would need to sequence each HSC clone to atleast ~60X to be confident that roughly half of the SNVs we are calling are true SNVs. Not only was this inadequate for our downstream analysis, but also not feasible due to financial and sample restrictions.

Like MuTect, the rate of true positives called by CaVEMan increases with coverage (Figure 12. D., green line). However, unlike MuTect, the rate of increasing true positives begins to plateau at around 30X where ~75% of true positive SNVs are called. Furthermore, the rate of false positive calls (red line) appears to be stable regardless of coverage, maintaining a proportion of 10-17% until 60X, after which the rate of false positives decreases with coverage until the gold standard coverage of 90X.

In summary, although MuTect calls many more SNVs than CaVEMan, regardless of the coverage or filtering parameters, CaVEMan outperformed MuTect both in terms of sensitivity and

specificity. While variant calls which passed pipeline filtering only plateaued at ~60X coverage for MuTect, SNVs which passed CaVEMan filtering plateaued at ~30X (purple line, Figures 12. A. and B.). Unexpectedly, we viewed an increase in filtered calls at 20X compared to higher coverages with MuTect (orange line, Figure 12. A.) and observed that the majority of these SNVs are false positives (Figure 12. C., red line). Furthermore, the rate of true positives (Figure 12. C., green line) with coverage increased in an almost exponential fashion, while the inverse was viewed for false positives.

Conversely, we viewed a gradual increase in the rate of filtered SNVs called per coverage with CaVEMan, which began to plateau at ~30X (orange line, Figure 12. B.). Similarly, the rate of true positives versus coverage initially increased dramatically until beginning to plateau at around 30X, where ~75% of true positives were called (Figure 12. D., green line). Furthermore, the false positive rate was stable for CaVEMan (10-17% false positive SNVs), regardless of the coverage (Figure 12. D., red line). Taken together, we determined that sequencing each HSC colony to ~30X was sufficient for our downstream analysis using CaVEMan and lower coverages would be sub-optimal for confidently calling true SNVs within an HSC clone.

1.1.4. Sensitivity and specificity assessment of INDEL caller

In addition to calling SNVs, we were also interested in looking at insertions and deletions (INDELs) present within our HSC colonies. As such, we performed sensitivity and specificity analyses of our chosen INDEL calling pipeline, Pindel, using the same deeply sequenced HSC colony. We decided to use this INDEL calling pipeline following the advice from our collaborator's at Sanger who have also used Pindel to call somatic variants within clonal samples. The concepts for the sensitivity and specificity analyses are the same as described above in our SNV caller comparisons. In Figure 12. E., we see an almost linear increase in INDELs called with coverage (purple line). However, after filtering out non-clonal mutations by applying a cut-off of >0.3 VAF, not only did the amount of INDELs called decrease dramatically per coverage, but the rate at which INDELs increased with coverage also reduced (orange line). This rate plateaued at ~30X, with very few additional INDELs being called beyond this coverage. Furthermore, when assessing the rate of true positives with coverage, we could see it begin to plateau at around ~50X (Figure

12. F., green line). However at 30X, 70% of the true INDELs were called. The rate of false positives versus coverage was less stable than viewed for CaVEMan, but the proportion of false positives is less than 33% of total INDELs per coverage, regardless of the coverage (Figure 12. F., red line). Therefore, for HSC colonies sequenced to ~30X, we can be confident that the majority of the calls are true positive (~70%), while false positives encompass a smaller proportion of the calls (~30%). Given the small number of INDELs called at 30X coverage (mean = 30 INDELs), we felt these proportions of true and false positives would be adequate for our ongoing analysis of mutation burden in HSCs under different experimental conditions.

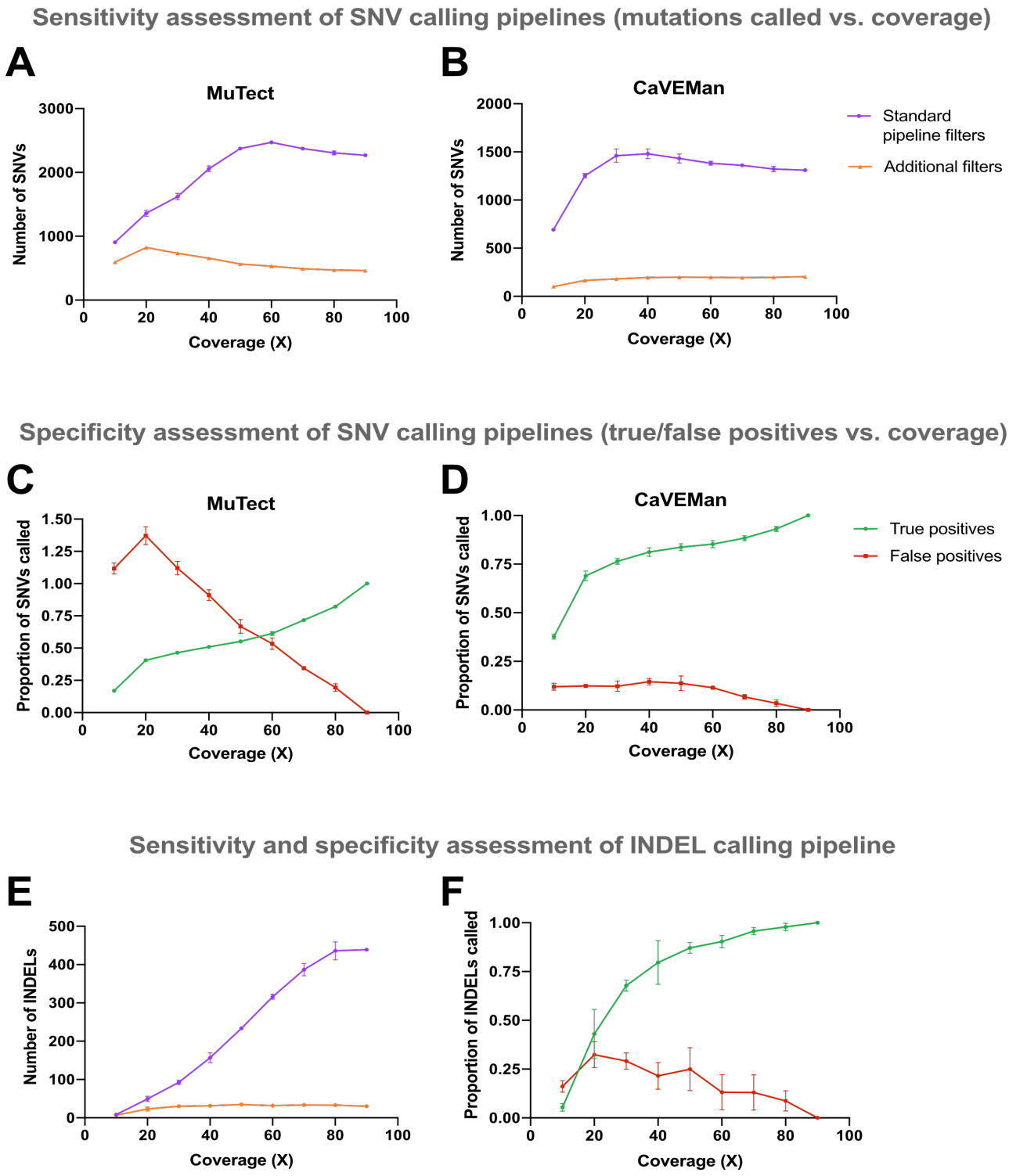


Figure 12. Effect of filtering parameters and coverage on the ability to call variants in HSC colonies using WGS. (A), **(B)** and **(E)** Sensitivity assessment of variant callers MuTect, CaVEMan and Pindel, respectively. Plots show the number of variants called at various downsampled coverages of a deeply sequenced HSC colony (against matched

coverage germline control). Purple line denotes all variants which passed the variant pipeline filters. Orange line denotes all variants which passed additional filtering. **(C)**, **(D)** and **(F)** Specificity assessment of variant callers MuTect, CaVEMan and Pindel, respectively. Plots show the proportion of true positive (green line) and false positive (red line) variants per coverage. **All plots**: downsampling per coverage was performed in triplicate. Each dot represents the mean and error bars indicate SD.

1.1.5. Assessment of overlapping SNVs called by MuTect and CaVEMan

Due to the requirements of our downstream analysis, it was important for our SNV calling analysis to be highly sensitive and specific. As such, we wanted to assess whether using the overlapping SNVs called by the two pipelines was more suitable for our analysis than only using one. In Figure 13. A. we observed that although MuTect (red line) called many more SNVs than CaVEMAn (blue line) at any given coverage, the number of overlapping SNVs was similar to those called by CaVEMAn at all coverages (green line). It should be noted that these were filtered SNV calls and that the red line in Figure 13. A. corresponds to the orange line in Figure 12. A. above, while the blue line in Figure 13. A. matches that of the orange line in Figure 12. B.. This was further demonstrated when we looked at the proportion of SNVs from each caller within the overlapping SNVs (Figure 13. B.). This proportion was calculated as the amount of SNVs called by pipeline A which were present within the subset of overlapping SNVs between pipeline A and B at coverage X. That is, how many of the total SNVs called by either MuTect/CaVEMan which were present within the MuTect and CaVEMan overlapping SNVs at any given coverage. Here we can we see that from 20X onwards, there was a plateau where just under 90% of the SNVs called by CaVEMan were present within the overlapping subset of SNVs (blue line). Therefore, CaVEMan was able to call 90% of the SNVs which were called by both pipelines and as such, only called an additional ~10% of SNVs which were not called by MuTect. Conversely, the proportion of MuTect filtered SNVs present within the overlapping SNVs from both pipelines was relatively low and increased with increasing coverage (Figure 13. B., red line). This illustrates that even at the highest coverage of ~90X, less than 40% of the variants called by MuTect are present within the overlap of both pipelines. Therefore, MuTect calls over 60% more SNVs at any given coverage which do not overlap with CaVEMan SNVs.

Additionally, we performed a similar analysis to that described above but this time looking at the overlap in true positives called by each pipeline. Therefore, MuTect-assigned true positives illustrated in Figure 12. C. above (green line), are depicted in Figure 13. C. as the red line. While CaVEMan-assigned true positives (Figure 13. C., blue line) correspond to the true positives depicted in Figure 12. D. (green line). Similar to Figure 13. A., there was an increase in the number of MuTect-assigned true positives with coverage (Figure 13. C., red line), while there was a plateau in the number of CaVEMan-assigned true positives (blue line), as well as overlapping true positives (green line) from ~30X onwards. The number of overlapping true positives was very close to that observed for CaVEMan-assigned true positives whereby from 20X onwards, over 85% of CaVEMan-assigned true positives were present within the overlapping true positives from both pipelines (Figure 13. D., blue line). Furthermore, although over 60% of MuTect-assigned true positives were present within the overlapping true positives from 10X, after 60X the proportion of MuTect-assigned true positives present within the overlapping true positives starts to decrease (red line). This was due to the observed ever-increasing rate of true positives called by MuTect with increasing coverage. This demonstrates that although MuTect calls more true positives with increasing coverage, the majority of these are additional true positives which are not present within the overlap of both pipelines.

To summarize, although using the overlap of SNV calls from two pipelines increases the specificity of calling true SNVs, we found that there was little difference between the results from the overlapping SNVs of both MuTect and CaVEMan, and those from CaVEMan alone. CaVEMan was able to call close to 90% of SNVs found within the overlapping SNVs from both pipelines, and over 90% of the overlapping true positives from 20X onwards. Therefore, due to the small number of SNVs called per HSC colony, using a combination of two pipelines to call SNVs does not increase the specificity of our analysis enough to warrant the extra work. Furthermore, this would be an unnecessary waste of time and resources and as such, for our analysis of HSC colonies we chose to use CaVEMan alone to call SNVs.

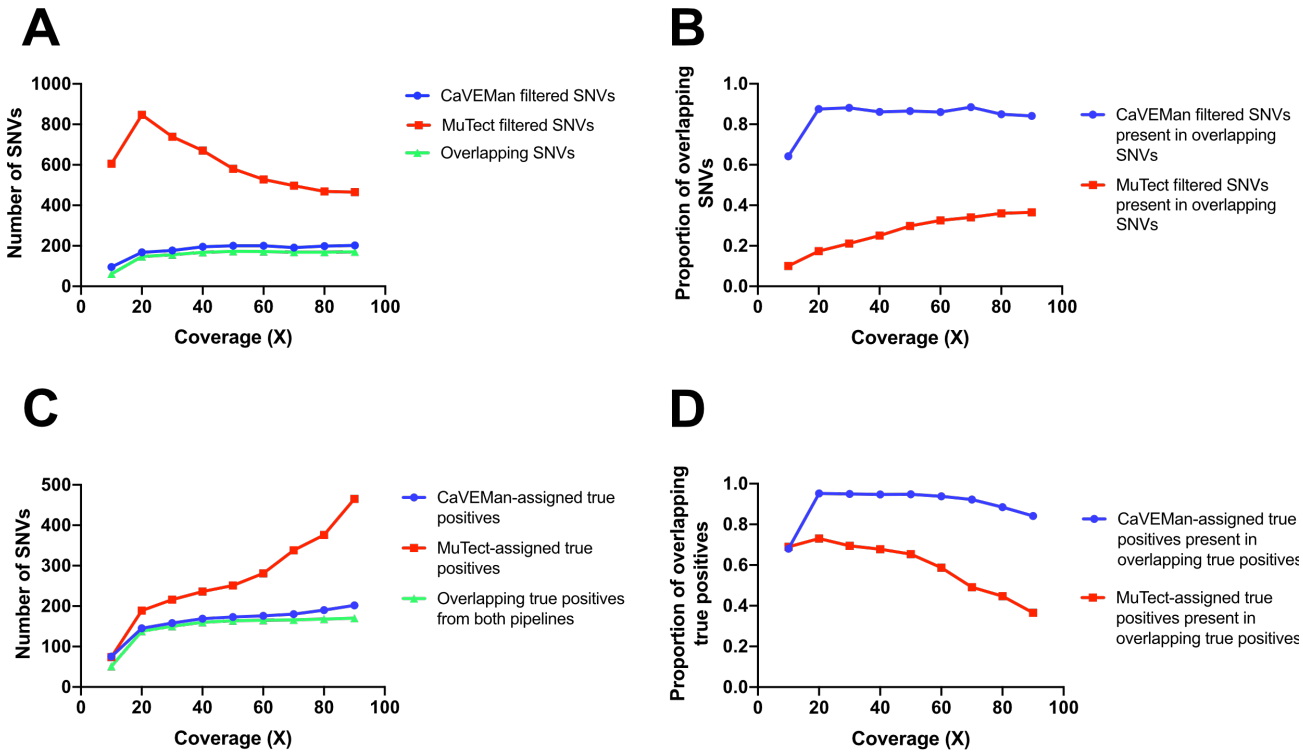


Figure 13. Assessment of results from individual SNV callers, MuTect and CaVEMan, compared to the overlapping results from both. (A) Comparison of the number of filtered SNVs called by MuTect (red line), CaVEMan (blue line) and the overlap of both (green line). **(B)** Proportion of MuTect (red line) and CaVEMan (blue line) filtered SNVs present within the overlapping SNVs from both pipelines. **(C)** Comparison of the number of MuTect-assigned true positive SNVs (red line), CaVEMan-assigned true positives (blue line) and overlapping true positives (green line). **(D)** Proportion of MuTect-assigned (red line) and CaVEMan-assigned (blue line) true positives present within the overlapping true positive SNVs from both pipelines.

1.2. Comparison of quality control output from sequenced HSC clones

Once we had established the optimal variant analysis parameters and determined a minimum coverage depth of ~30X for the accurate and sensitive analysis of our HSC colonies, we proceeded to sequence multiple colonies from various age and treated mice. In brief, we sequenced three mice per age/treatment group and aimed to sequence 2-4 HSC colonies per mouse, as well as corresponding germline control per mouse (tissue from the tail). Thus, in total, we have sequenced over 70 HSC colonies, and their matched germline controls, from young, old, pIpC- and TPO-treated mice, as well as label-retaining and non-label-retaining cells (LRC and non-LRC,

respectively) from old mice which have been labelled with GFP in their HSC compartment (scltTa H2B GFP). We additionally sequenced HSC colonies from a mouse which had been treated with *Mycobacterium avium*. However due to an issue with missing samples we were only able to sequence colonies from one mouse and have therefore excluded data from these colonies from our analysis. Besides HSC colonies which had clear contamination, all sequenced colonies were analyzed and their results presented in this thesis.

1.2.1. Sequencing coverage obtained was on average higher than the target coverage

By sequencing one HSC sample/germline control per lane on the Illumina HiSeqX Ten platform (150bp paired-end), we were able to obtain an average sequencing depth over 35X for HSCs from all groups (Figure 14. A.), and their corresponding germline controls. Although the library preparation protocol that was used stipulated a minimum DNA input of 100ng, the majority of the colonies were not large enough to obtain 100ng. As such, we optimized the fragmentation and PCR amplification steps of the protocol to account for lower DNA inputs (see methods section 4.1.2 and supplementary Figure 4). Thus, we successfully sequenced HSC colonies with DNA inputs between 25-100ng.

1.2.2. Higher duplicate rates correlated with lower sequencing coverage, but not lower DNA inputs

On closer inspection of the sequencing alignment quality control (QC) assessments (provided by the DKFZ ODCF), we noted that six HSC colonies did not pass quality control. These HSC colonies failed due to a high duplicate rate (duplicate rate cut-off threshold is 25%). Interestingly, 5 of the 6 HSCs were from either pIpC- or TPO-treated mice, while the remaining colony was from the non-LRC group. All of these six colonies had a coverage lower than 35X, with the lowest being 23X. This was the lowest coverage viewed across all of our sequenced HSC colonies. As expected, we saw a trend where samples with a higher duplicate rate had a lower coverage (Figure 14. B.). However, only ~50% of the variability can be accounted for when assessing the impact of duplicates on coverage using a simple linear regression (R^2 = 0.5381). Surprisingly, lower DNA input did not have a strong correlation with higher duplicate rate (R^2 = 0.2540) (Figure 14. C.).

Even though 4 of the 6 colonies which failed duplicate rate QC had DNA inputs lower than 50ng, there were three times as many HSC colonies which had DNA inputs lower than 50ng but still passed QC. Therefore, lower DNA input did not necessarily result in a higher duplicate, and subsequently lower coverage. Lastly, although we previously illustrated that coverage impacts on the number of variants called (Results section 1.2. above), we saw no correlation between the number of SNVs called and coverage (R^2 = 0.0041; Figure 14. D.). A similar result was viewed looking at INDELs (R^2 = 0.0022; Figure 14. E.). Taken together, we were able to achieve a sequencing depth for all of our colonies which was sufficient for an accurate analysis of mutation burden. Therefore, differences in mutation burden are not due to variability in coverage but are rather an intrinsic feature of the individual HSCs.

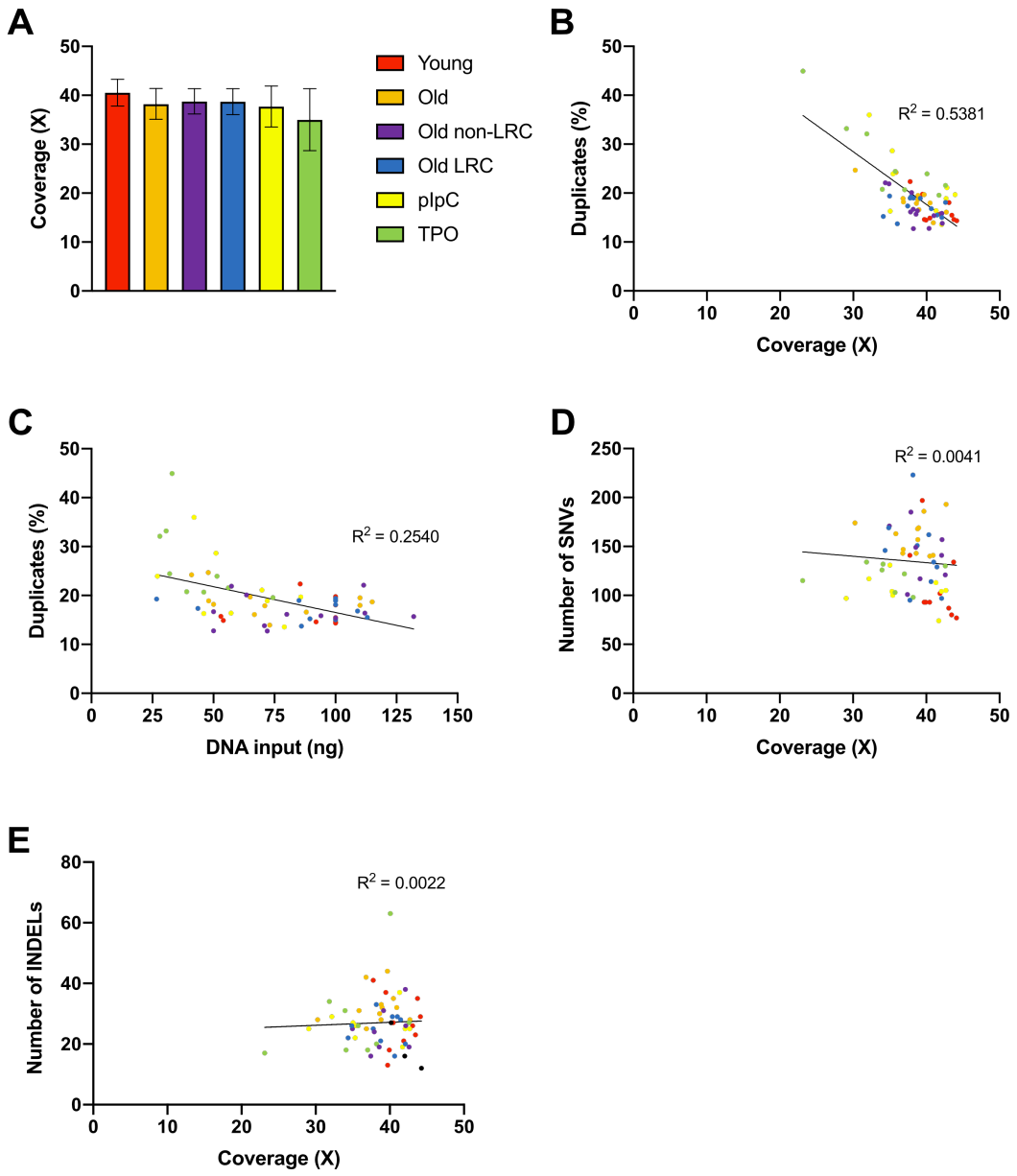


Figure 14. Effect of genomic DNA input and alignment quality control measures on mutation burden. (A) Mean coverage of HSC colonies across age and treatment groups. Error bars indicate SD. **(B)** Percentage duplicates and coverage per HSC colony. **(C)** Percentage duplicates and DNA input (ng) per colony. **(D)** Number of SNVs and coverage per HSC colony. **(E)** Number of INDELs and coverage per HSC colony. For plots **(B)** - **(E)**: each dot represents one HSC. Color of dots represents group using the same legend as in **(A)**. Simple linear regression model was performed on the HSC colonies and R^2 value is indicated on the plot.

2. Age-associated mutation burden accumulation within murine hematopoietic stem cells

2.1. SNV mutation acquisition increases with age

To assess whether the rate of SNV acquisition is associated with ageing, we sequenced HSC clones from young (~8 months of age) and old (~22 months of age) C57BL/6 mice (Figure 15. A.). From here on, young and old mice are denoted as "young" and "old", respectively. Three mice were sequenced per age group and four HSC colonies, as well as matched germline control, were sequenced per mouse (n = 12 HSC colonies per age group). We viewed a statistically significant difference in the number of SNVs between young and old mice HSCs ($p = 0.0001$; unpaired two-tailed t test; Figure 15. B.), with a positive correlation between the age of the mouse and HSC SNV burden ($p < 0.05$; simple linear regression; Figure 15. C.). This indicates that there is a gradual accumulation of SNVs within the HSC compartment of a mouse during its lifetime. Using the results from the simple linear regression model, we could estimate an annual rate of 39.8 SNVs per year per HSC clone (95% confidence intervals (CIs) are 22.3-57.2, respectively; Figure 15. D.).

We then repeated the same analysis as above, this time taking into account our biological replicates. This was done to validate that we did not introduce a bias in our analysis by inflating the effective sample size when assessing SNV differences using all HSC colonies per group as our data input. Therefore we compared the SNV burden of the three mice from our young group, with the three mice from our old. We simply took the averages of all SNVs from the HSC colonies per mouse, and then repeated the unpaired t test analysis using the average SNV burden per mouse (n =3 for young group; n = 3 for old group). We viewed similar results whereby there was a statistically significant difference in SNV burden between young and old mice ($p = 0.0008$; unpaired two-tailed t test; Figure 16. A.), with a positive correlation of increasing SNV burden in HSCs with age ($p < 0.05$; simple linear regression; Figure 16. B.). These similar results were likely due to the SNV burden across HSCs within each group being relatively stable, therefore there is minimal skewing of the data. We performed a multiple comparisons test comparing the average SNV burden per mouse using an ordinary one-way ANOVA. There was no significant difference between mean SNV burden per mouse in the young ($p = 0.8277$; Figure 16. C. and D.) and old

group (p = 0.7193; Figure 16. E. and F.). Therefore, there is a stable distribution of SNV burden across HSCs for each group and no significant differences in SNV accumulation between mice from the same age group. As such, comparing HSCs between young and old groups does not inflate our effective sample to such a degree where our results are biased.

2.2. INDEL mutation acquisition increases with age

Although we detected a limited number of INDELs, we still observed a statistical difference in INDEL burden between HSC colonies from young and old mice (p value = 0.0387; unpaired two-tailed t test; Figure 15. E.), with a positive correlation between the number of INDELs and age ($p < 0.05$; simple linear regression; Figure 15. E.). Based on the results from the simple linear regression model, we could estimate an annual rate of 4.8 INDELs per year per HSC (95% confidence intervals (CIs) are 0.3-9.4, respectively; Figure 15. G.).

Similar to the SNV results above, we saw a stable distribution of INDELs across HSCs for each group and no significant differences in INDEL burden between young mice (p value = 0.6397; ordinary one-way ANOVA; Figure 17. C. and D.) and old mice (p value = 0.5577; ordinary one-way ANOVA; Figure 17. E. and F.). Furthermore, there was a statistical difference in INDEL burden between young and old mice (p = 0.0432; unpaired two-tailed t test; Figure 17. A.), where INDEL burden positively correlated with age ($p < 0.05$; simple linear regression; Figure 17. B.).

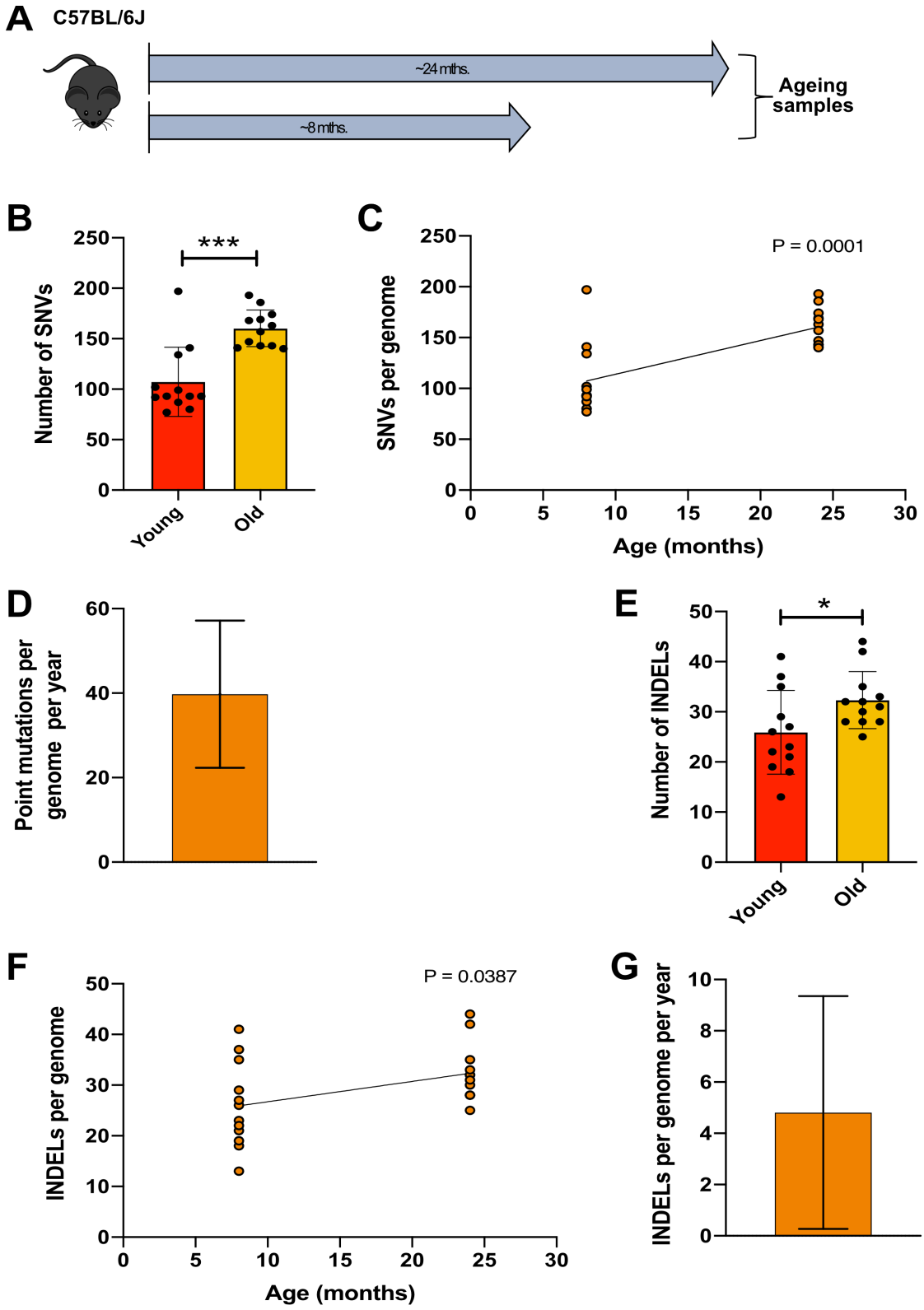


Figure 15. Age-associated increase in mutation burden within HSC colonies. (A) Schematic overview of ageing mouse cohort. Analyzed HSC colonies from three young (~8 month old; denoted as "Young") mice and three old

(~24 month old; denoted as “Old”) mice. **(B)** Mean number of SNVs in Young and Old HSCs. Each data point represents one HSC colony. Error bars indicate SD. Bar indicates a statistical difference between the number of SNVs in HSCs from Young and Old mice (unpaired two-sided t test; *** = $p < 0.001$). **(C)** Correlation of the number of SNVs per genome with age. Each dot represents a single HSC colony. Simple linear regression was performed and the p value is indicated above the plot (unpaired two-sided t test). **(D)** Annual SNV rate estimated using the linear regression model in **(C)**. Error bars indicate the 95% confidence intervals of the slope estimation. **(E)** Mean number of INDELs in Young and Old HSC colonies. Each data point represents one colony. Error bars indicate SD. Bar indicates a statistical difference in INDEL burden between Young and Old HSCs (unpaired two-sided t test; * = $0.01 < p < 0.05$). **(F)** Correlation of the number of INDELs with age. Each dot represents a single colony. Simple linear regression was performed and the p value is indicated above the plot (unpaired two-sided t test). **(G)** Annual rate of INDELs estimated using the linear regression model in **(F)**. Error bars indicate the 95% confidence intervals of the slope estimation.

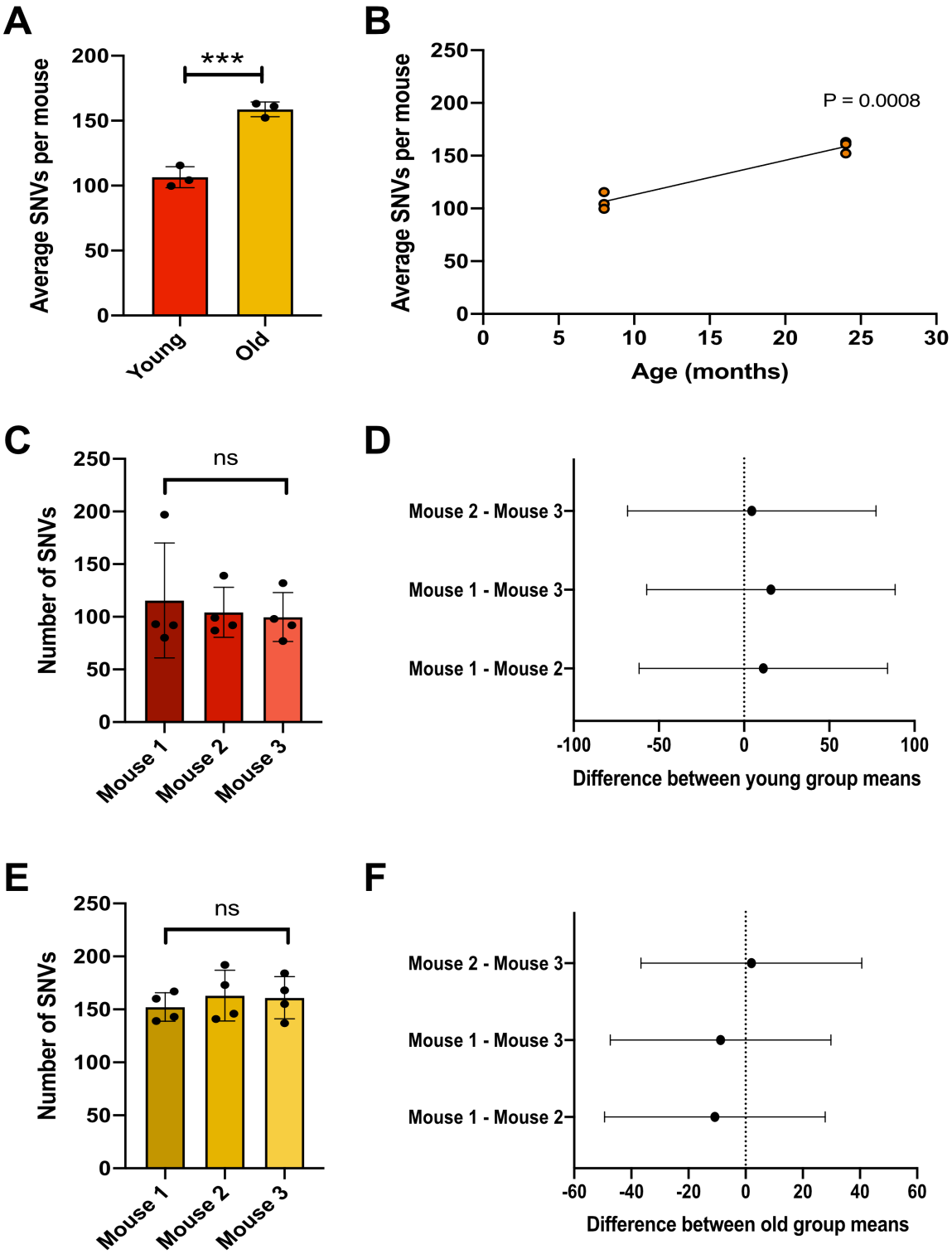


Figure 16. Comparison of SNV burden between mice from young and old groups. (A) Mean number of SNVs in Young and Old mice. Each data point represents one mouse. Error bars indicate SD. Bar indicates a statistical difference between the number of SNVs in mice from Young and Old groups (unpaired two-sided t test; *** = p <

0.001). **(B)** Correlation of the number of SNVs per mouse with age. Each dot represents a single mouse. Simple linear regression was performed and the p value is indicated above the plot (unpaired two-sided t test). **(C)** Comparison of the mean SNV burden per young mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(D)** Plot depicting the difference in means between mice from the young group. Bar depicts 95% confidence interval. **(E)** Comparison of the mean SNV burden per old mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(F)** Plot depicting the difference in SNV means between mice from the old group. Bar depicts 95% confidence interval.

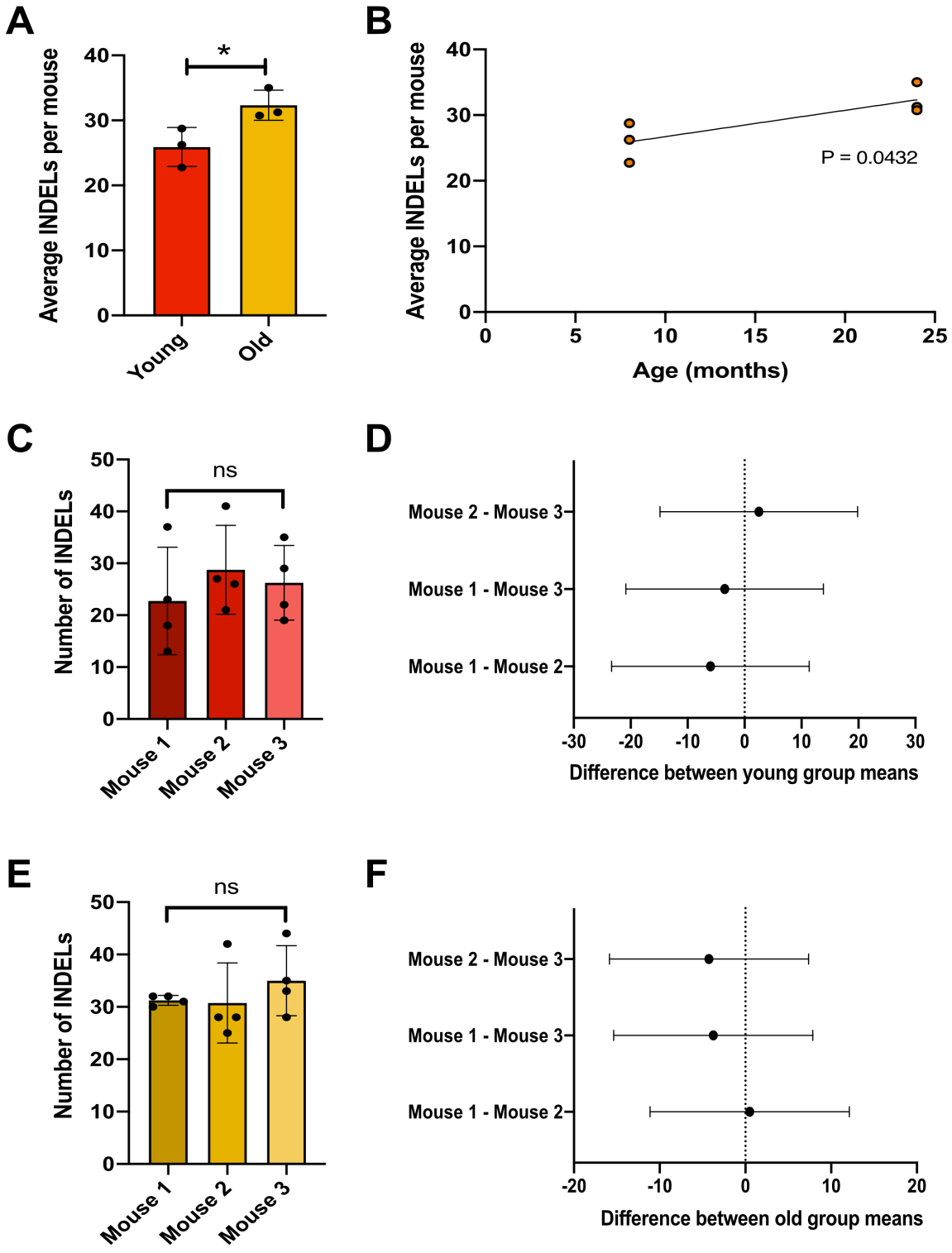


Figure 17. Comparison of INDEL burden between mice from young and old groups. (A) Mean number of INDELs in Young and Old mice. Each data point represents one mouse. Error bars indicate SD. Bar indicates a statistical difference between the number of INDELs in mice from Young and Old groups (unpaired two-sided t test; * = 0.01 <

$p < 0.05$). **(B)** Correlation of the number of INDELs per mouse with age. Each dot represents a single mouse. Simple linear regression was performed and the p value is indicated above the plot (unpaired two-sided t test). **(C)** Comparison of the mean INDEL burden per young mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(D)** Plot depicting the difference in means between mice from the young group. Bar depicts 95% confidence interval. **(E)** Comparison of the mean INDEL burden per old mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(F)** Plot depicting the difference in INDEL means between mice from the old group. Bar depicts 95% confidence interval.

3. Mutation burden within HSCs increases with replication history

3.1. SNV mutation acquisition increases with proliferation

Next we wanted to assess whether the extent of DNA replication and HSC has undergone is a primary driving force behind mutation acquisition rate during ageing. To do this, we utilized a label-retention model, in which the HSC compartment of mice is transiently labelled with a GFP-H2B fusion protein which, when switched off using doxycycline treatment, is diluted with every subsequent cell division[54]. Using this model, we were able to separate and expand label-retaining (LRC) and non-LRC HSCs from three old mice (~22 months of age) (Figure 18. A.). These represent HSCs which have remained predominantly dormant throughout their lifetime and those which have undergone HSC cycling, respectively. From here on, label-retaining colonies will be denoted as LRC HSCs and non-label-retaining colonies will be denoted as non-LRC ("old LRC" and "old non-LRC" in figures). Twelve LRC (4 HSCs from each mouse; n = 12) and eleven non-LRC (4 HSCs from 2 mice, 3 HSCs from 1 mouse n = 11) HSC colonies from the same three mice were sequenced, as well as their matched germline controls. There was a statistically significant increase in SNV burden in non-LRC HSCs compared to LRC HSCs ($p = 0.0093$; unpaired two-tailed t test; Figure 18. B.). Furthermore, we saw a similarity in the average SNV burden between the non-LRC HSCs and normal old HSCs (mean of old group = 160.3; mean of non-LRC group = 155.9 SNVs), while the average SNV burden of the LRC HSCs was significantly closer to that of the young HSCs (mean of LRC group = 122.3; mean of young group = 107.3 SNVs). Considering that the mice from this label-retention model are ~22 months of age and mice from the old group are ~24 months old, and the majority of HSCs are non-LRCs after such a lengthy chase period, it was unsurprising that their SNV burden is similar. On the other hand, HSCs from the same ~22 month old mice which

still retained their GFP label had a much lower SNV burden which closer represented those from ~8 month old young mice. This could indicate that replication is a key mechanism of ageing and mutation burden within HSCs, and that HSCs which have replicated less throughout their lifetime exhibit a "younger" SNV burden phenotype. As such, there was no statistical difference in the number of SNVs between old and non-LRC HSCs ($p = 0.6486$; unpaired two-tailed t test; Figure 18. B.), but SNV burden between old and LRC HSCs was statistically different ($p = 0.0009$; unpaired two-tailed t test; Figure 18. B.). Furthermore, no statistical significance was viewed between SNV numbers from young and LRC HSCs ($p = 0.2587$; unpaired two-tailed t test; Figure 18. B.), but SNV burden of young HSCs was statistically different from those of non-LRC ($p = 0.0012$; unpaired two-tailed t test; Figure 18. B.).

The correlation between the age of young and non-LRC, and the age of young and LRC HSCs was plotted using a simple linear regression model (Figure 18. C.). There was a positive correlation between age and HSC SNV burden when comparing young and non-LRC HSCs ($p < 0.05$). However no statistically significant correlation of SNV burden with age was viewed when comparing young and LRC HSCs ($p > 0.05$). This indicates that there is a gradual accumulation of HSC SNVs with replication, while SNV acquisition remains stable within dormant HSCs. Using these results, we could estimate an annual rate of 41.6 SNVs per year for non-LRC HSCs (95% confidence intervals (CIs) are 18.5-64.8, respectively), which is similar to the annual SNV rate viewed in our old group (Figure 18. D.). Conversely, the estimated annual rate when comparing SNV burden between young and LRC HSCs is much lower at 12.9 SNVs per year (95% confidence intervals (CIs) are -10.1376 - 35.8560, respectively).

When taking into account SNV burden per mouse for each group, SNV burden between LRC and non-LRC HSCs was statistically significant ($p = 0.0409$; unpaired two-tailed t test; Figure 19. A.). Furthermore, SNV burden correlated with age for HSC colonies from the non-LRC group ($p < 0.05$; simple linear regression; Figure 19. B.), but there was no correlation between SNV burden and age observed when assessing HSC colonies from the LRC group ($p > 0.05$; simple linear regression; Figure 19. B.). We again saw a stable distribution of SNV burden across HSCs for each group and

no significant differences in SNVs between analyzed mice from our label-retention model (p value = 0.5240; ordinary one-way ANOVA; Figure 20. A. and B.). Furthermore, when assessing differences in mice, taking into account the label status of each HSC, we saw no differences in SNVs between the mice for non-LRC HSCs (p value = 0.6850; ordinary one-way ANOVA; Figure 20. C. and D.) and for LRC HSCs (p value = 0.2570; ordinary one-way ANOVA; Figure 20. E. and F.). These results show no different conclusions compared to the results mentioned above where SNV burden was assessed across all HSCs for each group.

3.2. INDEL mutation acquisition increases with proliferation

Like the SNV results above, we noted a significant increase in INDEL burden within non-LRC HSCs compared to LRC HSCs (p value = 0.0241; unpaired two-tailed t test; Figure 21. A.). Therefore, INDEL burden also increased in HSCs with replication. Furthermore, we observed no significant difference in INDEL burden between old and non-LRC HSCs (p = 0.0.0577; unpaired two-tailed t test), although the overall distribution of INDELs appears to be lower than those observed in the old group. Conversely, the number of INDELs is significantly lower in LRC HSCs compared to HSCs from the old group (p = 0.0001, respectively; unpaired two-tailed t test). As no statistical difference was observed between INDELs viewed in young and non-LRC, and young and LRC HSCs, we did not perform a linear regression model to determine annual INDEL rate (p = 0.5954 and p = 0.1928, respectively; unpaired two-tailed t test).

Similar to the results above, the average INDEL burden per mouse for LRC and non-LRC HSCs was statistically significant (p = 0.0253; unpaired two-tailed t test; Figure 21. B.). Additionally, we saw a stable distribution of INDELs across HSCs for each group and no significant differences in INDEL burden between analyzed mice from our label-retention model (p value = 0.54600; ordinary one-way ANOVA; Figure 22. A. and B.). Furthermore, when taking into account the label status of each HSC, we saw no differences in INDELs between the mice for non-LRC HSCs (p value = 0.6251;

ordinary one-way ANOVA; Figure 22. C. and D.) and for LRC HSCs (p value = 0.6603; ordinary one-way ANOVA; Figure 22. E. and F.).

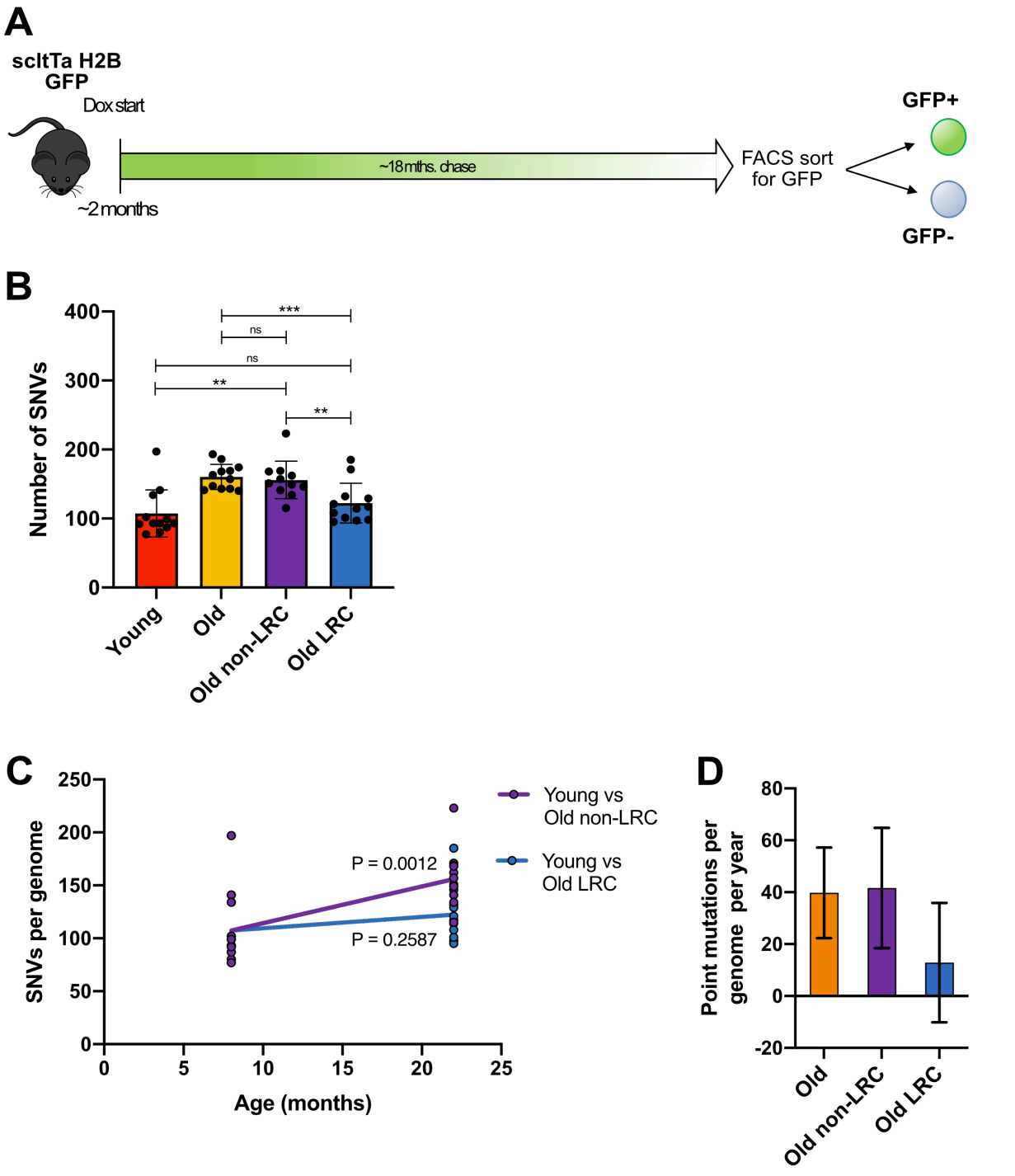


Figure 18. Replication-associated increase in mutation burden within HSC colonies. (A) Schematic overview of label-retention model. Label-retaining (GFP+) FACS sorted HSCs represent HSCs which have remained dormant, while non-label-retaining (GFP-) HSCs represent HSCS which have undergone replication. Analyzed label-retaining

(denoted as “old LRC”) and non-label-retaining (denoted as “old non-LRC”) HSC colonies from 3 ~22 month old mice. **(B)** Mean number of SNVs in Old, Old non-LRC and LRC HSCs. Each data point represents one HSC colony. Error bars indicate SD. Bar indicates statistical difference between groups (unpaired two-sided t test; ** = 0.001 < p < 0.01; *** = p < 0.001; ns = p > 0.05). **(C)** Correlation of the number of SNVs per genome with age. Each dot represents a single HSC colony. Simple linear regression was performed and the p value is indicated in the plot (unpaired two-sided t test). Correlation between Young and Old non-LRC HSCs is indicated by a purple line and correlation between Young and Old LRC HSCs is indicated by a blue line. **(D)** Annual SNV rate estimated using the linear regression model in **(C)**. Error bars indicate the 95% confidence intervals of the slope estimation. **(E)** Mean number of SNVs in Young, Old non-LRC and LRC HSCs. Each data point represents one HSC colony. Error bars indicate SD. Bar indicates a statistical difference between groups (unpaired two-sided t test).

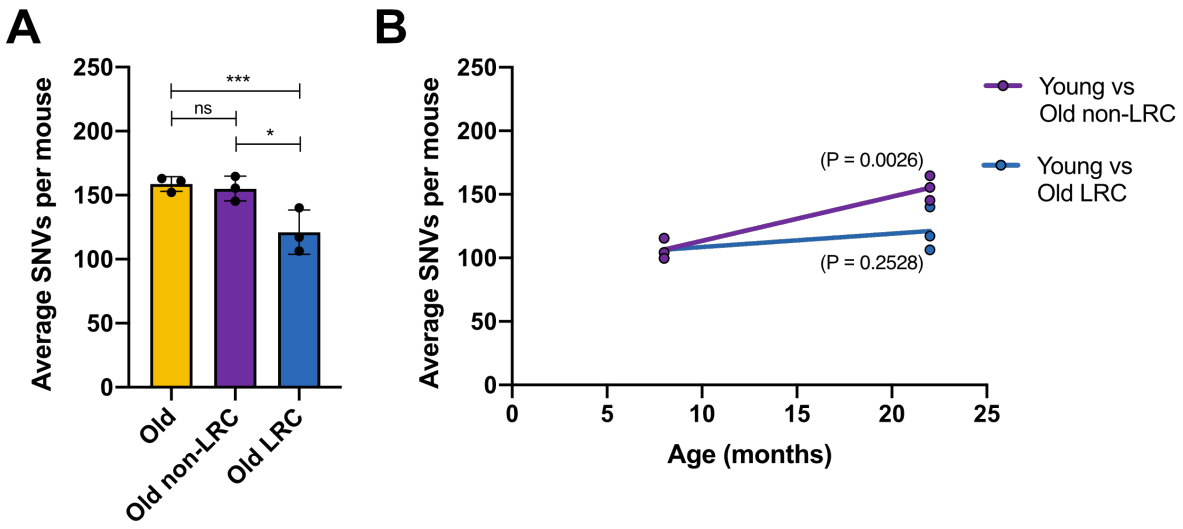


Figure 19. Comparison of SNV burden between mice from LRC and non-LRC groups reveals an increase in SNVs with replication. (A) Mean number of SNVs in Old, Old non-LRC and Old LRC mice. Each data point represents one mouse. Error bars indicate SD. Bar indicates a statistical difference between groups (unpaired two-sided t test; * = 0.01 < p < 0.05; *** = p < 0.001; ns = p > 0.05). **(B)** Correlation of the number of SNVs per mouse with age. Each dot represents a single mouse. Simple linear regression was performed and the p value is indicated above the plot (unpaired two-sided t test).

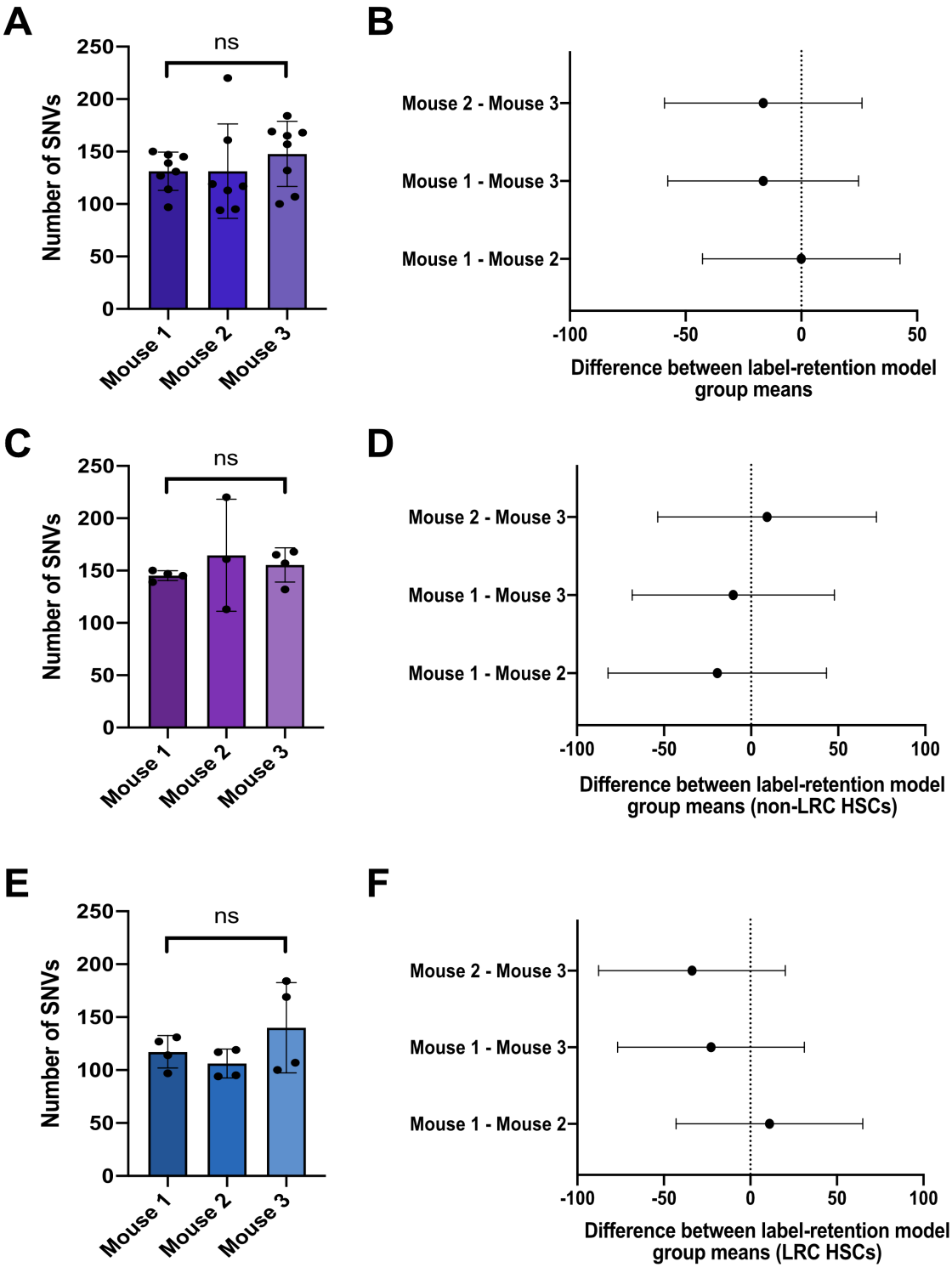


Figure 20. Comparison of SNV burden between mice from LRC and non-LRC groups. (A) Comparison of the mean SNV burden per mouse from the label-retention model using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(B)** Plot depicting the difference in SNV means between mice from the label-retention model.

Bar depicts 95% confidence interval. **(C)** Comparison of the mean SNV burden per mouse for non-LRC HSCs using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(D)** Plot depicting the difference in SNV means between mice from the non-LRC HSCs group. Bar depicts 95% confidence interval. **(E)** Comparison of the mean SNV burden per mouse for LRC HSCs using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(F)** Plot depicting the difference in SNV means between mice from the LRC HSCs group. Bar depicts 95% confidence interval.

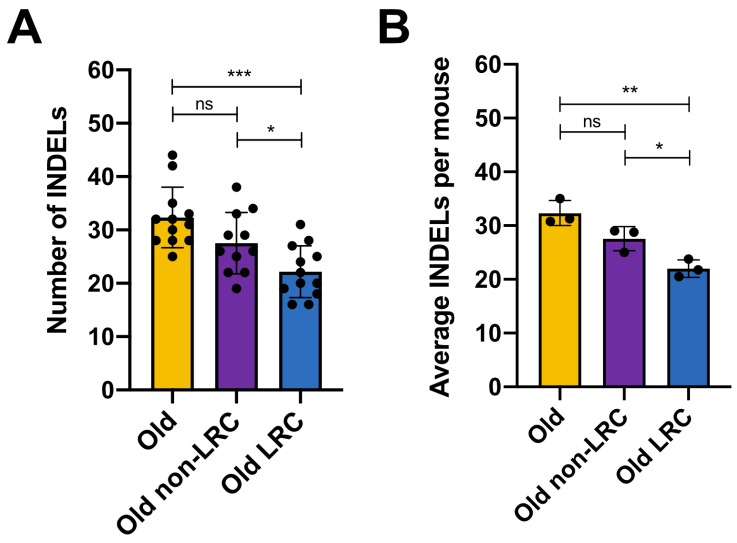


Figure 21. Replication-associated increase in mutation burden within HSC colonies. (A) Mean number of INDELs in Old, Old non-LRC and Old LRC HSCs. Each data point represents one HSC. Error bars indicate SD. Bar indicates a statistical difference between groups (unpaired two-sided t test; * = $0.01 < p < 0.05$; *** = $p < 0.001$; ns = $p > 0.05$). **(B)** Mean number of INDELs in Old, Old non-LRC and Old LRC mice. Each data point represents one mouse. Error bars indicate SD. Bar indicates a statistical difference between groups (unpaired two-sided t test; * = $0.01 < p < 0.05$; ** = $0.001 < p < 0.01$; ns = $p > 0.05$).

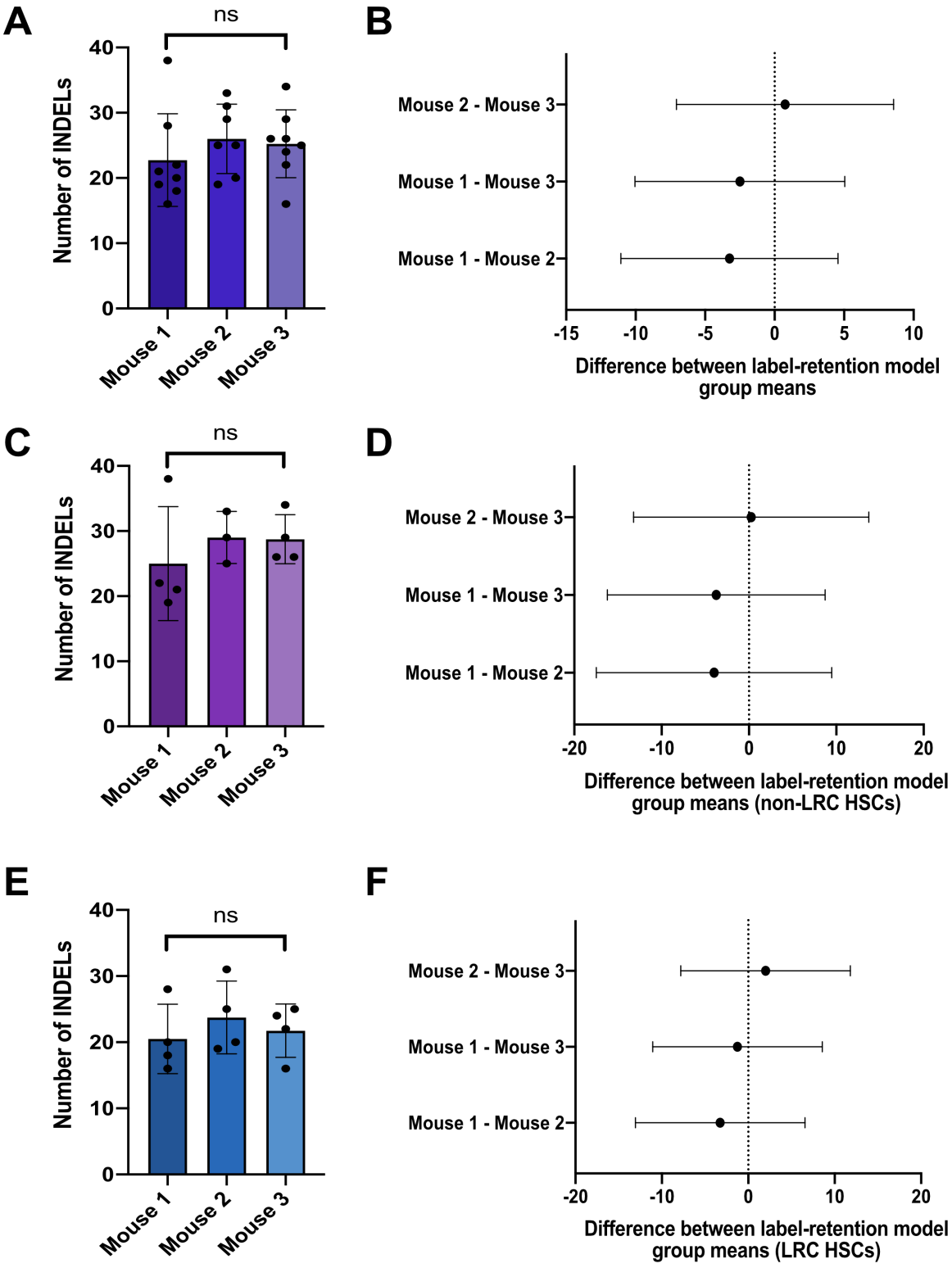


Figure 22. Comparison of INDEL burden between mice from LRC and non-LRC groups. (A) Comparison of the mean INDEL burden per mouse from the label-retention model using a one-way ANOVA. Ns bar indicates no significant

difference ($p > 0.05$). **(B)** Plot depicting the difference in INDEL means between mice from the label-retention model. Bar depicts 95% confidence interval. **(C)** Comparison of the mean INDEL burden per mouse for non-LRC HSCs using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(D)** Plot depicting the difference in INDEL means between mice from the non-LRC HSCs group. Bar depicts 95% confidence interval. **(E)** Comparison of the mean INDEL burden per mouse for LRC HSCs using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(F)** Plot depicting the difference in INDEL means between mice from the LRC HSCs group. Bar depicts 95% confidence interval.

4. Mutation burden within HSCs does not increase with induced HSC proliferation

4.1. Comparison of mutation burden within HSCs from age-matched untreated, pIpC- and TPO-treated mice shows no difference in SNV and INDEL mutation acquisition

Based on the results from our label-retention model where we observed an increase in mutation burden within murine HSCs with replication, we wanted to identify whether induced HSC proliferation results in a subsequent increase in mutation burden. As such, we chose two separate treatment regimens, Polyinosinic:polycytidylic (pIpC) and Thrombopoietin (TPO) treatment, which have previously been shown to induce HSC cycling[40]. Figure 23. A. illustrates these treatment regimens. In short, mice were treated with three rounds of either pIpC or TPO. One round of treatment encompasses two injections with a prescribed dosage of agonist per week for four weeks (see Methods section 2.2 and 2.3 for treatment dosages), followed by a four week recovery period. After three rounds of treatment, the mice are ~8 months of age. Three mice were sequenced per treatment group and at least two HSC colonies, as well as their matched germline controls, were sequenced per mouse (pIpC n = 7, 3 HSCs from 1 mouse and 2 HSCs from 2 mice; TPO n = 9, 4 HSCs from 1 mouse, 3 HSCs 1 mouse, 2 HSCs from 1 mouse). No statistically significant differences in the number of SNVs were viewed between age-matched untreated HSCs (~8 month old young group) and HSCs from pIpC- ($p = 0.8127$; unpaired two-tailed t test) or TPO-treated mice($p = 0.2446$; unpaired two-tailed t test; Figure 23. B.). Similarly, we viewed no difference in the number of INDELs between young and pIpC-treated ($p = 0.6983$;

unpaired two-tailed t test), and young and TPO-treated HSCs (p = 0.5515; unpaired two-tailed t test; Figure 23. C.).

When assessing the differences in mutation burden between mice from the same group, we again saw no difference in the number of SNVs and INDELs between young and pIpC-treated (p = 0.5696 for SNVs and p = 0.5338 for INDELs; unpaired two-tailed t test), and young and TPO-treated HSCs (p = 0.2326 for SNVs and p = 3896 for INDELs; unpaired two-tailed t test; Figures 24. A. and B.). Furthermore, we saw no differences in SNV (p value = 0.8975; ordinary one-way ANOVA; Figure 25. A. and B.) and INDEL (p value = 0.3488; ordinary one-way ANOVA; Figure 26. A. and B.) burden between pIpC-treated mice, and no differences in SNV (p value = 0.4465; ordinary one-way ANOVA; Figure 25. C. and D.) and INDEL (p value = 0.5869; ordinary one-way ANOVA; Figure 26. C. and D.) burden between TPO-treated mice. As such, the distribution of SNVs and INDELs across HSCs for each group was stable resulting in no different conclusions than the ones made above when assessing mutation burden of all HSCs within a group.

4.2. Reduced HSC colony size from treated mice could have potentially introduced a selection bias

Considering that we saw an increase in mutation burden with proliferation, it was an unexpected result that pIpC- and TPO-induced HSC cycling did not result in differences in mutation burden. However, it must be noted that an unintentional bias in sample selection has occurred. As previously mentioned, we were able to optimize our library preparation protocol in order to sequence smaller HSC colonies with genomic content between 25-100ng. Although this was a sufficient threshold for all other analyzed groups, we were not able to sequence the majority of the pIpC- and TPO-treated colonies because they were too small to obtain the required minimum input DNA amount of 25ng. This is due to a phenomenon we have previously observed where HSC colonies from pIpC- and TPO-treated mice have a decreased clonogenicity potential. That is, HSCs sorted from pIpC- and TPO-treated mice gave rise to smaller colonies with lower cell counts compared to PBS/untreated, age-matched controls (Figure 23. D. and E.). This difference in HSC colony size between PBS- and pIpC-treated mice, and PBS- and TPO-treated mice was statistically

significant (unpaired two-tailed t test; clonogenicity analysis performed by Ruzhica Bogeska and Marleen Büchler-Schäff for pIpC and TPO, respectively). Figure 23. F. illustrates this difference between HSC colony size where we can see that colonies from PBS-treated mice are much larger than those from pIpC- and TPO-treated mice.

As such, when selecting HSC colonies to sequence, we were only able to prepare libraries and sequence larger colonies from pIpC- and TPO-treated mice. Therefore, we have potentially selected more "functional" HSCs which have a better proliferation capacity, and excluded less-functional and possibly more damaged HSCs from our analysis.

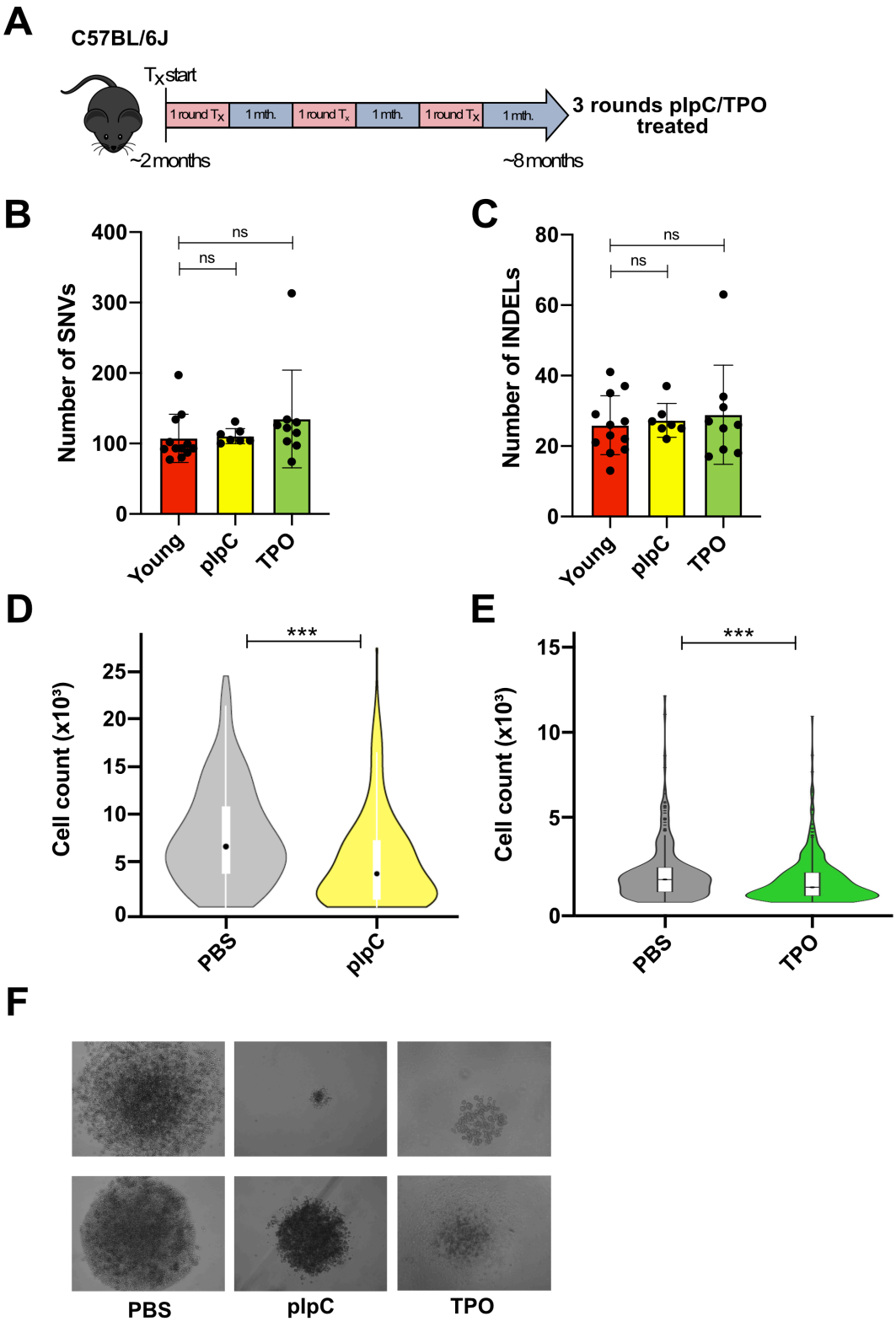


Figure 23. Murine HSCs exposed to replication-inducing agonists show no increase in mutation burden and a decrease in clonogenicity potential. (A) Schematic overview of pIpC and TPO treatment regimen. Mice are treated with three rounds of treatment (Tx) after which mice are ~8 months old. Three pIpC-treated mice (denoted as "pIpC)

and three TPO-treated mice (denoted as "TPO") were analyzed. **(B)** Mean number of SNVs in Young, pIpC and TPO HSCs. Each data point represents one HSC colony. Error bars indicate SD. Ns bar indicates no significant difference between groups (unpaired two-sided t test; $p > 0.05$). **(C)** Mean number of INDELs in Young, pIpC and TPO HSC colonies. Each data point represents one colony. Error bars indicate SD. ns bar indicates no significant differences in INDEL burden between groups (unpaired two-sided t test; $p > 0.05$). **(D)** Clonogenicity assay comparing HSC colony sizes between PBS- and pIpC-treated mice. Dots in the middle of the boxes indicate the average colony size per group. Bar indicates a statistical difference between the two groups (unpaired two-sided t test; *** = $p < 0.001$). **(E)** Clonogenicity assay comparing HSC colony sizes between PBS- and TPO-treated mice. Middle line with dot within box and plot indicates the average colony size per group. Bar indicates a statistical difference between the two groups (unpaired two-sided t test; *** = $p < 0.001$). **(F)** Microscopy images of two PBS- and two pIpC-treated HSC colonies (all images taken at 2.5X magnification and not edited after).

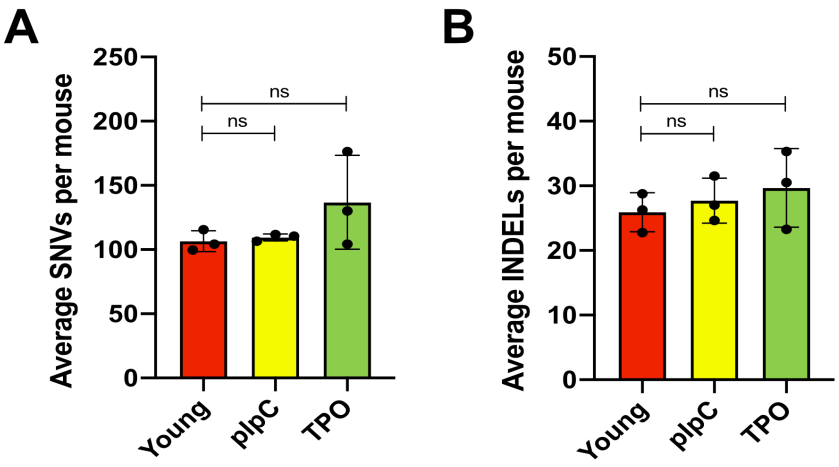


Figure 24. Comparison of SNV and INDEL burden between mice from young, pIpC- and TPO-treated groups reveals. **(A)** Mean number of SNVs in Young, pIpC- and TPO-treated mice. Each data point represents one mouse. Error bars indicate SD. Ns bars indicate no statistical difference between groups (unpaired two-sided t test; $p > 0.05$). **(B)** Mean number of INDELs in Young, pIpC- and TPO-treated mice. Each data point represents one mouse. Error bars indicate SD. Ns bars indicate no statistical difference between groups (unpaired two-sided t test; $p > 0.05$).

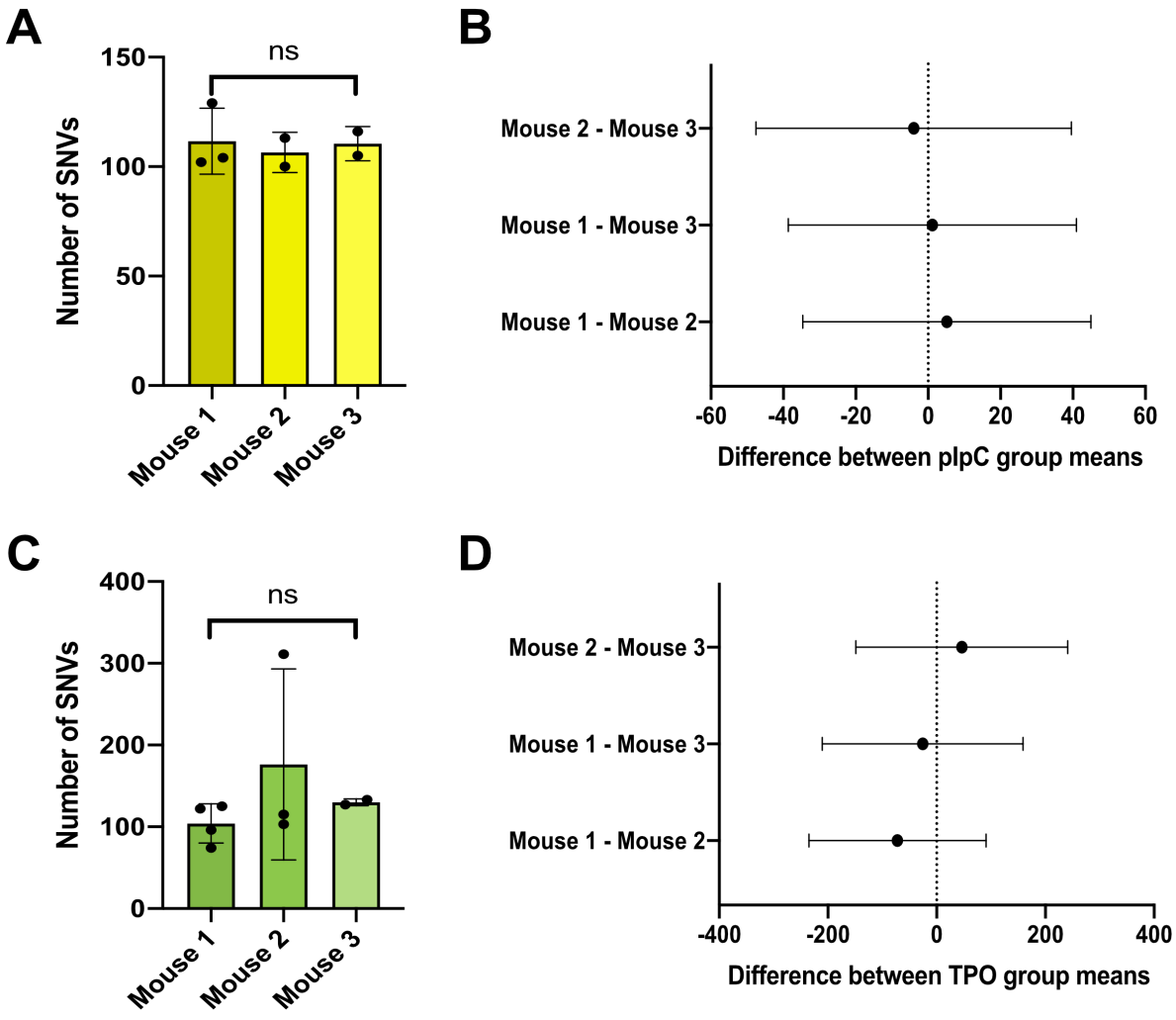


Figure 25. Comparison of SNV burden between mice from Young, pIpC- and TPO-treated groups. (A) Comparison of the mean SNV burden per pIpC-treated mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(B)** Plot depicting the difference in SNV means between pIpC-treated mice. Bar depicts 95% confidence interval. **(C)** Comparison of the mean SNV burden per TPO-treated mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(D)** Plot depicting the difference in SNV means between TPO-treated mice. Bar depicts 95% confidence interval.

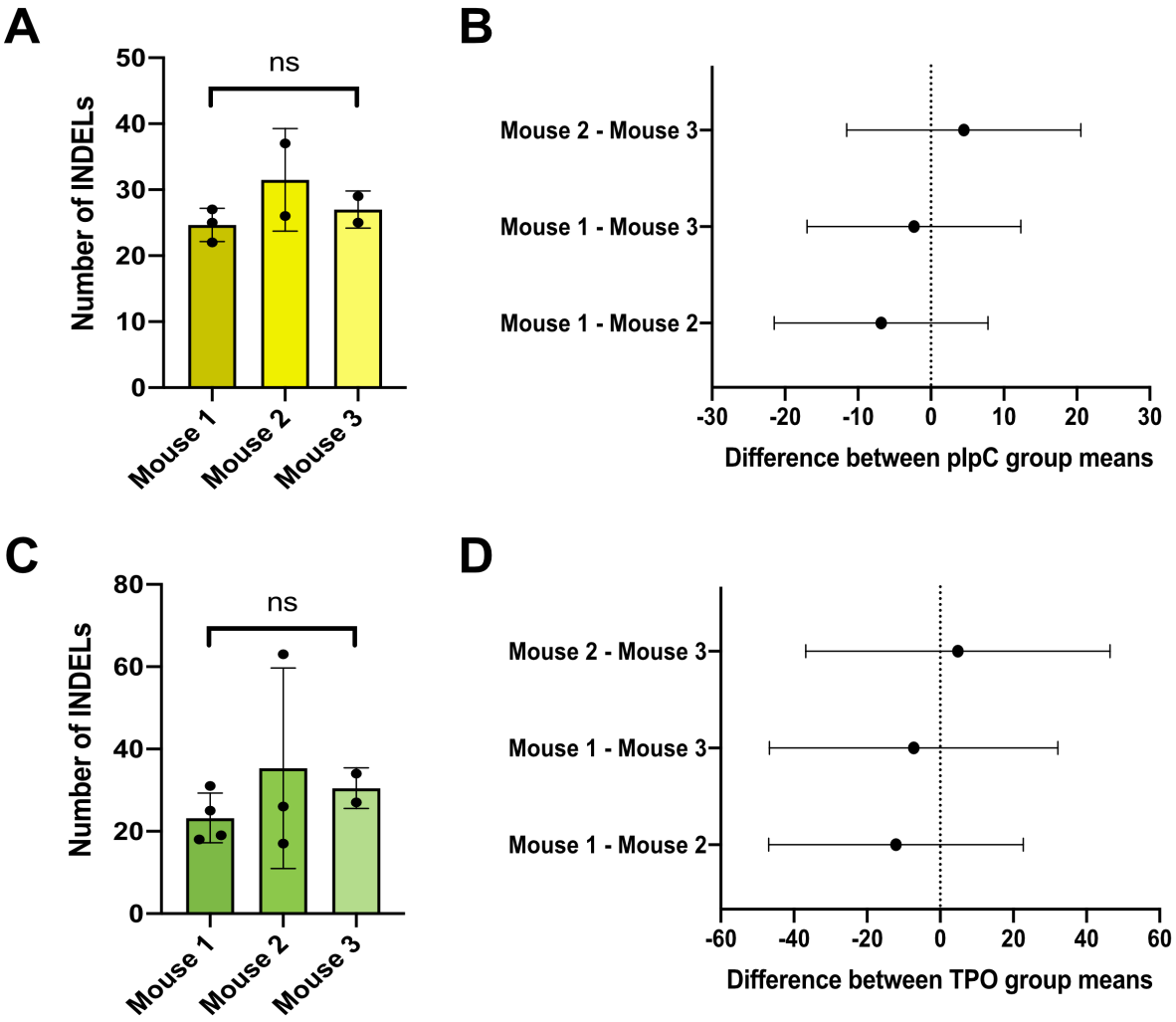


Figure 26. Comparison of INDEL burden between mice from Young, pIpC- and TPO-treated groups. (A) Comparison of the mean INDEL burden per pIpC-treated mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(B)** Plot depicting the difference in INDEL means between pIpC-treated mice. Bar depicts 95% confidence interval. **(C)** Comparison of the mean INDEL burden per TPO-treated mouse using a one-way ANOVA. Ns bar indicates no significant difference ($p > 0.05$). **(D)** Plot depicting the difference in INDEL means between TPO-treated mice. Bar depicts 95% confidence interval.

5. Assessment of SNV burden among age and treatment groups controlling for discrepancies in coverage by normalizing for callable genome size

As mentioned before, we observed an increase in mutation burden with age and proliferation. This was the case for both SNVs and INDELs. However, we did note a few outliers within each

group where certain HSCs had dramatically more SNVs than the majority of HSC colonies for that group. For example, three colonies (each from a different mouse) from the young group had much higher SNV burdens than the rest of the colonies from that group. These outliers within the young group and across all groups can be seen in Figure 27. A. Similarly, we saw variability in INDELs in HSC colonies from the same groups (Figure 27. B.). As such, we wanted to assess whether these outliers are true phenomenon representing HSCs with higher SNV burden or if this variability was rather due to variability in coverages.

Although we previously noted that the average sequencing coverage per group was fairly stable (Figure 14. A.), when we looked at the coverage per HSC colony we not only viewed variability in coverage across all HSCs, but also within groups (Figure 27. C.). In order to determine whether this variability in coverage plays a role in the variability in SNV burden within HSCs from the same groups, we decided to assess SNV burden per callable genome. Here the callable genome represents the size of the genome which was actually used for SNV calling. Details of determining the size of the callable genome per HSC is described in the Methods (section 5.4) but briefly, we subtract the regions which were not analyzed by CaVEMan from the reference genome in order to obtain the size of the callable genome in giga base pairs (Gbs). After identifying the callable genome sizes for each HSC colony, we were able to compare the number of SNVs per Gb per HSC colony. Encouragingly, we observed very similar results than those described before (Figure 27. D. and E.). Again when performing unpaired two-tailed t tests, we saw significant differences between young and old HSCs ($p = 0.0001$; unpaired two-tailed t test), non-LRC and LRC HSCs ($p = 0.0093$), young and non-LRC ($p = 0.0019$), and old and LRC HSCs ($p = 0.0004$; Figure 27. D.). Furthermore, no statistical differences were observed between young and LRC HSCs ($p = 0.3299$), old and non-LRC HSCs ($p = 0.4618$), young and pIpC-treated HSCs ($p = 0.8143$), and young and TPO-treated HSCs ($p = 0.8143$) (all comparisons performed using unpaired two-tailed t test). Additionally, we observed that the high SNV burden outliers still remain. This was due to the fact that even though there were some differences in global coverages across HSC colonies, the actual callable genome size was relatively similar between colonies from the same groups (Figure 27. F.). Therefore, we did not view any outlier colonies with higher SNV burden which also had a

larger callable genome, but rather the callable genome size was stable within groups. As such, this would indicate that these higher SNV burden HSC colonies are actually true events, demonstrating heterogeneity between HSCs from the same mouse that has not been viewed before.

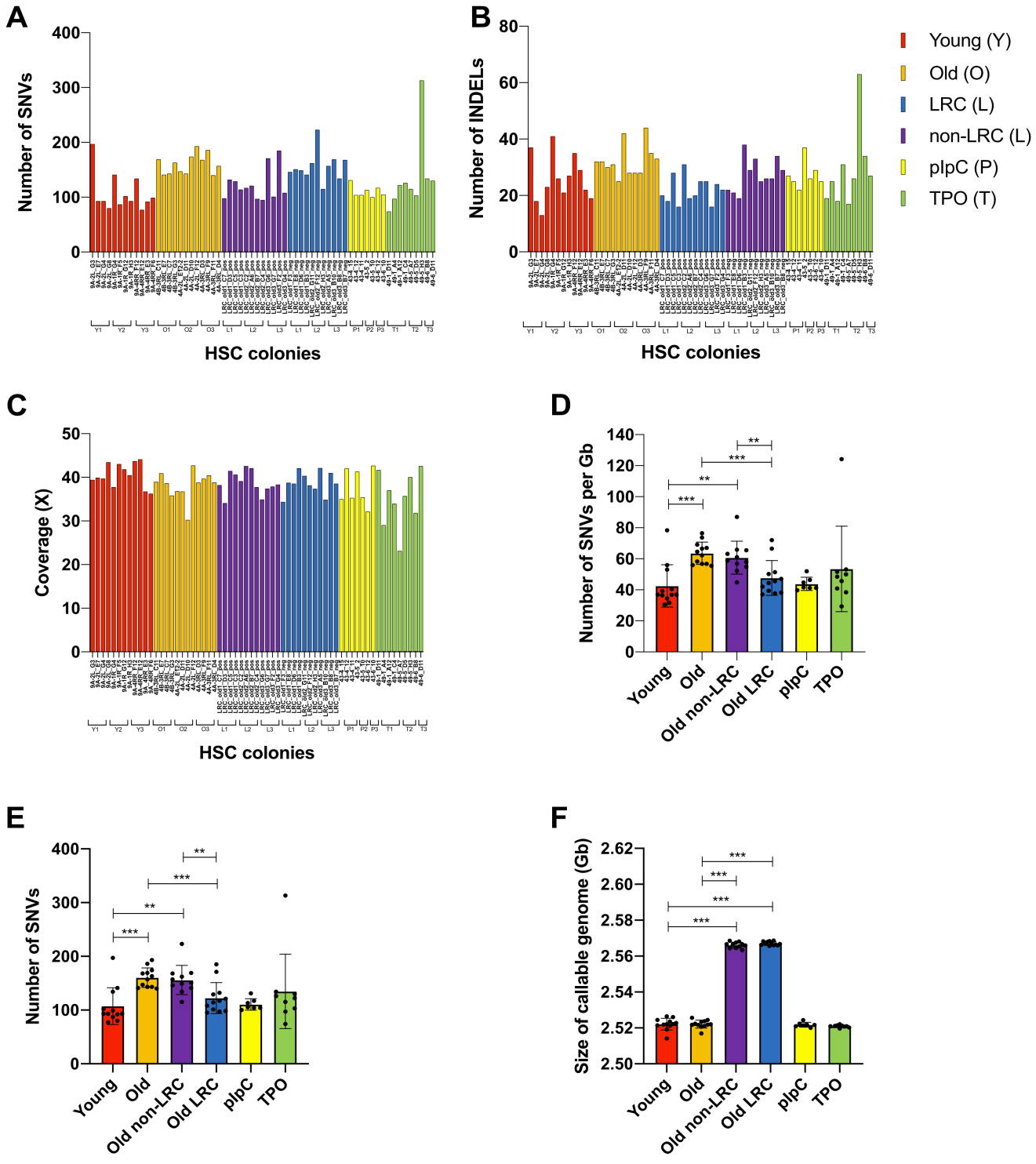


Figure 27. Variability in SNV burden across HSC colonies does not depend on differences in global coverage. (A) Number of SNVs per HSC across all analyzed groups. Figure legend is indicated on the top left. Letters in brackets in the legend represent an abbreviated version of each group name. Bars indicate which HSC colonies come from which mouse, with the code underneath indicating group, mouse and colony. **(B)** Number of INDELs per HSC across all

analyzed groups. Figure legend is indicated on the top left. Letters in brackets in the legend represent an abbreviated version of each group name. Bars indicate which HSC colonies come from which mouse, with the code underneath indicating group, mouse and colony. **(C)** Coverage per HSC across all analyzed groups. Figure legend is indicated on the top left. Letters in brackets in the legend represent an abbreviated version of each group name. Bars indicate which HSC colonies come from which mouse, with the code underneath indicating group, mouse and colony. **(D)** Mean size of callable genome (Gb) per group. Each dot represents one HSC. Error bars indicate SD. Bar indicates a statistical difference between groups (unpaired two-sided t test; ** = $0.001 < p < 0.01$; *** = $p < 0.001$). **(E)** Number of SNVs per group. Each dot represents one HSC. Error bars indicate SD. Bar indicates a statistical difference between groups (unpaired two-sided t test; ** = $0.001 < p < 0.01$; *** = $p < 0.001$). **(F)** Size of the callable genome per group. Error bars indicate SD. Bar indicates a statistical difference between groups (unpaired two-sided t test).

6. Mutational profile and signature analysis revealed an increase in signatures correlated with ageing

6.1. Mutation spectra did not differ between groups

Although we saw a significant increase in SNVs with ageing and proliferation, the genomic distribution of these SNVs remained stable throughout the groups (Figure 28. A.). That is, the majority of SNVs occurred within non-coding regions of the genome (intergenic, intronic and non-coding RNA within introns), while we observed less SNVs within protein-coding genes (exonic SNVs). We observed no other significant differences in the distribution of SNVs within other genomic regions between the groups. Unsurprisingly, we observed no statistically relevant differences in the number or type of exonic SNVs between the groups; that is, synonymous, nonsynonymous or stop gain variants (Figure 28. B.). This was due to the low number of exonic variants present within HSCs across the groups, ranging from 0-6 exonic SNVs per HSC colony. That being said, we did observe the presence of more nonsynonymous SNVs across all groups except in the young group (6 nonsynonymous SNVs and 7 synonymous SNVs across 12 HSC colonies). Furthermore, we observed a large difference in the number of nonsynonymous versus synonymous SNVs within the TPO-treated group (14 nonsynonymous SNVs and 3 synonymous SNVs across 9 HSC colonies). But again it must be noted that this was highly variable between the individual HSCs within groups and numbers were too low for statistical relevance.

As seen in Figure 28. C. and Figure 29., the mutational spectra of transitions and transversions did not differ between the HSC groups. Likewise, the mutational spectra of different INDEL types did not differ between the groups (Figure 28. D.). Taken together, these results suggest that the underlying mechanisms of mutation acquisition are the same throughout the groups. Considering that these colonies were all generated from the same cell type, HSCs, this could explain the lack of variability in the mutational spectra of HSCs from our various groups.

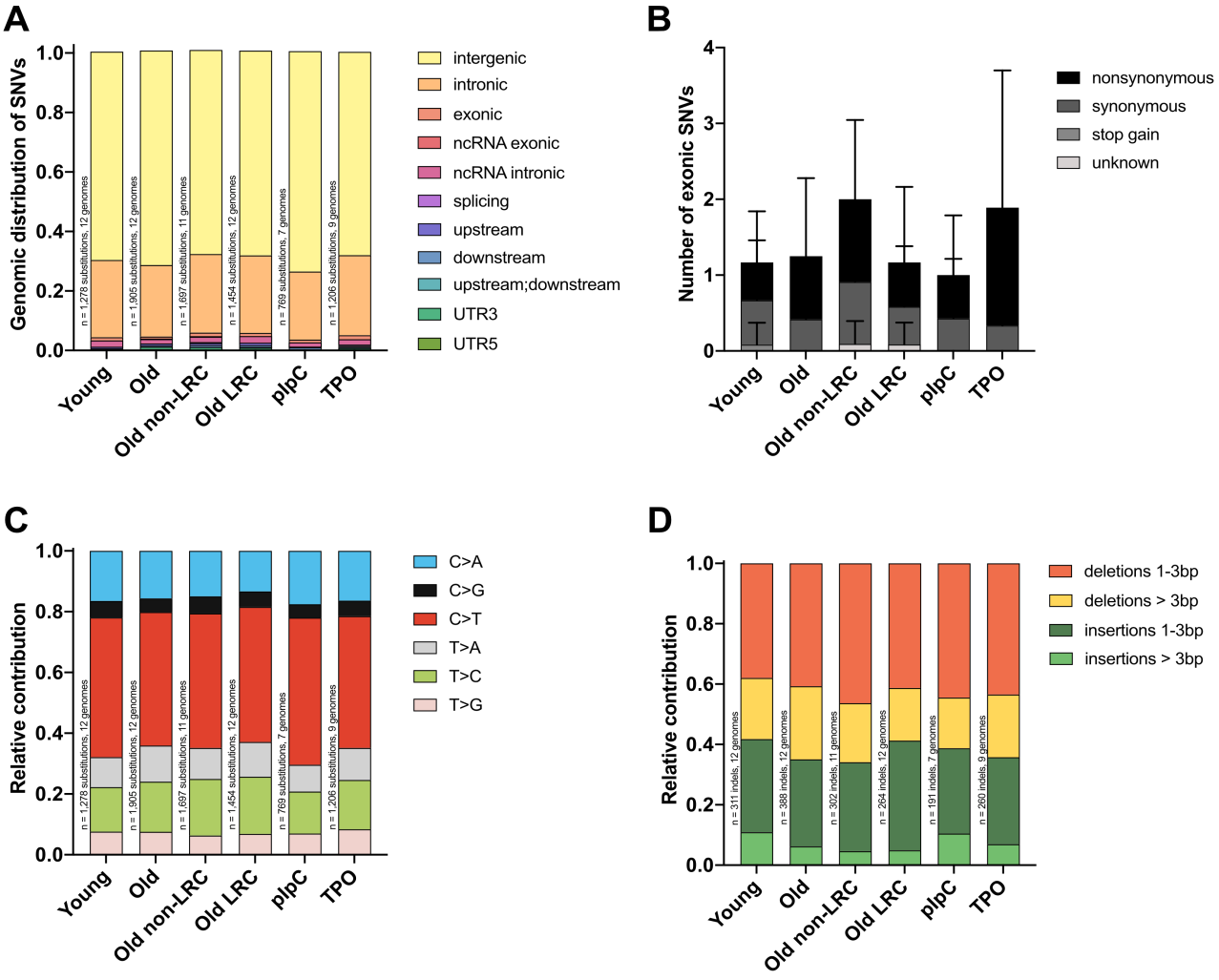


Figure 28. Genomic distribution of mutational changes revealed no distinct pattern between groups. (A) Distribution of SNVs per group found in each genomic region. The total number of SNVs and analyzed HSCs per group is indicated on the side of each bar. **(B)** Mean number and type of exonic SNVs for all HSCs per group. Error bars indicate SD. No statistically significant difference was observed between any groups (unpaired t test). **(C)** Relative contribution of each SNV mutation type per group. The total number of SNVs and analyzed HSCs per group is

indicated on the side of each bar. **(D)** Relative contribution of each INDEL mutation type per group. The total number of INDELs and analyzed HSCs per group is indicated on the side of each bar.

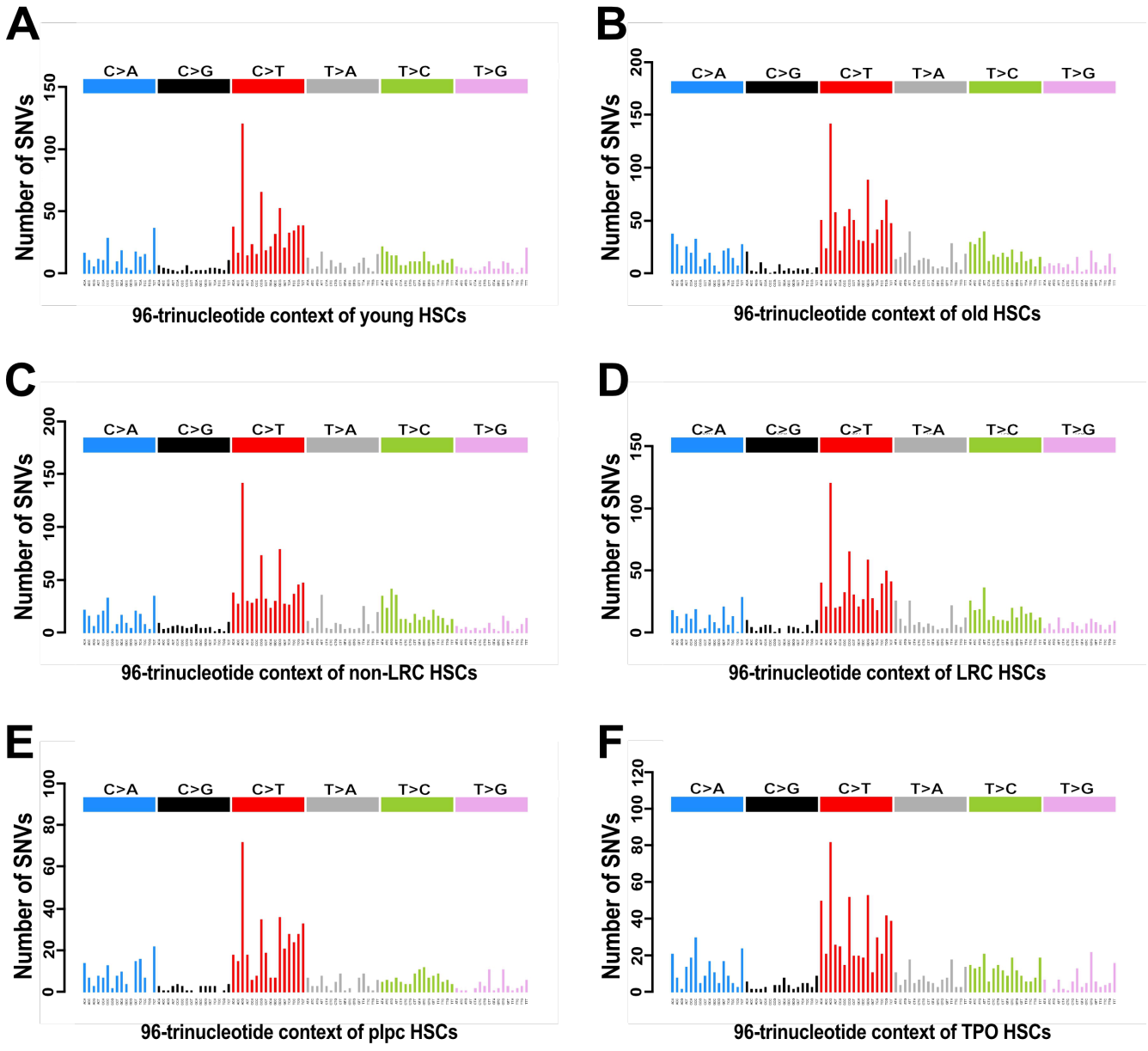


Figure 29. Mutational profiles of total SNVs per group. Total 96-trinucleotide context for all SNVs in: **(A)** young, **(B)** old, **(C)** non-LRC, **(D)** LRC, **(E)** plpC, and **(F)** TPO HSCs.

6.2. Contribution of mutational signatures between samples

6.2.1. Mutational signature analysis of individual HSCs

Although the types of mutations and mutation profiles did not differ much between groups, we wanted to assess whether we saw any differences with regards to mutational signatures. To identify these signatures, we performed a supervised analysis of mutational signatures within our HSCs using the package YAPSA[55] and the updated COSMIC mutational signatures[44]. COSMIC encompasses 49 SNV mutational signatures which were identified *de novo* from data from the PCAWG (Pan-cancer Analysis of Whole Genomes) consortium using two different calling algorithms, SigProfiler and SignatureAnalyzer. YAPSA utilizes the linear combination decomposition (LCD) function to identify the exposure of known signatures within a particular mutation catalogue. That is, when assessing our known SNVs within our cohort in the context of these known mutational signatures in COSMIC, we can use YAPSA to estimate the exposure of each mutational signature within our cohort's SNV catalogue. This is highly useful as YAPSA is not constrained by SNV catalogue or cohort size so we are theoretically able to assess the mutational signatures within individual HSCs. That being said, the mutational load per HSC was very low which we predicted would likely affect the confidence of the detected signatures.

This was confirmed when we looked at the mutational signatures per HSC across all groups (Figure 30. A.). Identifiers for each HSC colony are indicated along the x-axis and correspond to a list of identifiers (supplementary Table 2). From initial observation, there appeared to be no distinct patterns among HSCs from the same groups. A special feature of YAPSA is the ability to perform signature analysis using signature-specific cutoffs. It is known that different signatures have different detectability and those with high detectability will appear more often as false positives. In order to account for these differences in detectability, YAPSA introduced the use of pre-determined signature-specific cutoffs which can improve on the sensitivity and specificity of the identified mutation signature calls. Thus, after identifying exposures for each mutational signature, we apply these signature cutoffs. Any signatures which have an exposure contribution higher than their corresponding signature-specific cutoffs, are identified. The supervised analysis

of signature calling is then rerun using this subset of identified signatures and the contribution of each signature per sample is presented. From Figure 30. B., it was observed that seven signatures were identified at higher exposures than the signature-specific cutoffs. However, when we performed a cluster analysis to identify groups of samples which were exposed to similar mutational processes, i.e. clustering HSCs according to their exposures of these seven signatures, we viewed no distinct pattern for the HSCs which were clustering together (Figure 31. A.). That is, HSCs from the same mice/age/treatment did not cluster together but rather clustering seemed to be random when assessing mutation signatures across all HSCs. Additionally, we observed that the majority of clustering seemed to depend on the exposures of signatures 5, 1, 40, 3 and 18, which all exhibited atleast a few HSCs with exposure levels above 0.2. Signatures 2 and 7c exhibited lower signature exposures across all HSCs. This was further illustrated when we evaluated the confidence intervals (CIs) of exposures for each mutational signature per HSC (Figure 31. B.). YAPSA provides 95% CIs using the concept of profile likelihood and plots these according to each signature exposure per sample. In simple terms, YAPSA changes the exposure of a signature in an individual sample by a small value and then reruns the entire analysis using this alternative data model containing this newly altered exposure value and the remaining signatures. Therefore, this alternative data model has one degree of freedom less than the initial model. Then log-likelihoods are computed from the differences between the initial and the alternative model and a likelihood ratio test is performed to yield a p-value for the exposure value change, from which 95% CIs can be extrapolated. We observed a high variability in signature exposures within HSCs from each group and across all groups (Figure 31. B.). Furthermore, in the majority of the HSCs, most of the SNVs for each HSC contributed to signature 5 exposure, with lower levels of exposure for signatures 1, 3, 18 and 40. Signatures 2 and 7c had minimal contribution across the HSCs. Due to the low mutation burden within individual HSCs, there was a high degree of variance for signature exposures. As such we decided to perform the same analyses but rather grouping SNVs per mice and then per group, in order to increase our mutation burden per sample to hopefully decrease this observed high variability in signature exposures.

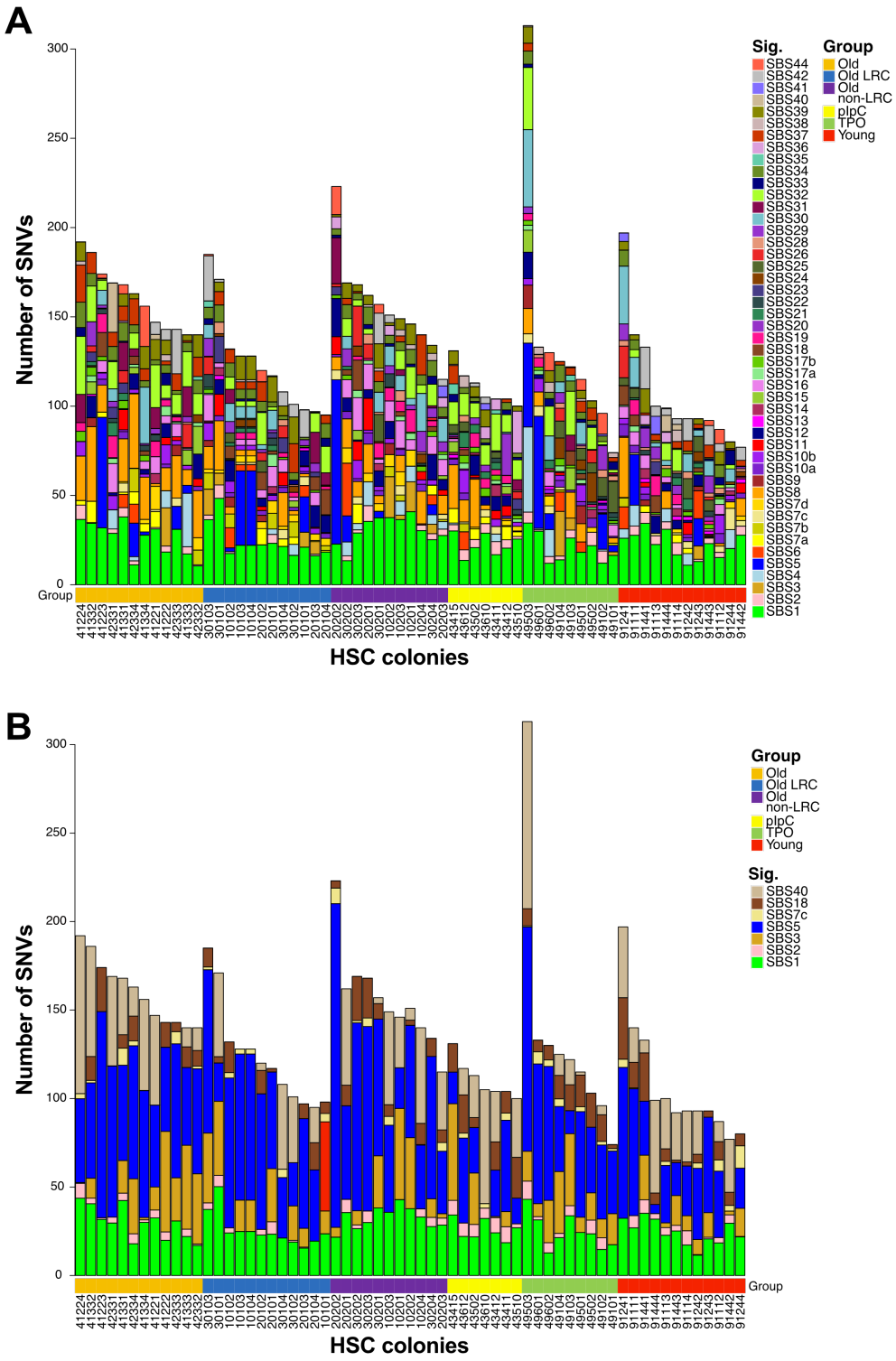


Figure 30. Mutational signatures observed per HSC colony. Absolute exposures of **(A)** all mutation signatures, and **(B)** signatures which passed a signature-specific cutoff for each HSC colony. Legend indicates HSC colony groups. Values indicated along the x-axis are identifiers for each HSC colony. Sig = signature; SBS = single base substitution.

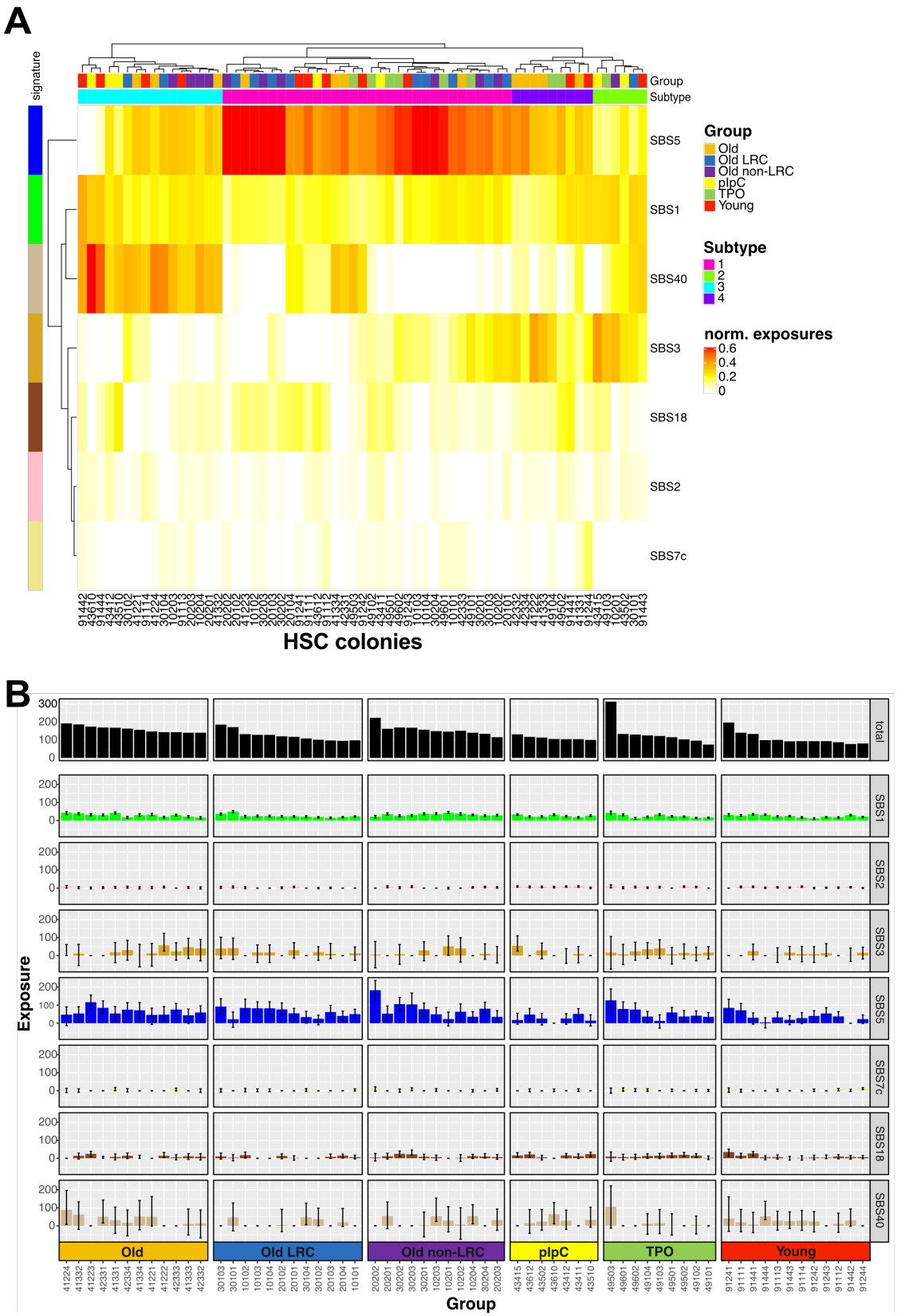


Figure 31. Distribution of identified mutation signatures across all HSCs. (A) Clustering of HSCs and signatures based on the relative exposures of identified signatures across all HSCs. Legend indicates groups, exposure level and

clustering subtypes. Values indicated along the x-axis are identifiers for each HSC colony. Sig = signature; SBS = single base substitution; norm = normalized. **(B)** Confidence intervals of signature exposures across all HSCs. Error bars indicate 95% CIs.

6.2.2. Mutational signature analysis of individual mice

Here, we grouped SNVs identified for each mouse and performed the above-described mutational signature analysis again. That is, SNVs identified for each HSC were grouped according to their corresponding mouse. The exception was for LRC and non-LRC HSCs. Although LRC and non-LRC HSCs were sequenced from the same three mice, we grouped the SNVs from LRC HSCs for one mouse together, and SNVs from non-LRC HSCs from the same mouse together, and repeated this for each mouse. Therefore, instead of us analyzing three mice/samples from the label-retention model, we essentially analyzed six samples; three LRC samples and three non-LRC samples. As such, we analyzed SNVs from six groups, three mice per group.

As seen before when assessing mutation signatures per HSC, we initially observed no distinct patterns among mice from the same groups (Figure 32. A.). After applying signature-specific cutoffs, we observed six mutation signatures whose exposures were above the pre-determined thresholds (Figure 32. B.). Six of these signatures were the same as those identified before in the HSC mutation signature analysis. These are signatures 1, 2, 3, 5, 18 and 40. Signature 7c was no longer viewed when grouping SNVs per mouse, therefore this signature was likely a false positive due to the decreased number of SNVs per HSC. When we performed a cluster analysis, we observed an interesting, albeit surprising pattern (Figure 33. A.). From this analysis, we observed four subtypes in our data. Subtype 1 was comprised of mice which had high signature 5 exposure. Signature 5 has an unknown mutational process but this signature has been reported to act in a "clock-like" manner in humans. That is, the exposure of this signature seems to correlate with the age of individual (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/SBS5.tt). This subtype was comprised of two old mice, one young mouse, two TPO-treated mice, two non-LRC and two LRC mice. Considering that the majority of these mice are ~22 months and older, with the exception of the one young mouse and two-treated mice, the high exposure of this signature in these samples appears to correlate with the concept of it acting as an ageing/clock signature.

Subtype 2 contains four mice, one non-LRC, one LRC, one old and one young. This subtype seems to be characterized by a moderate exposure, ~0.4, of signature 5 and 40. The mutational process of signature 40 is also unknown but has been correlated with the age of patients for some types of human cancers (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/SBS40.tt). Subtype 3 is comprised of three mice with the lowest exposures for signature 5, moderate to high exposure of signature 40 and the highest exposure of signature 1 of all the mice (>0.2). Furthermore, these three mice seem to cluster according to similar exposures of signatures 18 and 2. Signature 1 occurs via enzymatic or spontaneous deamination of 5-methylcytosine to thymine which results in G:T mismatches in double-stranded DNA (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/SBS1.tt). Failure to repair this DNA damage results in C>T substitution fixation. Like signatures 5 and 40, signature 1 correlates with the age of an individual. Signature 18 has been associated with damage by reactive oxygen species, while signature 2 has been attributed to activity of APOBEC cytidine deaminases in human cancer (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/SBS18.tt, https://cancer.sanger.ac.uk/cosmic/signatures/SBS/SBS2.tt). The activation of these deaminases in cancer has been linked to previous viral infection, tissue inflammation or retrotransposon jumping. Interestingly, this subtype is comprised of three ~8 month old mice, two of which were pIpC-treated mice which could explain the higher exposure of signature 2 (the other was one young mouse). Lastly, subtype 4 was comprised of two mice, one pIpC- and one TPO-treated, which had a moderate signature 5 exposure (>0.3) and the highest exposures of signature 3 across all mice (>0.3). Signature 3 has been associated with defective homologous recombination-based (HR) DNA damage repair which often exhibits as small INDELs, genome rearrangements, as well as SNVs (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/SBS3.tt). However, both of these mice had the lowest INDEL burden compared to other pIpC- and TPO-treated mice and had comparable INDEL numbers to young untreated mice (Figures 26. A. and C., Mouse 1; Figure 24. B.).

Lastly, we observed a high variability in signature exposures between mice from the same groups (Figure 33. B.), especially with regards to signatures 3, 18 and 40 where we observed that some

mice from within each group had minimal to no exposure of these signatures. This is the likely reason for the surprising cluster patterns we observed, where one or two mice from one group would cluster away from each other. This was unexpected due to mice from the same group being housed together in the same conditions.

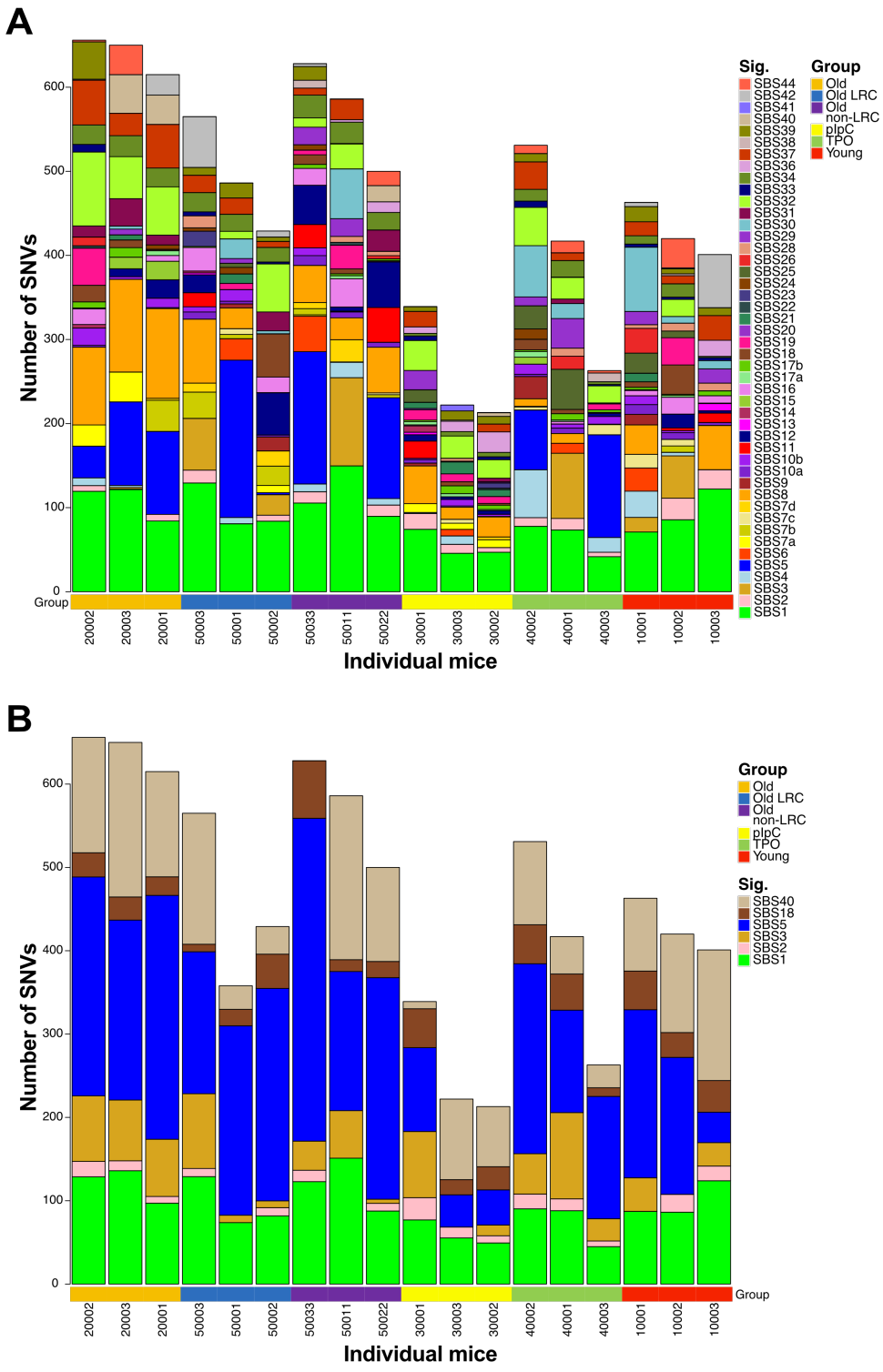


Figure 32. Mutational signatures observed per mouse. Absolute exposures of **(A)** all mutation signatures, and **(B)** signatures which passed a signature-specific cutoff for each mouse. Legend indicates groups. Values indicated along the x-axis are identifiers for each mouse. Sig = signature; SBS = single base substitution.

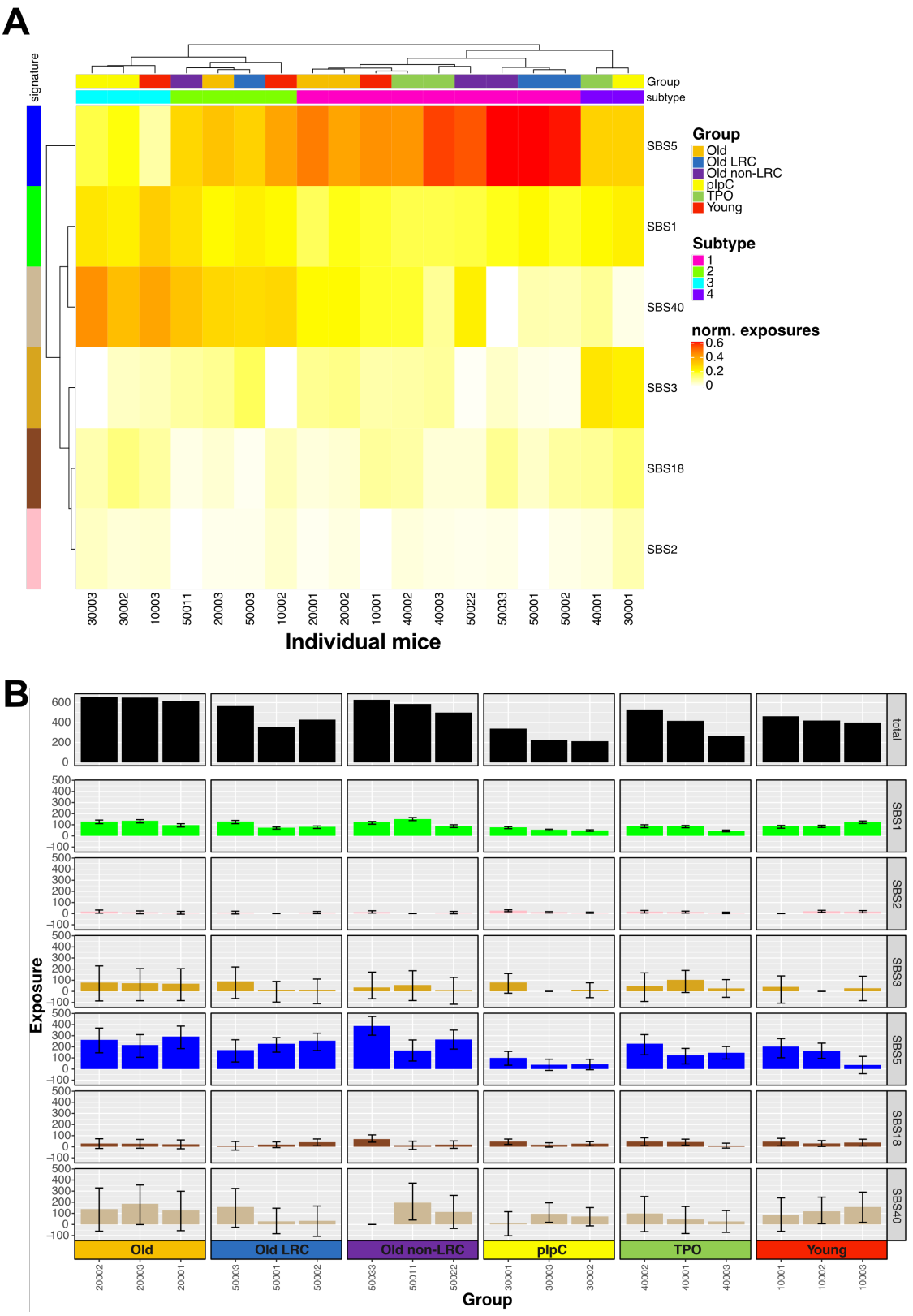


Figure 33. Distribution of identified mutation signatures across all mice. (A) Clustering of mice and signatures based on the relative exposures of identified signatures across all mice. Legend indicates groups, exposure level and

clustering subtypes. Values indicated along the x-axis are identifiers for each mouse. Sig = signature; SBS = single base substitution, norm = normalized. **(B)** Confidence intervals of signature exposures across all mice. Error bars indicate 95% CIs.

6.2.3. Mutational signature analysis of individual groups

Finally, we performed the mutational analysis on each group. That is, we grouped SNVs from HSCs from mice of the same group together and assessed the differences in mutation signatures across the groups. Again, we separated the LRC and non-LRC HSCs into separate groups. It must be noted that we had sequenced fewer HSC colonies from pIpC- and TPO-treated mice due to DNA input limitations. Therefore the amount of SNVs for each of these groups was lower than that seen in other groups. That being said, although we only sequenced 9 HSC colonies from TPO-treated mice, we had a total SNV burden similar to that of the young group where we sequenced 12 HSCs (total SNVs for TPO=1206; total SNVs for young=1278; Figure 34. A.). From our initial mutation signature analysis, we observed a large contribution of signature 5 within the old, LRC, non-LRC and TPO-treated groups (Figure 34. A.). After applying signature-specific cutoffs and rerunning the analysis, we observed that signature 5 was the main contributor to the mutational spectra of all groups, except for the young and pIpC-treated groups where signatures 1, 5 and 40 appeared to have equal contributions (Figure 34. B.). Furthermore, as with the mutational signature analysis of mice above, signatures 2, 3 and 18 were also present.

After clustering the groups according to exposure to these six signatures, we again observed four subtypes (Figure 35. A.). Subtype 1 contained SNVs from the young group and was composed of moderate to high exposure of signature 40, as well as lower signature 5 exposure compared to the old, non-LRC, LRC and TPO-treated groups. Interestingly, although distinct subtypes (young = subtype 1; pIpC = subtype 3), the young and pIpC-treated groups clustered together under the same node. This was due to their lower exposure of signature 5, and higher exposures of signatures 1 and 40, which were the highest observed across all groups. The pIpC-treated group was classified as a distinct subtype from the young group due to its higher exposure of signatures 3, 18 and 2, which have been associated with defective HR DNA damage repair, DNA damage by reactive oxygen species, and APOBEC cytidine deaminase activity linked to previous viral

infection, tissue inflammation and retrotransposon jumping. The TPO-treated group also exhibited higher exposures of signatures 3, 18 and 2 which indicates a similar underlying mutational process within these two treatment groups. However, that is where the similarities stopped as the TPO-treated group clustered with the old group, which was classified by a high signature 5 exposure. Although, the old and TPO-treated groups had similar exposure levels to signatures 1 and 5, they differed with regards to the four remaining signatures. Namely, the old group had lower levels of signatures 3, 18 and 2 than the TPO-treated group, but a higher exposure of signature 40. Higher exposures of signatures 5 and 40 indicate that mutations present within this group are correlated with age. Lastly, subtype 4 contained the LRC and non-LRC groups and exhibited the highest exposure of signature 5 compared to all other groups. There was minimal difference in exposures across all signatures between these two groups which indicates that although we observed a higher SNV burden within non-LRC HSCs compared to LRC HSCs, the underlying mutational processes of these groups are the same. This makes sense considering these HSCs are from the same mice. Therefore difference in mutation burden between LRC and non-LRC HSCs was likely due to replication alone and not differing mutational processes.

Not surprisingly, we observed less variance in exposures per signature across all groups, with the exception of signatures 3 and 40 (Figure 35. B.). It is unsure why this occurred but was likely due to the low mutation burden across all HSCs. The limited number of SNVs present within HSCs limits the capacity of this analysis to accurately assess mutational signatures present per HSC and mouse. As such, combining SNVs according to their experimental groups to perform this supervised mutational signature analysis, was the best option for performing a clustering analysis of biological relevance. However, more sequencing of HSCs per group would be required to resolve the observed variances.

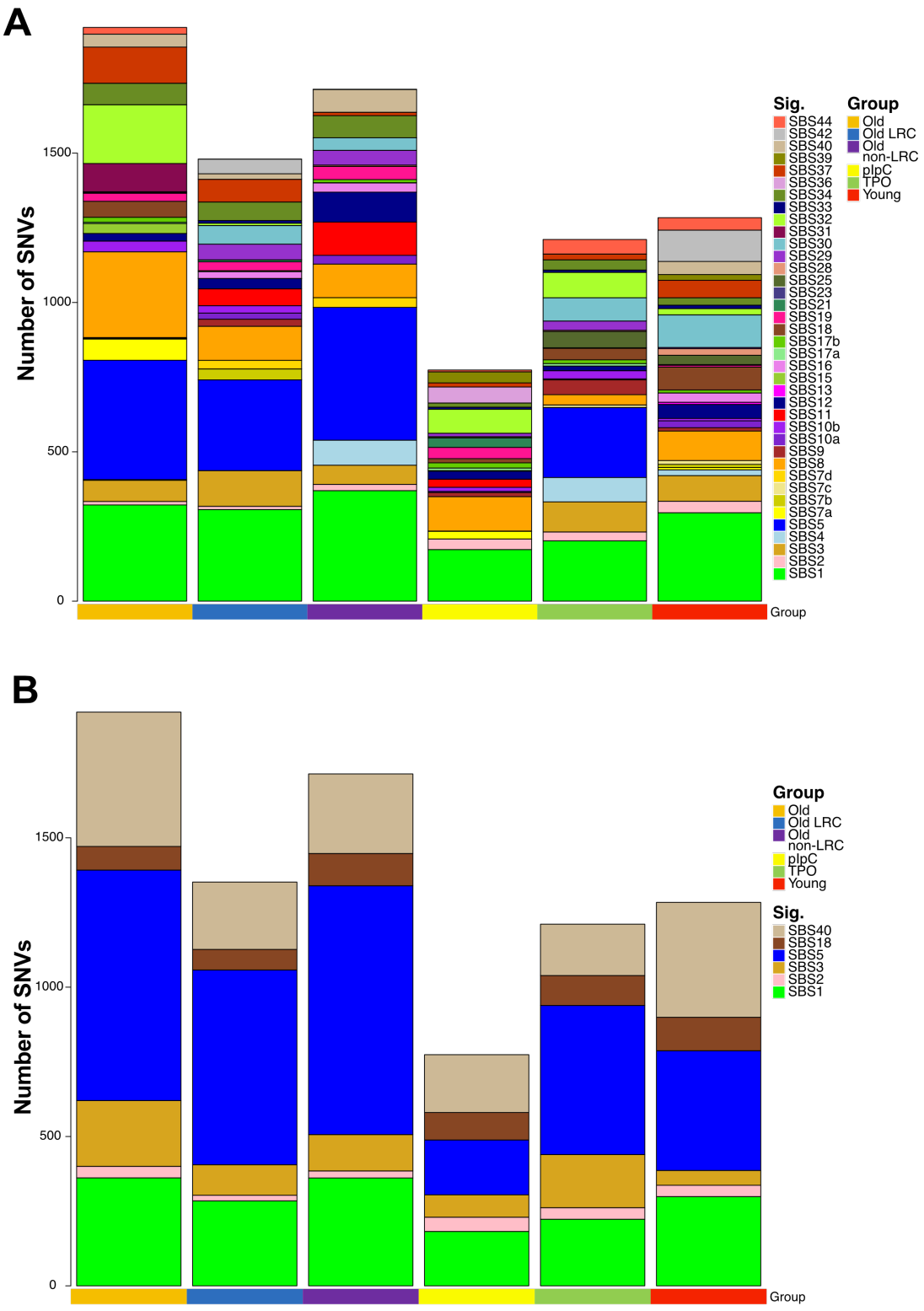


Figure 34. Mutational signatures observed per group. Absolute exposures of **(A)** all mutation signatures, and **(B)** signatures which passed a signature-specific cutoff for each group. Legend indicates groups. Sig = signature; SBS = single base substitution.

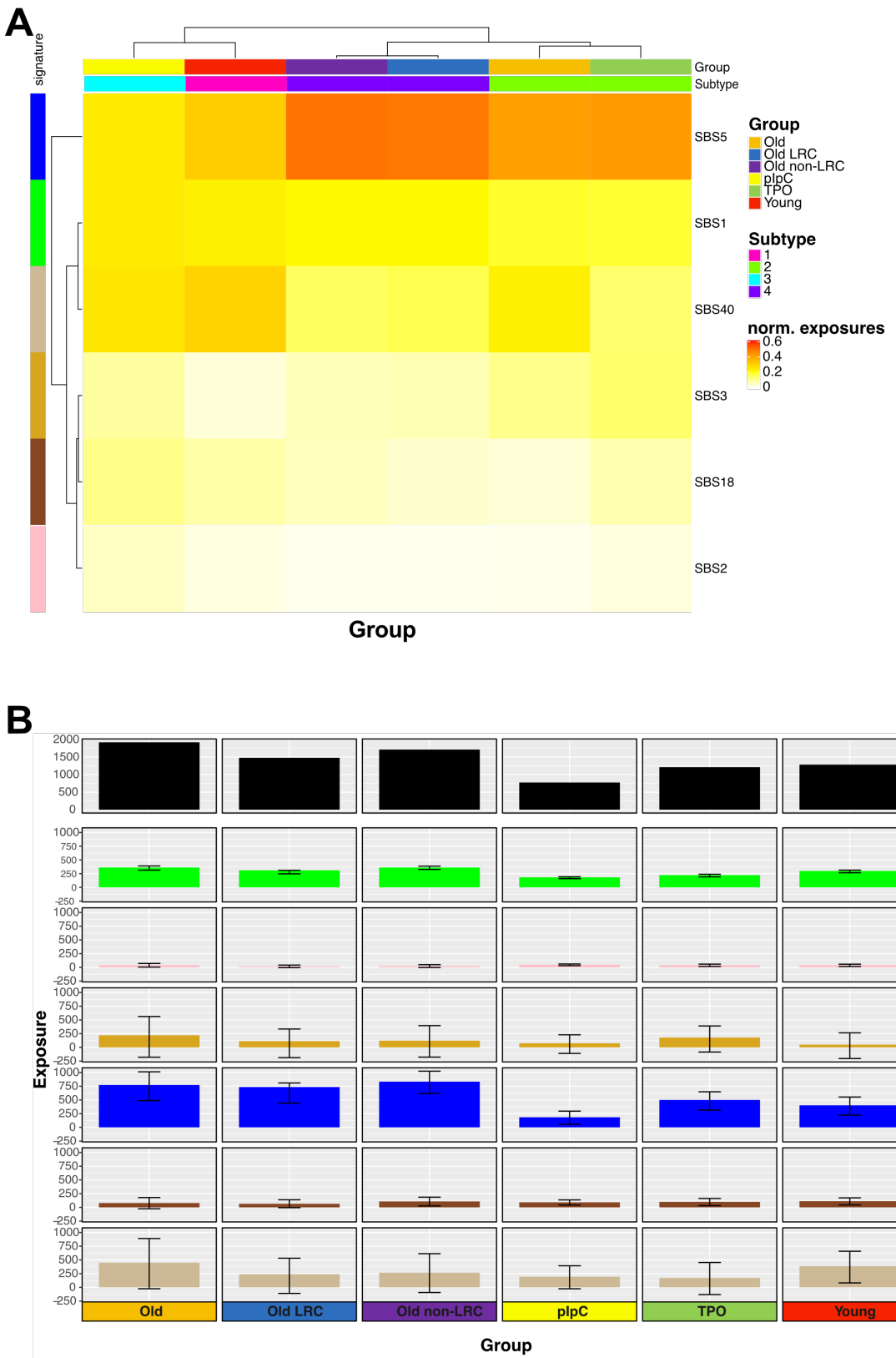


Figure 35. Distribution of identified mutation signatures across all groups. (A) Clustering of groups and signatures based on the relative exposures of identified signatures across all groups. Legend indicates groups, exposure level

and clustering subtypes. Sig = signature; SBS = single base substitution; norm = normalized. **(B)** Confidence intervals of signature exposures across all groups. Error bars indicate 95% CIs.

7. Telomere length and mitochondrial DNA copy number analyses revealed no significant changes with ageing of HSCs

Besides assessing the quantitative and qualitative differences in murine HSC mutation burden with age, replication and stress; we additionally aimed to assess whether we could detect the presence of other ageing phenotypes using our WGS data. Two such ageing phenotypes which have been viewed in humans, include telomere shortening and the reduction of mitochondrial DNA copy number (mtDNA-CN). Using TelomereHunter[56] to assess telomere content per HSC genome, we initially observed what appeared to be a shortening of telomeres within the old group compared to the young, however this was not statistically significant ($p = 0.2685$; unpaired two-tailed t test; Figure 36. A.). Nonetheless, we did observe a significant decrease in telomere length between young and non-LRC HSCs ($p = 0.0159$; unpaired two-tailed t test), and young and LRC HSCs ($p = 0.0435$; unpaired two-tailed t test). Considering that HSCs from these two groups were isolated from ~22 month old mice, this may indicate that telomere shortening is occurring with age but we likely need to increase our sample size in order to view clear differences between the groups. No difference in telomere content was observed between non-LRC and LRC HSCs ($p = 0.5814$; unpaired two-tailed t test), demonstrating that telomere shortening with age is likely not a replication-dependent mechanism. No other significant differences in telomere content were viewed between the groups.

To determine whether there are differences in mtDNA-CN in HSCs between our analyzed groups, we performed an analysis described by Longchamps *et al.*[57]. In brief, mtDNA-CN estimates are calculated by comparing the proportion of mtDNA mapped reads, to the total number of mapped reads per HSC genome. We noted no differences in mtDNA-CN between young and old ($p = 0.9142$; unpaired two-tailed t test), and non-LRC and LRC HSCs ($p = 0.8513$; unpaired two-tailed t test; Figure 36. B.). However, we did observe an increase in the mean proportion of mtDNA-CN in pIpC HSCs compared to: young ($p = 0.0307$; unpaired two-tailed t test), old ($p = 0.0234$;

unpaired two-tailed t test), and TPO-treated HSCs (p = 0.0205; unpaired two-tailed t test). Nevertheless, it is important to note that we did observe a large variability in mtDNA-CN estimates in HSC colonies from the pIpC-treated group. Together with the fact that we sequenced fewer HSC colonies from pIpC-treated mice compared to the other analysis groups, additional pIpC HSCs need to be sequenced to accurately assess this variance.

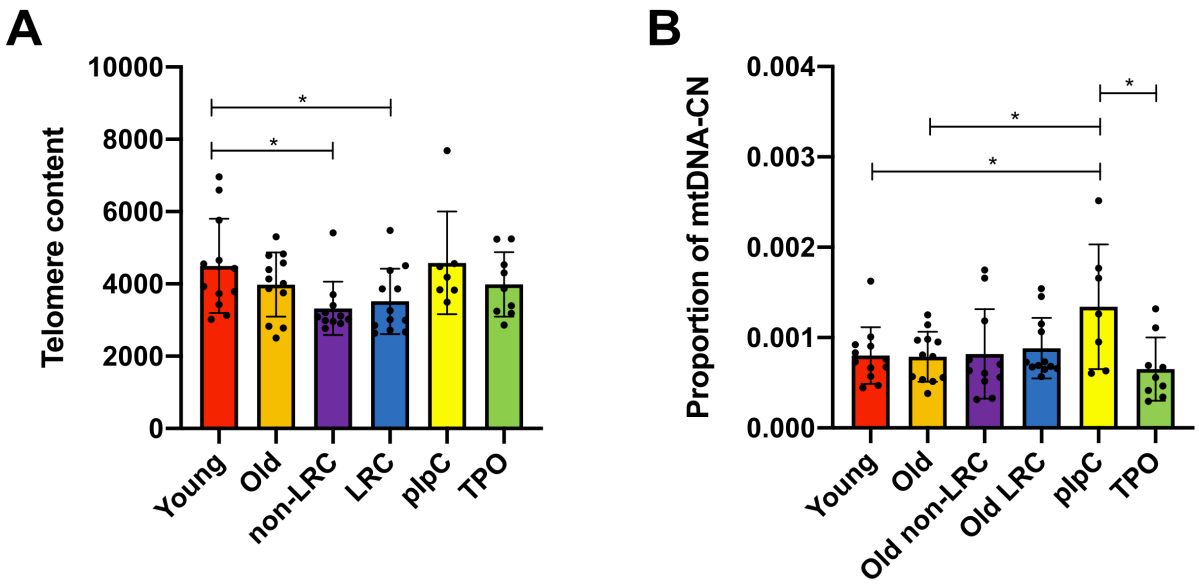


Figure 36. Age-associated telomere shortening and reduction in mitochondrial DNA copy number was not observed in murine HSCs. (A) Telomere content estimates per HSC across all analyzed groups. Each dot represents one HSC. Error bars indicate SD. Bars indicate statistically significant differences between groups (unpaired two-sided t test). **(B)** Proportion of mtDNA-CN within each HSC genome, across all groups. Each dot represents one HSC. Error bars indicate SD. Bars indicate statistically significant differences between groups (unpaired two-sided t test).

V. Discussion

Ageing of stem cells has been linked to an increase in genomic instability and a decrease in functionality[8]. Stem cells have often been proposed as the likely cells of origin of various cancers[33,58,59] and the random acquisition of mutations within ASCs is thought to have a large impact on cancer risk[33]. Furthermore, advanced ageing, progeroid syndromes have been linked to an accumulation in DNA damage[33], and individuals with these syndromes could potentially have differences in cellular mutation burden and type compared to healthy individuals[59].

The advent of next generation sequencing (NGS) technologies has provided researchers with a means of assessing the quantitative and qualitative differences in mutations within organisms. As such there has been an explosion in research focused on identifying driver mutations of cancer, but less so for normal cells. With regards to the hematopoietic system, multiple studies have identified leukemic driver mutations within HSCs[28,29], as well as candidate driver mutations for pre-malignant phenotypes like clonal hematopoiesis[25] and CHIP[30,31]. This is largely because of the fact that it is easier to assess somatic mutations within these samples due to their clonal nature. However, little focus has been placed on assessing the mutational differences within individual HSCs during normal hematopoiesis and ageing. Furthermore, it is not yet understood how the pre-malignant evolution of HSCs proceeds with time and which mediators drive this.

In this study we performed WGS on *in vitro* expanded murine HSC colonies from: (i) young and old mice, (ii) mice from a label-retention model whereby we could separate HSCs which had cycled more from those which had remained predominantly dormant throughout their lifetime, as well as (iii) mice which had been treated with agonists known to drive HSC cycling. Our goal was to identify whether there are quantitative and qualitative differences in mutations between these various groups in order to understand the possible physiologic mediators of HSC ageing and subsequent attrition.

1. Mutation burden accumulates within HSCs during ageing

Separate studies have reported an age-associated increase in DNA damage foci within murine HSCs[34,36,60], as well as an increase in SNVs within human HSPCs[23,24]. A previous study using a similar single-cell culture method as we did, observed a gradual increase in SNV burden within human HSPCs with age, at a rate of approximately 14 SNVs per year[24]. Likewise, we too observed a statistically significant increase in SNV burden within murine HSCs with age. Using a linear regression model, we were able to extrapolate an annual rate of ~40 SNVs per year. This is much higher than that of human HSCs and actually closer resembles that observed in human ASCs from more active tissues like the colon, liver and small intestine[46]. This is likely due to the shorter lifespan of mice and possibly demonstrates an overall increase in cell activity in mice compared to humans. Considering the maturational rate of mice does not correlate linearly to that of humans but is overall faster than humans[61], such an explanation seems feasible. Interestingly, we also observed a significant increase in INDEL burden within murine HSCs with age, which has not yet been viewed in human HSPCs[24]. Using a linear regression model, we extrapolated an annual INDEL rate of 5 INDELs per year, however the 95% confidence intervals spanned a large range from 0.3-9.4 INDELs, likely due to the limited number of INDELs we detected per HSC.

It must be noted that although a linear regression can be performed using only two time points, the interpretation is somewhat tricky as it cannot be accurately determined whether the change over time is truly linear, or rather exponential. As such, increasing our dataset to include more timepoints would help us in determining a true mutation rate. As part of our collaboration with the Martincorena lab at the Wellcome Trust Sanger Institute, we have done exactly that- we have sent HSC colonies from various aged mice to the Martincorena lab and together, hope to assess the true mutation rate within murine HSCs with age.

2. HSC replication history directly correlates with mutation burden accumulation during ageing

Once we had established that murine HSCs accumulate mutations with age, the obvious next step in our study was to determine which physiological sources mediate these DNA changes. It is well known that unavoidable, random mutations occur during DNA replication[33], therefore we hypothesized that replication may be a possible mechanism of mutation acquisition within HSCs. As previously mentioned, various studies have observed an increase in DNA damage response foci within old murine HSCs[9,34,36]. However, the mechanism which drives this DNA damage accumulation still remains questionable, with one study stating that this increase is due to an impairment in DNA damage repair mechanisms with age[9], while another argued that these foci actually represent replication stress[34]. An additional study hypothesized that HSCs accumulate DNA damage during quiescence and HSC replication activates the DNA repair pathways, resulting in the repair of this damage[36]. However, none of these studies took into account the proliferation status of these HSCs. Using a label-retention model, we were able to definitively identify and separate HSCs which had cycled more (non-LRC HSCs) throughout their lifetime from those which had remained predominantly dormant (LRC HSCs). From our results, we observed a significant increase in SNV and INDEL burden within HSCs which had cycled more compared to those which remained quiescent for the majority of their life. Furthermore, the SNV burden of these active HSCs was similar to that of the HSCs from our old group, illustrated by a lack of significant difference between the two groups. Conversely, the SNV burden of the predominantly dormant HSCs was similar to that of the young group and showed no significant difference. Considering the fact that HSCs from our label-retention model were isolated from ~22 month old mice and our old group of mice were ~24 months old, this highlights a direct link between replication and age-associated mutation burden. This observation was further emphasized when we assessed the annual SNV rate of our cycling and dormant HSCs. The SNV rate of cycling HSCs was very close to that of our old group at ~41 SNVs per year, while the rate of SNV acquisition within predominantly dormant HSCs was much lower than that of both the cycling and old group HSCs. Considering that this annual SNV rate within HSCs from our cycling and old groups is almost 3X

more than that viewed in humans, this could potentially indicate that murine HSCs replicate at a faster rate compared to humans, which could mediate their increased maturation rate.

Our results directly contradict the hypothesis that DNA damage accumulates within quiescent HSCs and that this damage is repaired upon replication[36]. Rather, a large proportion of mutations acquired during ageing appear to have been acquired during homeostatic replication of HSCs. Therefore demonstrating that dormancy protects the genome of HSCs. We are the first to definitively show this link between age-associated mutation burden and replication of HSCs.

3. Stress-induced HSC replication using pIpC and TPO treatment did not result in an advanced-ageing phenotype with regards to mutation acquisition

In line with our observation that active cycling of HSCs resulted in an increase in mutation acquisition, previous studies observed an increase in DNA damage response foci within HSCs when cycling was induced[38–41]. Here, these studies noted that treating mice with various non-DNA-damaging agonists which mimic endogenous and exogenous stressors, resulted in HSC activation and subsequent increase in DNA damage response foci. These agonists represented various mechanisms of stress-induction, including an inflammatory response triggered by pIpC treatment, and the stimulation of HSC proliferation for the replenishment of certain blood cell types by Thromboietin and G-CSF treatment, as well as chronic blood loss[40]. Although they observed an increase in HSC replication and subsequent accumulation of DNA damage response foci, the question still remained whether this resulted in an increase in mutation burden. Considering our observation that replication mediates mutation acquisition within HSCs, we hypothesized that stress-induced cycling of HSCs results in an increase in mutation burden and that this could be a potential mechanism of HSC mutation acquisition and ageing. However, when we induced HSC proliferation via pIpC and TPO treatment, we did not see an increase in mutation burden compared to age-matched, untreated HSCs.

A possible biological explanation for this observation could be that these HSCs are effectively repairing DNA damage, which would explain the accumulation of DNA response foci observed

after both pIpC and TPO treatment by Walter *et al.* (2015). Considering the study by Garaycoechea *et al.* (2018), who observed an increase in mutation burden upon treatment with alcohol only in the absence of effective repair mechanisms, this explanation could be possible as we have treated wild type mice with functional DNA repair mechanisms. In the future, additional sequencing of pIpC- and TPO-treated HSCs from mice absent of certain DNA repair mechanisms, would be beneficial for investigating this further. Our lab has previously demonstrated that pIpC treatment of mice with a defective Fanconi anaemia DNA repair pathway resulted in bone marrow failure, while those treated with saline were fine[40]. Therefore, it is possible that such agonists result in an increase in mutation burden only in the absence of functioning DNA repair mechanisms. Furthermore, long-term treatment with such agonists would be a valuable addition to this study. By doing this, we could investigate whether the age-associated mutation increase within HSCs is potentially mediated by the gradual impairment of certain DNA repair mechanisms with age as previously proposed[9], and whether this phenomenon can be further exacerbated by long-term stress-induced HSC cycling.

A more likely explanation is that we introduced a large, unintentional bias into our study. Due to the limitations of DNA input for our WGS library preparations, we were not able to sequence the majority of our pIpC- and TPO-treated HSC colonies as they were too small to obtain sufficient DNA. This is a phenomenon we have previously observed whereby treatment of mice with pIpC or TPO resulted in HSC colonies of decreased clonogenicity potential. That is, these HSCs appeared to be defective in their proliferation capacity and gave rise to smaller colonies with lower cell counts compared to saline-treated, age-matched mice. As such, we selected the larger colonies for sequencing and could have potentially chosen more "functional" HSCs with better proliferation capacities and excluded less-functional, possibly damaged HSCs from our analysis. Furthermore, due to the majority of these HSC colonies being too small to sequence, we sequenced fewer colonies from these groups compared to the other groups.

Although pIpC and TPO treatment have been shown to push the majority of HSCs into cycle, a portion of HSCs still remain dormant. It could be possible that we have sequenced these HSCs

which have remained predominantly quiescent and retained their potency. From our previous results from the label-retention model, we believe that quiescence protects the genome of HSCs, resulting in a lower mutation burden. If we did indeed unintentionally select HSCs which had remained predominantly dormant, this could explain why the mutation burden of these pIpC- and TPO-treated HSCs was so low. Unfortunately, it would be fairly complicated to prove this point. A more pertinent direction would be to rather focus on reducing this selection bias we have introduced. As such, we have initiated two follow-up experiments to overcome this.

Firstly, we will focus on sequencing the smaller colonies which we excluded from our analysis. Through our collaboration with the Martincorena lab and the Wellcome Trust Sanger Institute, we are looking into different library preparation protocols which would allow us to sequence smaller colonies to a similar whole-genome depth. Secondly, we have begun experiments using an apoptosis-deficient mouse model ($Bak^{-/-}$ $Bax^{f/f}$ SclCreERT2). Bak and Bax are known members of the Bcl-2 family which regulate apoptosis[62], therefore their removal would result in the absence of apoptosis within selected cells . Simply put, by knocking out apoptosis within the HSC compartment (the SCL transcription factor is only present within HSCs), regardless of the amount and type of DNA damage accrued within these HSCs, they will not die via apoptosis. Our hypothesis is that by knocking out apoptosis within HSCs, we would not only be able to culture more HSC colonies *in vitro* from these mice, but could also acquire larger HSC colonies as progeny cells from these apoptosis-deficient HSCs will not die from apoptosis during *in vitro* expansion. With regards to pIpC- and TPO-treatment, this will be important for overcoming the caveat of decreased *in vitro* clonogenicity potential of these HSCs.

4. Mutational processes within murine HSCs are stable with age and mimic those observed in human HSCs

As seen in other studies, somatic mutations are typically distributed within non-coding regions and are depleted in exonic regions[33]. This was also the case for our study when we assessed which regions of the genome contained somatic SNVs. From our results we observed that the majority of SNVs fell within non-coding regions, while very few coding variants were present

within the HSC genomes. Considering that both the murine and human genomes contain more than 90% non-coding DNA, and protein-coding regions only encompass about 5% of the genomes, this was an expected result[63]. Furthermore, we observed no difference in genomic region distribution of SNVs across our study groups. This demonstrates that there was no selection for certain genomic regions between the groups, but rather SNVs were spread randomly across the various HSC genomes. Although we identified very few exonic SNVs across the groups, we did observe a higher presence of nonsynonymous SNVs for the majority of the groups compared to synonymous SNVs. However, the differences between nonsynonymous and synonymous variant counts within groups was very low and showed no statistical relevance. Interestingly, although we sequenced fewer colonies from the TPO group and the mutation burden was not significantly different from that of HSCs from untreated, age-matched mice, we did observe an increase in exonic variants within the TPO group. This could potentially suggest that TPO-treatment is selecting for more exonic SNVs, indicating a possible impairment of certain DNA repair mechanisms following TPO-treatment. However, there was a high variability between TPO-treated HSCs and additional sequencing of more TPO colonies is required if we are to delve deeper into this.

With regards to the mutational spectra of transitions and transversions, as well as INDEL mutational spectra, we noted no differences between the HSC groups. A similar mutation profile has been observed in a previous study for both human HSPC and MPP colonies[24]. It can therefore be assumed that there is no variability in mutation spectra between these different groups because the mutational processes underlying these variants are similar. This is not surprising as they are all derived from the same cell type and would indicate that mutational processes which occur within human HSCs are similar to that of murine HSCs. This was further highlighted when we looked at the 96-trinucleotide mutational profile of SNVs from each group. Not only did we observe an uniformity in mutational profiles across the groups, we also noted that these mutational profiles looked very similar to those seen in human blood colonies and acute myeloid leukemia samples[45]. These mutational profiles were dominated by C>T and T>C transitions and have been observed in blood samples from ageing individuals with no history of hematologic

malignancies[24,45], as well as within blood samples from patients with age-related clonal hematopoiesis[27]. C>T and T>C transitions are dominant within the "clock-like" mutational signatures 1 and 5, so it was of no surprise when we observed an increase in both of these signatures within our HSC groups. Similar to the previously mentioned study assessing mutation burden within HSPC colonies from humans of different ages[24], we noted that signature 5 was the main contributor to the mutational spectra of all older groups, as well as the TPO-treated group. As such, for the most part it appears as if signature 5 increases with age within murine HSCs, which has also been noted in human HSCs. Interestingly, TPO-treated HSCs cluster with the older groups and have a higher exposure of signature 5 compared to the age-matched young and pIpC group. Taken together with the large heterogeneity we saw in mutation burden between TPO-treated HSC colonies, as well as the high incidence of exonic variants, it appears as if TPO-treatment could potentially affect the type of mutations acquired and overall mutational processes. Sequencing of additional TPO-treated HSC colonies will hopefully make this more clear. We observed a clear separation between our old groups, plus TPO, from our young groups based predominantly on a lower signature 5 exposure. pIpC-treated colonies closer resembled those from the age-matched young group, however these two groups were still represented as two distinct subtypes based on their mutational signatures. Clustering of these two groups was mostly due to their higher exposures of signatures 1 and 40 compared to the other groups. Considering both of these signatures have been correlated with age, this was an unexpected result for which we have no explanation as yet. Besides these three signatures, both the pIpC- and TPO-treated groups contained higher exposures of signatures 2, 3 and 18. These signatures have been correlated with APOBEC cytidine deaminase activity which has been linked to viral infection and tissue inflammation, defective homologous recombination (HR) DNA damage repair, and DNA damage by reactive oxygen species, respectively. One could assume that this represents a similarity in the underlying mutational processes between the two groups; however as we have highlighted before, we have introduced a large bias in our analysis of these two groups and require additional sequencing before we can make any conclusive remarks regarding stress-induced mutational mechanisms.

Beyond increasing our sample size of these two groups, in order to gain any meaningful insights regarding the mutational signatures and processes at play, we would need to increase the sample size of all of our groups. The limited number of SNVs present within HSCs limits the capacity of this analysis to accurately assess mutational signatures present across different groups. Furthermore, YAPSA performs a supervised signature analysis using mutation signatures which have been identified within human cancer samples. It could be possible that mice contain different mutational signatures than humans with regards to their trinucleotide context. Therefore we must extend this analysis to perform *de novo* mutation signature extraction. Furthermore, it would be interesting to look into INDEL mutational signatures at a later stage.

5. Differences in mutation burden among HSCs from the same groups indicates an unexplored heterogeneity

Even though we saw an overall similarity in the mutation profiles across the groups, we noted a heterogeneity within HSCs from the same mice and groups. As previously mentioned, we noted a large variability in mutation burden between TPO-treated HSC colonies. This heterogeneity was also observed within the other study groups, whereby some colonies had a much higher or lower mutation burden compared to the other HSCs from the same group. In order to assess whether this heterogeneity was a true phenomenon or rather due to differences in sequencing coverages, we normalized SNV counts per HSC according to their callable genome size. Calculating the callable genome size is highly important for downstream analysis because even if the global coverage of two samples is the same, the callable regions may differ between the two. This could be due to differences in procedures like DNA extraction, library preparation, and even sequencing run quality[46]. Therefore, to accurately assess the difference in mutation burden between two samples, the differences in callable genome size must be accounted for. This is unfortunately something which has not been taken into account in the majority of WGS studies, including similar studies utilizing clonally-expanded ASC colonies.

Following normalization, we saw no difference in our overall results and the observed heterogeneity in mutation burden within HSCs from the same group still remained. Therefore, it

appears as if this heterogeneity between HSCs from the same mice and groups is real. Previous single-cell studies have hinted at this phenomenon before[1,3], however none have shown this heterogeneity clearly. It will be an important next step in HSC and ASC research to look into this phenomenon of heterogeneity further, as this could potentially shed light on why some ASCs clonally expand and out-compete others or undergo malignant transformation.

6. Whole genome sequencing and analysis optimization is crucial for accurate somatic mutation assessment

Previous studies utilizing WGS of *in vitro* expanded clonal cultures performed little to no sequencing and analysis optimization[24,33,45,46]. As such, not only were their methods of DNA extraction, library preparation and sequencing different, but also their sequencing depth thresholds and downstream analysis protocols varied. Although all of these studies agreed upon applying a hard VAF cut-off of 0.25-0.3 VAF in order to filter out *in vitro* expanded variants, additional filtering parameters varied between each study. Of course, this is normal as analysis should be optimized for each study according to the dataset they have in hand, however there was a gross lack of such optimization among the majority of these studies.

We are the first of this type of study to have produced a benchmarking dataset specifically for the optimization of our sequencing and analysis parameters. By performing an extensive downsampling analysis, we were able to determine which SNV calling pipeline was best suited for our downstream analysis. Furthermore, by using a blacklist to exclude repetitive regions and focusing our filtering on read and mapping quality, we were able to filter out the majority of the artefacts present. This was demonstrated by the low rate of false positives detected at all coverages using CaVEMan and our stringent filtering parameters. Furthermore, we were able to determine the ideal sequencing coverage of 30X required for the accurate assessment of somatic mutations within HSC colonies. At 30X we were able to call the majority of true positive variants (~75%), while only including less than 10% of false positives.

An important note is that we did not view much of a difference when we compared overlapping SNV calls from both MuTect and CaVEMan, to SNVs from CaVEMan only. Not only did this illustrate that our optimized analysis was specific enough to use CaVEMan alone, but also demonstrated the fact that different SNV pipelines are specific for different types of datasets and that this needs to be taken into account before performing your analysis. For example, although MuTect was not sensitive and specific enough for our study, a previous benchmarking study using WGS data from tumor samples found that MuTect was highly specific for the detection of somatic SNVs within these cancer samples[50]. This highlights the need to standardize analysis pipelines to deal with different types of artefacts or better yet, to perform more artefact detection studies in order to develop extensive WGS artefact blacklists which can be used by everyone.

7. Limitations and future directions

One of the major requirements of this study was that single-sorted HSCs must self-renew and remain undifferentiated in order to give rise to large enough colonies for WGS[46]. Unfortunately this was not the case with many of our HSC colonies, especially those derived from our pIpC- and TPO-treated mice. Many of these cells either did not tolerate the *in vitro* expansion or lacked the potential to proliferate and give rise to colonies. As such, we cannot perform this analysis on a lot of cells, reducing the comprehensiveness of this study. Furthermore, previous studies utilizing such a clonal culture expansion method have noted the requirement of a large concentration of input DNA for the accurate mutation analysis of the original cell[33,46]. Although we were able to successfully optimize our method for much smaller DNA inputs than previous methods[46], this was still a limiting factor in our study. With regards to the pIpC- and TPO-treated colonies, this restriction introduced a large bias in our analysis. Furthermore, we are not able to validate our detected SNVs using targeted resequencing as all of the DNA extracted was used for library preparation for the majority of our colonies. An additional caveat of this clonal culture method is the possibility of sample contamination. As long as the HSC colony is clonal, the analysis presented here can accurately determine the mutation burden of the original HSC. That being said, we analyzed sequenced HSCs which appeared to be contaminated i.e. mutation burden was several times larger than the rest of the HSC colonies. These samples had to be discarded from

our analysis and although this did not occur often, this poses a limitation of the method as non-clonal samples can only be identified after sequencing.

As previously mentioned, in order to overcome this caveat of small colony sizes, we have introduced an apoptosis-deficient mouse model in the hope that knocking out apoptosis rescues this phenotype of small HSC colonies observed in pIpC- and TPO-treated mice. Furthermore, we plan to optimize the library preparation step of our WGS protocol to generate libraries from DNA inputs lower than 25ng. At the rate that sequencing technologies are developing, it is likely that WGS will become cheaper, allowing us to rectify some of our limitations and discrepancies between samples by increasing out sample set. As such, we are currently looking into transitioning over to the NovaSeq platform which will allow us to sequence more HSC colonies to 30X at a fraction of our current costs.

Besides increasing the number of HSC colonies sequenced per group, it will also be important to extend this study further by assessing the mutation spectrum of HSCs from mice which have been treated with different stressors, as well as mice from clinically relevant mouse models. We already have HSC colonies from a mouse infected with *M. avium* in hand and plan on getting more from a collaborator in the future. Furthermore it would be interesting to treat mice with cytotoxic stressors and chemotherapies with known mutagenic potential, as well as assess the mutation burden within HSCs from mice harboring known age-associated, leukemic driver mutations like DNMT3A and TET2[64]. The quantitative and qualitative assessment of mutations within these mice can help us understand the mutagenic effects that leukemic driver mutations, cancer therapies and a combination of both can have on the mutational spectra of an individual. Additionally we plan to sequence *in-vivo* expanded HSC colonies in order to determine whether *in vitro* expansion results in different types of mutations than those gained within a mouse during HSC expansion under normal hematopoiesis[42]. In our lab, we have performed single-cell HSC transplants using HSCs from label-retaining mice. These HSCs contain a label which is passed on to their progeny cells during division. Therefore, by transplanting single, label-retaining HSCs into a wild type mouse, we are able to identify and sort all progeny cells which originated from that

single HSC. This is essentially a clonal colony similar to that of our *in vitro* expanded colonies but without introducing potential *in vitro* expansion-specific artefacts. Not only will this allow us to strengthen our blacklist by identifying and adding *in vitro* expanded artefacts, but this will also be of some clinical relevance as various transplant methods require prior *ex-vivo* expansion before transplantation and it is not yet known how culture expansion can affect the genomic integrity of cells.

Lastly, we now have a large dataset of high-quality WGS data from various HSC colonies in hand. It will be important to extend our analyses further to look at larger genetic aberrations like structural and copy number variants, as well as assess whether the exonic variants we observed were present within any known cancer driver genes. Additionally, although we did not observe a difference in mtDNA-CN across the various groups except the pIpC group, assessment of mitochondrial DNA mutations will be an important analysis as these are known to contribute to ageing and age-associated diseases[16,17]. Furthermore, we did not observe a statistically significant decrease in telomere length with age, although we did observe a decrease in telomere length between the young and old LRC groups. This suggests that telomere shortening with age is occurring but we likely need to increase our sample size to observe clear differences between the groups. Surprisingly, when comparing active and predominantly dormant HSCs, we observed no difference in telomere length with increasing replication history. This was an unexpected result as telomere shortening is known to occur during cellular division and studies have observed a reduction in telomere length following increased HSC cycling post-transplantation[65]. It will be important to investigate this further by not only increasing our overall sample size, but also utilizing more sensitive methods of quantifying telomere length, as assessing telomere length using WGS data is a fairly new and developing field. Extending our downstream analyses, as well as our sample number and groups, will allow us to delve deeper into the potential mechanisms and drivers of age-associated mutation acquisition and attrition of HSCs.

8. Summary and working model

We have successfully optimized our experimental pipeline in order to accurately assess the quantitative and qualitative differences in somatic mutations between HSCs of different ages, replication history and treated with different stressors. Our method of clonally expanding single HSCs in culture has allowed us the opportunity to assess the mutational spectra of HSCs across the entire genome using WGS, an advancement which is not yet possible with single-cell sequencing technologies. By initially generating a benchmarking dataset, we were able to accurately assess the true and false positive rates across various coverages and not only identified the optimal coverage for our analysis, but also the necessary filtering steps for the accurate detection of true somatic variants within our HSC colonies. Therefore, we now have a tailored WGS analysis in hand which is not only highly sensitive but also highly specific for the detection of somatic mutation events within individual HSCs. We are the first of this type of study to have produced such a benchmarking dataset specifically for the prior optimization of our sequencing and analysis parameters.

Using our optimized analyses, we were able to definitively show that, like with human HSCs, there is a gradual age-associated increase in mutation accumulation within murine HSCs under homeostatic conditions. Furthermore, we demonstrated that this age-associated increase in mutation burden correlates with the replication history of the HSC. Namely, HSCs which have replicated more throughout their lifetime have a higher mutation burden than those which remain predominantly quiescent. These results clearly contradict the classical model of hematopoiesis which shows HSCs as remaining largely quiescent throughout their lifetime, rather relying on ST-HSCs and progenitors to replenish and maintain the hematopoietic system[1]. That is, if HSCs remained quiescent for the majority of their life, we would not observe this age-associated increase in mutation burden which appears to correlate with proliferative history. These results rather indicate that over time HSCs replicate and divide, resulting in an accumulation of DNA damage. With every division, it is probable that not all of this damage is effectively repaired by DNA repair mechanisms and is subsequently passed on to progeny cells and accumulates within the HSCs with age.

When we induced HSC replication using stress agonists pIpC and TPO, we did not observe an increase in mutation burden. Considering our results which demonstrated that increased HSC replication correlates with an increase in mutation burden, this was an unexpected result. However, we believe we have introduced a large selection bias within these cohorts due to our DNA input limits and the decreased clonogenicity potential of pIpC- and TPO-treated HSCs. Therefore, future work will focus on rectifying this caveat and increasing our sample set to include more clinically relevant phenotypes.

To conclude, we have developed and validated our analysis pipeline and are the first of this type of study to perform such an extensive benchmarking optimization analysis. Using our optimized pipeline, we are the first to definitively observe a causal link between replication history and age-associated mutation acquisition within HSCs. As such, we have clearly demonstrated the power of such a study to investigate ageing and the physiologic mediators which drive age-associated genomic instability. Going forward, as researchers delve into the process of mutation acquisition more, performing such an analysis will be critical in the setting of understanding phenotypes like malignant transformation, as well as looking deeper at the molecular heterogeneity in disease settings.

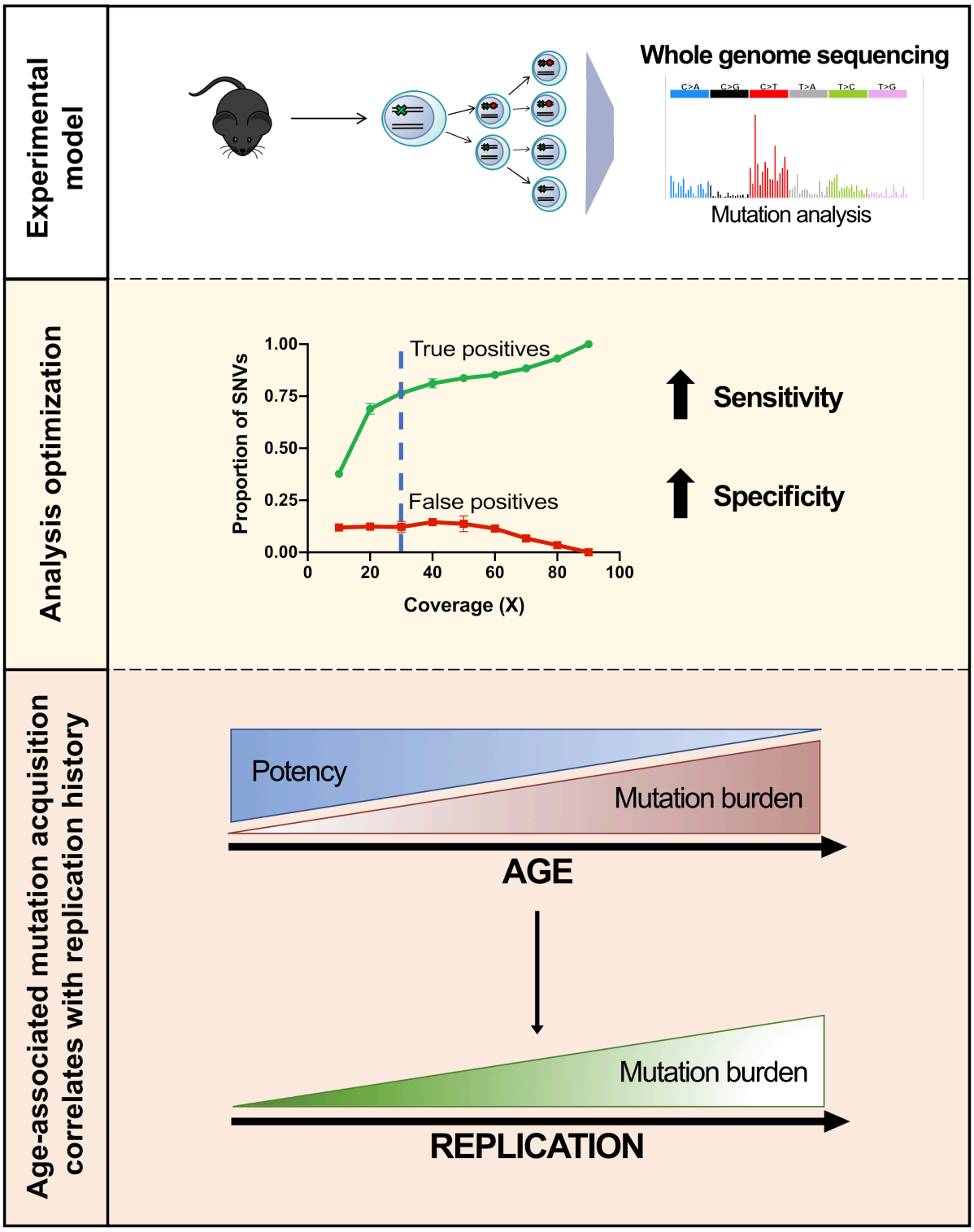


Figure 37. Graphical overview of our optimized experimental and analysis pipelines to accurately assess the quantitative and qualitative differences in mutation burden between HSCs.

VI. Materials and methods

1. Mouse models and collection of primary material

1.1. Mouse lines

Mice used in this project were housed in pathogen-free, individually ventilated cages within the DKFZ animal facility. All experimental procedures involving said mice, were approved by the Animal Care and Use Committees of the German "Regierungspräsidium Karlsruhe für Tierschutz und Arzneimittelüberwachung". We obtained C57BL/6J mice from Harlan Laboratories, Charles River Laboratories, or Janvier Laboratories and these mice were utilized in the ageing, pIpc- and TPO-treatment experiments (all C57BL/6J mice used for this project were female). For the label-retention assay, scltTa H2B GFP mice were bred on a C57BL/6J background (all scltTa H2B GFP mice used in this project were male). The genotype of scltttTa H2B GFP was confirmed with an established PCR protocol. At the time of starting each experiment, mice were between 8-16 weeks of age.

1.2. Organ harvest and isolation of lineage negative bone marrow cells

At the endpoint for each mouse experiment, mice were sacrificed by cervical dislocation and spine, tibias, femurs and hips were collected from each mouse. The flesh from the bones was removed as much as possible, as well as the spinal cord, with careful note to keep bones from individual mice separate. Additionally, a portion (~1cm in length) of tail from each mouse was harvested and stored at -20°C for later DNA extraction and sequencing as a germline control for each mouse. First, both femurs for each mouse were flushed in 2mLs of 2% FCS (Gibco) PBS (Sigma-Aldrich) and a 30μL aliquot was taken for measuring cell numbers using a Hemavet 950 (Drew Scientific) or Scil Vet ABC+ (scil animal care company GmbH). The remaining bones were then crushed by mortar and pestle in a solution of Iscove's modified Dulbecco's medium (IMDM, Life Technologies), followed by filtering of this cell suspension through a 40μm filter (Greiner bio-one). This was then pooled with the remaining cell suspension from the flushed femurs and spun down at 1500rpm for 5 min at 4°C. The cell pellet was resuspended in 4mL 2% FCS PBS and the total cell count was measured using a Hemavet or Scil Vet. A lineage depletion antibody cocktail

(see Table 1; all antibodies from BD Bioscience) was then added to the cells at a concentration of 112.6μL per 1 x 10^8 cells and incubated on a roller at 4°C for 30 min. Cells were washed with 50mL of ice-cold PBS and centrifuged at 1500rpm for 5 min at 4°C post-incubation, and the cell pellet was resuspended in 3mLs of 2% FCS PBS for subsequent lineage depletion. In brief, 1mL of pre-washed Dynabeads suspension (Thermo Fisher Scientific) was added to the cell suspension and incubated for 45 min at 4°C on a roller. The bead and cell suspension mixture was placed in a Dynal magnet (Thermo Fisher Scientific), allowing lineage positive cells which were bound by antibodies within the lineage depletion mix and subsequently the Dynabeads, were bound to the side of the tube by the magnet and the remaining cell suspension containing lineage negative bone marrow cells was harvested into a fresh tube for an additional separation step in the Dynal magnet. The beads in the initial tube were resuspended in 7mLs of ice-cold 2% FCS PBS and placed in the magnet for another round of depletion. This step was repeated one more time and the three cell suspensions, containing lineage negative cells, were pooled and spun down. The cell pellet was resuspended in 1mL of 2% FCS PBS, lineage negative cell count was measured using a Hemavet or Scil Vet, and the remaining cell suspension was kept on ice for subsequent processing.

Table 1. Lineage depletion antibody mix*.

Antibody	**Dilution**
CD5 (Ly-1) Biotin IgG2a, k	1:200
B220 (CD45R) IgG2a, k	1:300
Mac-1 (CD11b), IgG2b, k	1:320
CD8a (Ly-2), IgG2a, k	1:200
Gr-1 (Ly-6G), IgG2b, k	1:350
Ter-119 (Ly-76), IgG2b, k	1:320

* all antibodies are from BD Bioscience.

2. Treatments to induce HSC cycling

2.1. Label-retention assay

For the label-retaining cell (LRC) assays, 8-week-old sclttTa H2B GFP mice were treated with doxycycline citrate (Sigma) which was supplemented in the mice' drinking water at a concentration of 2mg/mL, along with 20mg/ml of sucrose. Experimental mice were fed the doxycycline-supplemented water for an 18-month label chase period.

2.2. Polyinosinic:polycytidylic acid treatment

In order to mimic an inflammatory response, mice were injected intraperitoneally (i.p.) with 5mg/Kg of polyinosinic:polycytidylic acid (pIpC, InvivoGen), which had been resuspended in sterile saline according to the manufacturer's instructions. The treatment regime is illustrated in Figure 23. Mice were treated with three rounds of pIpC. One round of treatment encompasses two injections with a prescribed dosage of agonist per week for four weeks, followed by a four week recovery stage[40]. At the end point of the experiment, pIpC-treated mice were ~8 months old.

2.3. Thrombopoietin treatment

Treatment of mice with Thrombopoietin (TPO, PeproTech) was carried out using the same regime as described above for pIpC, however TPO dosage was 200μg/Kg as recommended in Walter *et al.*, 2015. TPO was dissolved in sterile saline as per the manufacturer's instructions and diluted to the desired concentration before treatment. Mice were treated with three rounds of TPO treatment and at the end point of the experiment, these mice were ~8 months old.

3. Methodology to assess mutation burden at the hematopoietic stem cell level

3.1. Fluorescence activated cell sorting of LT-HSCs

The lineage depleted cell suspension was stained at a density of 1 x 10^7 cells/mL with the LT-HSC antibody mix (diluted in 2% FCS PBS) described in Table 2 and incubated on a roller at 4°C for 30

min. The cells were then washed with 2% FCS PBS, spun down at 1500rpm for 5 min at 4°C, and the cell pellet was resuspended to a concentration of 1 x 10^7 cells/mL in 2% FCS PBS and filtered into a FACS tube using a 40μm strainer. Single LT-HSCs were sorted using either the Aria or Fusion (both BD Bioscience) fluorescence activated cell sorters (FACS) at ~2500 events/sec using a 100μm nozzle. A gating scheme for sorting LT-HSCs can be found in the supplemental information (supplementary Figure 1). LT-HSCs were sorted for Lin^-, $sca\text{-}1^+$, $c\text{-}kit^+$, $CD150^+$, $CD48^-$, $CD34^-$.

Table 2. LT-HSC antibody mix*.

Antibody	Fluorochrome
Lineage: CD5, CD8a, CD11b, B220, Gr-1, Ter119, Streptavidin	PE-Cy7
cKit/CD117	BV711
Sca1	APC-Cy7
CD150	PE-Cy5
CD48	PE
CD34	FITC

* all antibodies are from BD Bioscience.

3.2. *In vitro* culture of HSC colonies

Single LT-HSCs were sorted into 96-well round bottom plates (Sigma-Aldrich) containing 150μL per well of expansion medium comprised of serum-free medium (StemSpan SFEM) (STEMCELL Technologies), Penicillin-Streptomycin (Sigma-Aldrich), L-Glutamine (Gibco) and the necessary recombinant murine (rm) cytokines which facilitate the growth and *in vitro* differentiation of HSCs into erythroid, megakaryocytic and myeloid cells (Table 3). These sorted single LT-HSCs were allowed to expand *in vitro* for a period of 12-14 days in an incubator under hypoxic conditions (5% O_2, 37°C, 5% CO_2), after which size of these HSC colonies were manually assessed using light microscopy (Leica DM IL LED microscope). Individual HSC colonies of an appropriate size (estimated ~500 cells or more) were then collected into separate 1.5mL Eppendorf

microtubes, spun down at 1500rpm for 3-5 min in order to discard the supernatant and the cell pellet was snap-frozen with liquid nitrogen and subsequently stored at -80°C.

Table 3. Expansion medium for single LT-HSC colony growth*.

StemSpan serum Free Expansion Medium (SFEM)			1 mL
	Final concentration	Stock concentration	Volume (μL)
penicillin/streptomycin	1%	-	10
L-Glutamine	1%	-	10
Flt3 ligand	10 ng/mL	100 ng/μL	0.1
SCF	50 ng/mL	100 ng/μL	0.5
TPO	10 ng/mL	100 ng/μL	0.1
IL-3	5 ng/mL	100 ng/μL	0.05
IL-11	10 ng/mL	100 ng/μL	0.1
EPO	0.3 IU/mL	1 IU/μL	0.3
IL-7	20 ng/mL	100 ng/μL	0.22

*All cytokines from PeproTech. Pen/strep and L-glut from Sigma-Aldrich and Gibco, respectively.

4. Whole Genome Sequencing of HSC colonies

4.1. Sequencing library preparations

4.1.1. Tagmentation protocol for sequencing of benchmarking dataset

DNA was isolated from one large HSC colony from a two year old, pIpC-treated mouse (8 rounds of treatment), using the DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's instructions. DNA libraries for WGS were generated according to the tagmentation-based protocol[66]. This protocol was initially designed for whole-genome bisulfite sequencing, however we adapted it for our purposes by excluding the bisulfite conversion step. This method utilizes a hyperactive Tn5 transposase in combination with DNA adapters to simultaneously fragment and ligate adapters to the ends of the input DNA fragments. For each library, we used up to 5ng of DNA input and adapted the amount of transposase required for the amount of DNA input. Finished libraries were validated using a High Sensitivity DNA chip and BioAnalyzer (Agilent Technologies) according to the manufacturer's instructions. In order to obtain a very high sequencing depth for our benchmarking datasets, we sequenced five individually prepared single

libraries and one quadriplex library for both the HSC colony and germline control. Each library was sequenced on one lane using the Illumina HiSeq V4 platform (125bp paired-end). Bioanalyzer results from the prepared libraries for the HSC colony and germline control can be found in supplementary Figures 2 and 3, respectively.

4.1.2. TruSeq Nano protocol for sequencing of HSC clones and germline controls for mutation comparisons

DNA was isolated from the HSC colonies and corresponding germline control (tail) using the DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's instructions, with the adjustment of eluting the DNA in 60μL (two elution steps of 30μL each) resuspension buffer provided by the Illumina library prep kit. Due to the small size of the HSC colonies, 5μL of the DNA was used to measure the DNA concentration was measured with a Qubit in order to get an accurate reading. DNA libraries for WGS were generated using the Illumina TruSeq Nano DNA Library Prep kit according to the manufacturer's instructions with two minor adjustments; depending on the concentration of input DNA, the number of PCR amplification cycles was between 8-10 cycles, and the last washing step was repeated twice in order to prevent carry-over of magnetic beads in the final DNA library. Although the Illumina protocol specified that 100ng of DNA should be used as input, we were not always able to obtain this amount of DNA from our HSC colonies. As such, we optimized the sonification step (sonification of DNA carried out using the Covaris M220) of the library prep protocol to allow lower DNA inputs between 25-100ng (supplementary Figure 4). Sonicated DNA and finished DNA libraries were validated using a High Sensitivity DNA chip and BioAnalyzer (Agilent Technologies) according to the manufacturer's instructions. All HSC colonies, and corresponding controls, were sequenced (150bp paired-end, one lane per sample) using the Illumina HiSeqX Ten sequencers. Bioanalyzer results from the prepared libraries for the HSC colonies and corresponding germline controls can be found in supplementary Figures 5 - 19.

5. Whole genome sequencing analysis

5.1. Read Alignment

Sequence reads were mapped against the GRCm38mm10 reference genome using the Burrows-Wheeler Aligner v0.7.15 mapping tool[67]. Sequence reads were marked for duplicates and merged using Sambamba v0.6.5 and aligned bam files were subsequently indexed using SAMtools v0.1.19. Read alignment and quality score analyses were performed by the DKFZ Omics IT and Data Management Core Facility (ODCF).

5.2. Variant calling and filtering

Raw single nucleotide variants (SNVs) for were called using two different SNV calling pipelines; MuTect[51] v.1.1.4 and CaVEMan[52] via the cgpCaVEManWrapper. Both SNV pipelines were used in our benchmarking analysis in order to determine ideal sequencing and filtering parameters to accurately call true variants of the original single LT-HSC within the HSC colony. Following the benchmarking analysis, CaVEMan was used to call SNVs within individual HSC colony samples against their corresponding controls. SNV calling parameters and raw SNV filters for each pipeline are as follows. MuTect was run using default settings and the additional option "excludeIntervals", whereby a file which defined regions of consistently higher coverage was used as a blacklist to exclude certain regions of the genome from analysis. This blacklist was generated by Dr. Charles Imbusch at the DKFZ using the BEDTools utilities[68]. Raw SNV calls were filtered to exclude variants containing the string "REJECT", followed by filtering for variants which were covered by 15 or more reads and a minimum variant allele frequency (VAF) of 0.25 to exclude *in vitro* expanded SNVs. CaVEMan was run with the default settings using a blacklist file of regions to ignore, which was generated by the Wellcome Trust Sanger Institute using multiple "control" mouse genomes to search for repetitive and homologous regions to be excluded from analysis. Raw SNV calls from CaVEMan which were annotated with "PASS" were used and additionally filtered with the following parameters; ASRD>= 0.94, CLPM ==0, and a minimum of 0.3 VAF. In brief, ASRD accounts for mapping quality of the reads covering the variant allele i.e. a read length adjusted alignment score, and CLPM filters for sequence read quality, that is, it gives a value for how many variant supporting reads contain soft clipped bases. With our filtering

parameters, we essentially filter out any variants which have an alignment score of less that 94%, and contain any variant supporting reads which have soft clipped bases. These parameters were recommended by the Martincorena lab at the Wellcome Trust Sanger Institute for analysis of SNVs within clonal samples. For the identification of insertions and deletions (INDELS) within HSC colonies, Pindel[53] was run using singularity to execute the INDEL calling pipeline from the cgpCaVEManWrapper. Default settings for running Pindel were used and due to the clonality of our samples, we chose to filter INDELs which had a minimum VAF of 0.3 to exclude *in vitro* accumulated mutations.

5.3. Coverage downsampling for benchmarking analysis

Downsampling of coverage was performed on one deeply sequenced LT-HSC colony (initial coverage of 89X) and a deeply sequenced tail control (initial coverage of 97X) from the same mouse. Downsampling fractions were calculated to obtain samples in approximately 10X increments, that is; 10X, 20X, 30X, continuing until 89X (Table 4). Variants were called for each coverage fraction using the appropriate coverage-matched HSC:control downsampled BAM and indexed BAM files. For downsampling, the Picard v1.61 (http://broadinstitute.github.io/picard/) DownsampleSAM tool was used, adjusting for differing coverage fractions within the "PROBABILITY" input. Downsampling per coverage fraction was repeated in triplicate using different seeds to generate three bam files per sequencing coverage, each containing different reads which were selected at random via the Picard downsampling algorithm. Prior to variant calling, using CaVEMan, MuTect and Pindel, the downsampled bam files were sorted and indexed using Sambamba v0.6.6 with default settings[69]. Coverage of downsampled fractions was assessed using Samtools -depth option[70].

Table 4. Downsampling fractions calculated for benchmarking analysis.

Sample	Fraction to downsample by	Coverage (X)
HSC colony	0.113	10
HSC colony	0.225	20
HSC colony	0.339	30
HSC colony	0.45	40
HSC colony	0.575	50
HSC colony	0.675	60
HSC colony	0.8	70
HSC colony	0.9	80
HSC colony	0	89
Tail control	0.104	10
Tail control	0.206	20
Tail control	0.312	30
Tail control	0.412	40
Tail control	0.530	50
Tail control	0.624	60
Tail control	0.739	70
Tail control	0.832	80
Tail control	0.912	89

5.4. Normalization for mutations per actual genome covered analysis

In order to calculate the actual coverage of the genome used for calling SNVs via CaVEMan, we utilized the command-line tool Mosdepth[71] via Anaconda (linux version Anaconda3.2019.07). Calculating the callable genome was broken down into three steps. Firstly, for each HSC colony we created a bedfile of the regions which were analyzed by CaVEMan using the BEDTools "substract" option. CaVEMan generates a bedfile of the regions which were excluded from the SNV calling analysis, and by subtracting these regions from the reference genome, we could generate a file of the analyzed regions. Secondly, this bedfile was then used as the input for Mosdepth to calculate coverage from, with default settings and a critical read fragment depth threshold of 4 i.e. Mosdepth calculates the coverage of the genome by analyzing regions with atleast 3X read depth. Lastly, the intersection of the CaVEMan-analyzed regions file and the

mosdepth file of regions which had a coverage of at least 3X, was generated and this output file was used to calculate the callable genome size. With this, we were able to normalize SNV counts per HSC colony according to their callable genomes.

5.5. Mitochondrial DNA copy number analysis

Mitochondrial DNA copy number (mtDNA-CN) assessment of each HSC colony was performed according to the protocol described by Longchamps *et al.*[57], whereby mtDNA-CN was calculated per sample as the proportion of mtDNA mapped reads compared to the total number of mapped reads per genome. For this, we utilized SAMtools[70] v1.10 options "idxstats" to determine the number of mtDNA mapped reads and "flagstat" to determine the total number of mapped reads per HSC colony. Therefore, mtDNA-CN was represented as a proportion and these were compared between sample groups.

5.6. Telomere length analysis

Assessment of telomere length for each HSC colony was performed using the TelomereHunter software[56]. TelomereHunter was run using Anaconda3 v5.1.0 using the default settings.

5.7. Mutational profile and signature analysis

Supervised mutation signature profiling of each HSC colony was performed using the YAPSA (Yet Another Package for Signature Analysis) package[55] on the most updated mutational signature catalogue[44] (https://cancer.sanger.ac.uk/cosmic/signatures). Relevant mutational signatures were selected using a signature-specific probability cut-off developed within YAPSA. Cluster analysis of these signature exposures was performed using the ComplexHeatmap package[72].

5.8. Mutation annotation analysis

For each HSC colony, identified SNVs called by CaVEMan were annotated using the ANNOVAR tool[73]. ANNOVAR settings used to annotate these variants from mouse HSCs were as follows; "—buildver mm10 –protocol refGene –operation g" in order to generate gene-based annotation output for each HSC colony. Each variant was classified either as exonic, splicing, ncRNA, UTR5,

UTR3, intronic, upstream, downstream or intergenic and frequencies of each type were compared between sample groups.

6. Statistical analysis

R v4.0.0, via an RStudio Server account provided by DKFZ ODCF, and GraphPad Prism v8.4.3 were used for performing the statistical analysis and graphical representation of the results for this project, unless otherwise indicated. Unless indicated differently, the data are presented as mean values +/- standard deviation (SD). Statistically significant differences are indicated as follows: ns ($p > 0.05$), * ($0.01 < p < 0.05$), ** ($0.001 < p < 0.01$), and *** ($p < 0.001$). For comparisons of two groups, two-sided unpaired non-parametric t-tests were performed. For comparisons between more than two groups, an ordinary one-way analysis of variance (ANOVA) was performed. In order to assess whether there was a correlation between age and mutation rate for our various groups, we used a simple linear regression model.

VII. References

1. Liggett, L. A. & Sankaran, V. G. Unraveling Hematopoiesis through the Lens of Genomics. *Cell* **182**, 1384–1400 (2020).
2. Orkin, S. H. & Zon, L. I. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell* **132**, 631–644 (2008).
3. Wilson, N. K. *et al.* Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* **16**, 712–724 (2015).
4. Nestorowa, S. *et al.* A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood* **128**, e20-31 (2016).
5. Velten, L. *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nature Cell Biology* **19**, 271–281 (2017).
6. Olsson, A. *et al.* Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* **537**, 698–702 (2016).
7. Laurenti, E. & Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**, 418–426 (2018).
8. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The Hallmarks of Aging. *Cell* **153**, 1194–1217 (2013).
9. Rossi, D. J. *et al.* Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* **447**, 725–729 (2007).
10. Burtner, C. & Kennedy, B. Progeria syndromes and ageing: What is the connection? *Nature reviews. Molecular cell biology* **11**, 567–78 (2010).

11. Jaiswal, S. *et al.* Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *New England Journal of Medicine* **371**, 2488–2498 (2014).
12. Forsberg, L. A. *et al.* Age-Related Somatic Structural Changes in the Nuclear Genome of Human Blood Cells. *The American Journal of Human Genetics* **90**, 217–228 (2012).
13. Rossi, D. J., Jamieson, C. H. M. & Weissman, I. L. Stems cells and the pathways to aging and cancer. *Cell* **132**, 681–696 (2008).
14. Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).
15. Behrens, A., van Deursen, J. M., Rudolph, K. L. & Schumacher, B. Impact of genomic damage and ageing on stem cell function. *Nat Cell Biol* **16**, 201–207 (2014).
16. Park, C. B. & Larsson, N.-G. Mitochondrial DNA mutations in disease and aging. *J Cell Biol* **193**, 809–818 (2011).
17. Zhang, R., Wang, Y., Ye, K., Picard, M. & Gu, Z. Independent impacts of aging on mitochondrial DNA quantity and quality in humans. *BMC Genomics* **18**, (2017).
18. Mengel-From, J. *et al.* Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Hum Genet* **133**, 1149–1159 (2014).
19. Shammas, M. A. Telomeres, lifestyle, cancer, and aging. *Curr Opin Clin Nutr Metab Care* **14**, 28–34 (2011).
20. DNA Damage Response - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/neuroscience/dna-damage-response.

21. Chatterjee, N. & Walker, G. C. Mechanisms of DNA damage, repair and mutagenesis. *Environ Mol Mutagen* **58**, 235–263 (2017).
22. Moehrle, B. M. *et al.* Stem Cell-Specific Mechanisms Ensure Genomic Fidelity within HSCs and upon Aging of HSCs. *Cell Rep* **13**, 2412–2424 (2015).
23. Welch, J. S. *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264–278 (2012).
24. Osorio, F. G. *et al.* Somatic Mutations Reveal Lineage Relationships and Age-Related Mutagenesis in Human Hematopoiesis. *Cell Reports* **25**, 2308-2316.e4 (2018).
25. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* **371**, 2477–2487 (2014).
26. Xie, M. *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med* **20**, 1472–1478 (2014).
27. Zink, F. *et al.* Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood* **130**, 742–752 (2017).
28. Shlush, L. I. *et al.* Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**, 328–333 (2014).
29. Jan, M. *et al.* Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci Transl Med* **4**, 149ra118 (2012).
30. Steensma, D. P. *et al.* Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* **126**, 9–16 (2015).
31. Heuser, M., Thol, F. & Ganser, A. Clonal Hematopoiesis of Indeterminate Potential. *Dtsch Arztebl Int* **113**, 317–322 (2016).

32. Geiger, H., David, S., Nattamai, K. J. & Jan, V. Quantification of Genomic Mutations in Murine Hematopoietic Cells. in *Genetic Modification of Hematopoietic Stem Cells: Methods and Protocols* (ed. Baum, C.) 423–436 (Humana Press, 2009). doi:10.1007/978-1-59745-409-4_28.
33. Blokzijl, F. *et al.* Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* **538**, 260–264 (2016).
34. Flach, J. *et al.* Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
35. Mazouzi, A., Velimezi, G. & Loizou, J. I. DNA replication stress: Causes, resolution and disease. *Experimental Cell Research* **329**, 85–93 (2014).
36. Beerman, I., Seita, J., Inlay, M. A., Weissman, I. L. & Rossi, D. J. Quiescent hematopoietic stem cells accumulate DNA damage during aging that is repaired upon entry into cell cycle. *Cell Stem Cell* **15**, 37–50 (2014).
37. Bernitz, J. M., Kim, H.-S., MacArthur, B., Sieburg, H. & Moore, K. Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions. *Cell* **167**, 1296-1309.e10 (2016).
38. Essers, M. A. G. *et al.* IFNalpha activates dormant haematopoietic stem cells in vivo. *Nature* **458**, 904–908 (2009).
39. Baldridge, M. T., King, K. Y., Boles, N. C., Weksberg, D. C. & Goodell, M. A. Quiescent haematopoietic stem cells are activated by IFN-gamma in response to chronic infection. *Nature* **465**, 793–797 (2010).
40. Walter, D. *et al.* Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature* **520**, 549–552 (2015).

41. Bogeska, R. *et al.* Hematopoietic stem cells fail to regenerate following inflammatory challenge. *bioRxiv* 2020.08.01.230433 (2020) doi:10.1101/2020.08.01.230433.
42. Garaycoechea, J. I. *et al.* Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature* **553**, 171–177 (2018).
43. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
44. Alexandrov, L. B. *et al.* The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101 (2020).
45. Lee-Six, H. *et al.* Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).
46. Jager, M. *et al.* Measuring mutation accumulation in single human adult stem cells by whole-genome sequencing of organoid cultures. *Nat Protoc* **13**, 59–78 (2018).
47. Behjati, S. *et al.* Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422–425 (2014).
48. Gawad, C., Koh, W. & Quake, S. R. Single-cell genome sequencing: current state of the science. *Nature Reviews Genetics* **17**, 175–188 (2016).
49. Costello, M. *et al.* Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res* **41**, e67 (2013).
50. Alioto, T. S. *et al.* A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing. *Nat Commun* **6**, 10001 (2015).

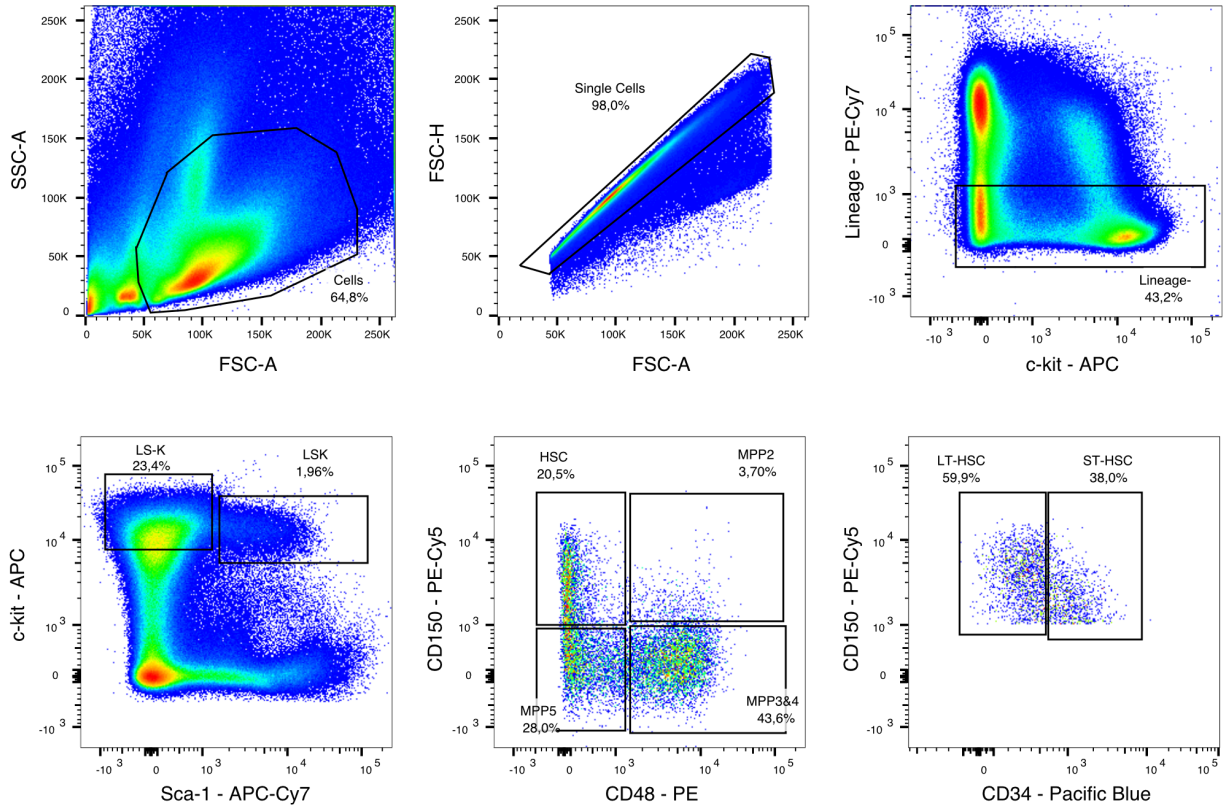
51. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213–219 (2013).
52. Jones, D. *et al.* cgpCaVEManWrapper: Simple Execution of CaVEMan in Order to Detect Somatic Single Nucleotide Variants in NGS Data. *Curr Protoc Bioinformatics* **56**, 15.10.1-15.10.18 (2016).
53. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).
54. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
55. Huebschmann, D., Jopp-Saile, L., Andresen, C., Gu, Z. & Schlesner, M. *YAPSA: Yet Another Package for Signature Analysis.* (2020).
56. Feuerbach, L. *et al.* TelomereHunter – in silico estimation of telomere content and composition from cancer genomes. *BMC Bioinformatics* **20**, 272 (2019).
57. Longchamps, R. J. *et al.* Evaluation of mitochondrial DNA copy number estimation techniques. *PLOS ONE* **15**, e0228166 (2020).
58. Barker, N. *et al.* Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009).
59. Hoeijmakers, J. H. J. DNA Damage, Aging, and Cancer. *New England Journal of Medicine* **361**, 1475–1485 (2009).
60. Rossi, D. J. *et al.* Hematopoietic stem cell quiescence attenuates DNA damage response and permits DNA damage accumulation during aging. *Cell Cycle* **6**, 2371–2376 (2007).

61. The Mouse in Biomedical Research - 2nd Edition. https://www.elsevier.com/books/the-mouse-in-biomedical-research/fox/978-0-12-369454-6.
62. Karch, J. *et al.* Bax and Bak function as the outer membrane component of the mitochondrial permeability pore in regulating necrotic cell death in mice. *eLife* **2**, (2013).
63. Why Mouse Matters. *Genome.gov* https://www.genome.gov/10001345/importance-of-mouse-genome.
64. Ferrone, C. K., Blydt-Hansen, M. & Rauh, M. J. Age-Associated TET2 Mutations: Common Drivers of Myeloid Dysfunction, Cancer and Cardiovascular Disease. *Int J Mol Sci* **21**, (2020).
65. Allsopp, R. C., Cheshier, S. & Weissman, I. L. Telomere Shortening Accompanies Increased Cell Cycle Activity during Serial Transplantation of Hematopoietic Stem Cells. *J Exp Med* **193**, 917–924 (2001).
66. Wang, Q. *et al.* Tagmentation-based whole-genome bisulfite sequencing. *Nature Protocols* **8**, 2022–2032 (2013).
67. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
68. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
69. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
70. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

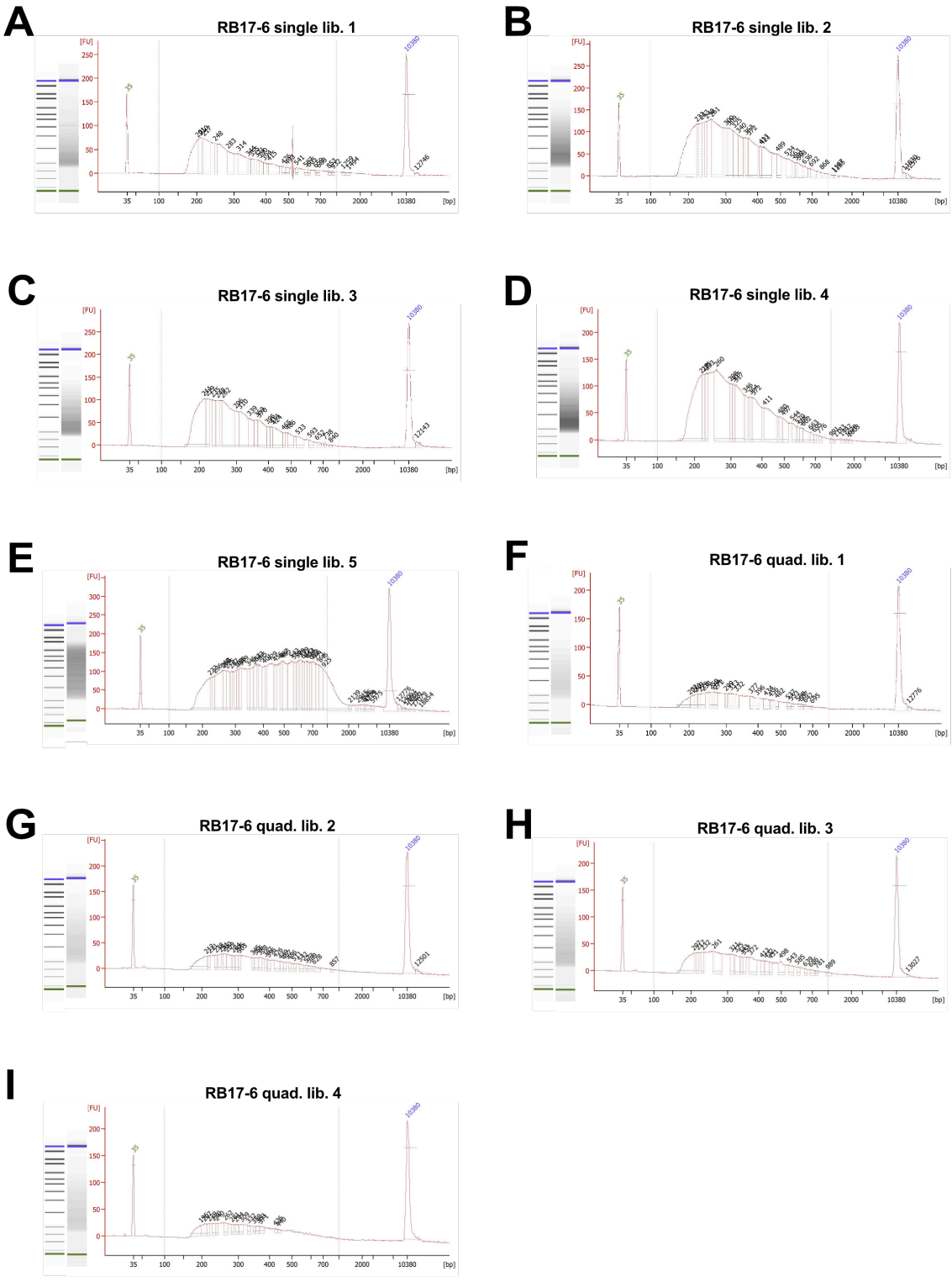
71. Pedersen, B. S. & Quinlan, A. R. Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics* **34**, 867–868 (2018).
72. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
73. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164–e164 (2010).

VIII. Supplementary

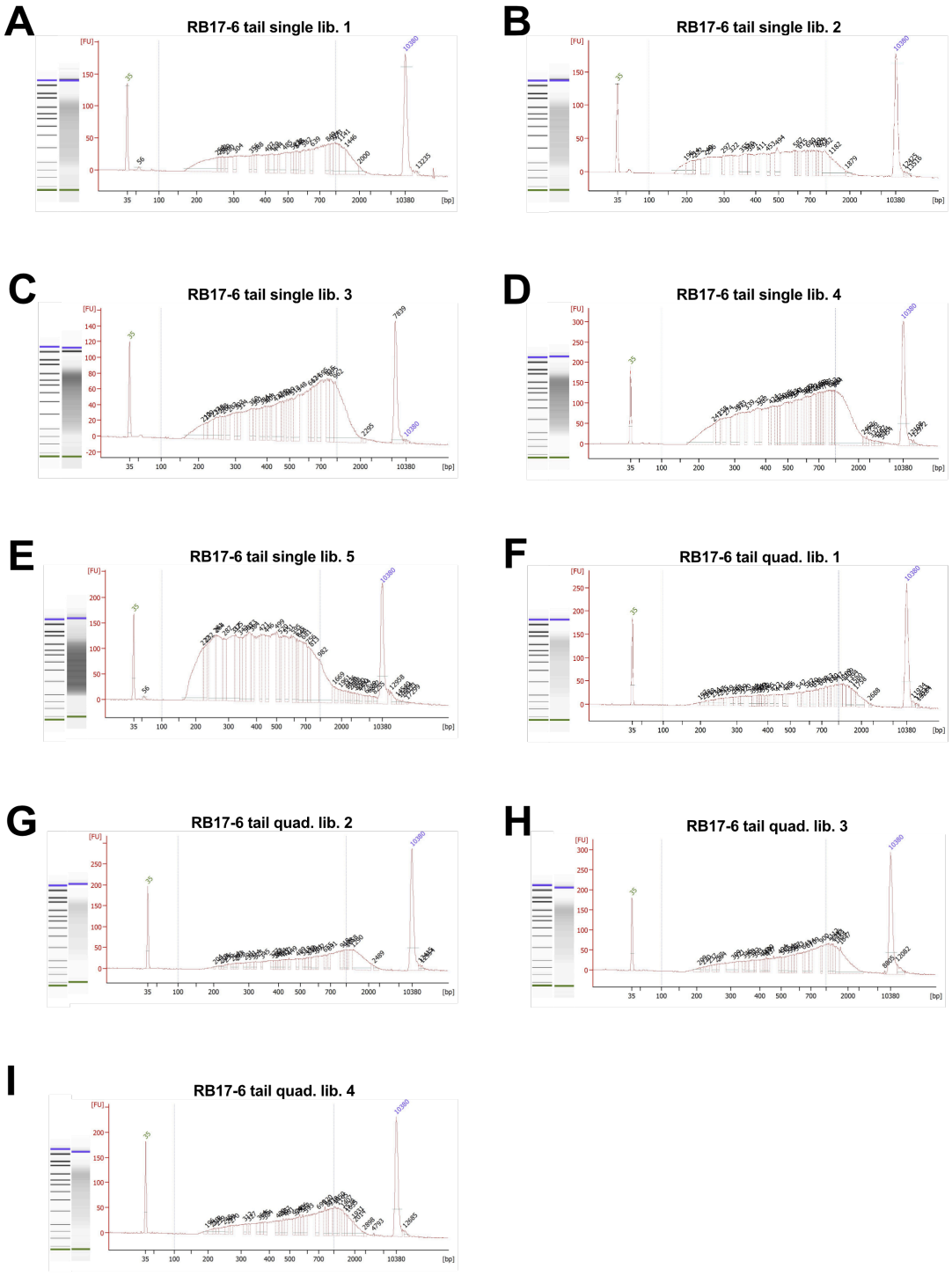
1. Supplementary figures



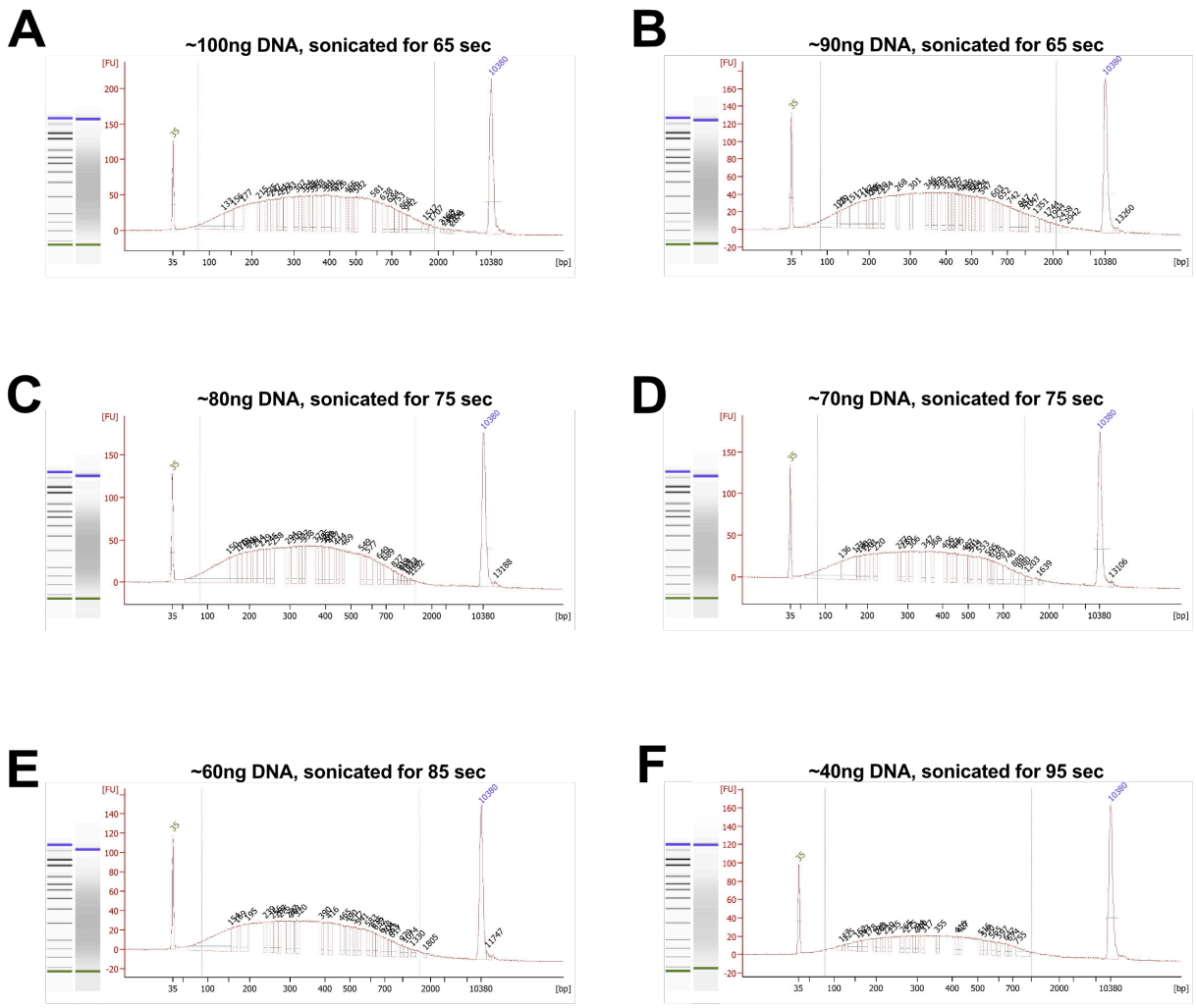
Supp. Figure 1. Sorting scheme for LT-HSC sort.



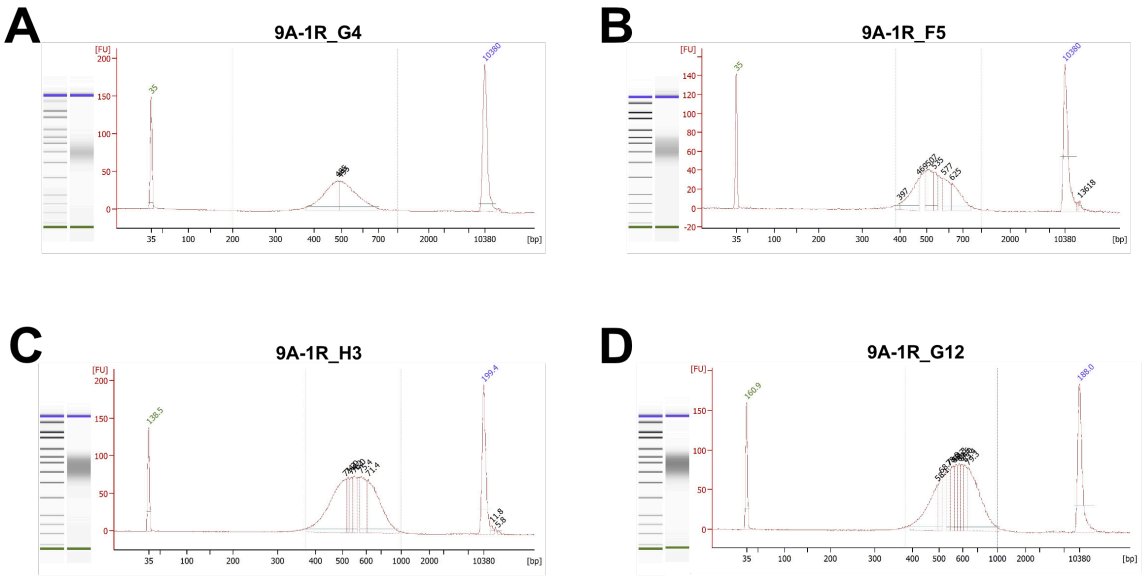
Supp. Figure 2. Bioanalyzer images from sequencing libraries prepared for the benchmark colony dataset. A – I. Fragment distribution of single and quadriplex libraries. Ladder is on the far left. Titles above indicate library name.



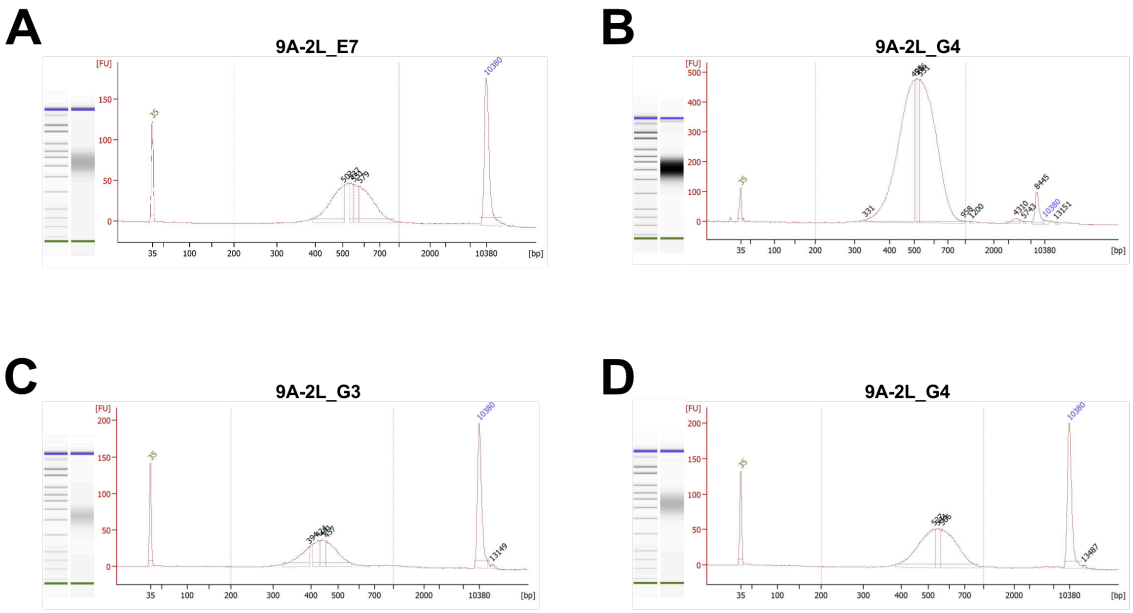
Supp. Figure 3. Bioanalyzer images from sequencing libraries prepared for the benchmark tail dataset. A – I. Fragment distribution of single and quadriplex libraries. Ladder is on the far left. Titles above indicate library name.



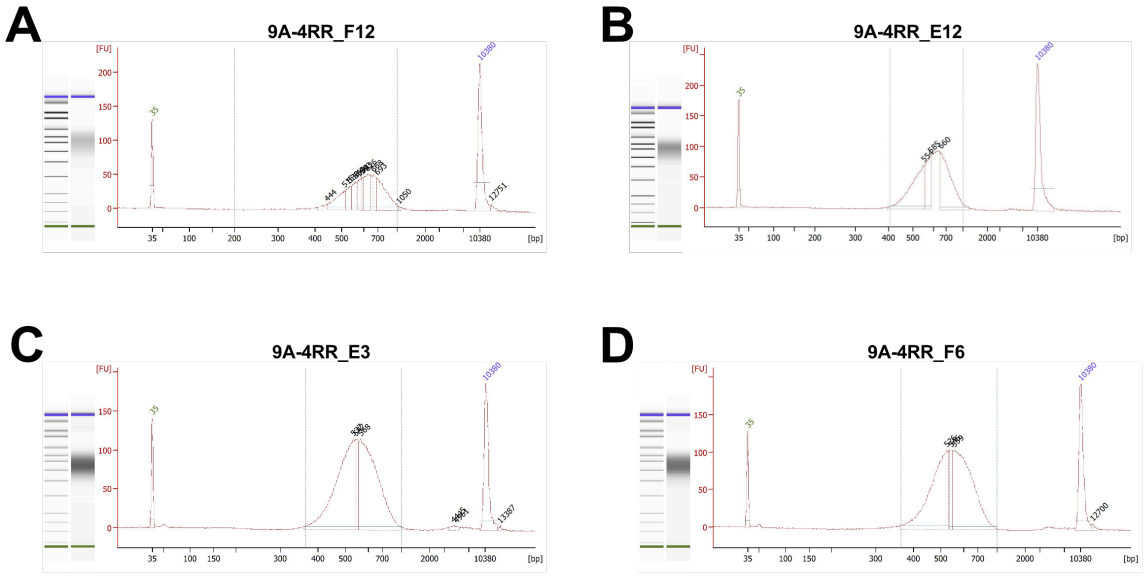
Supp. Figure 4. Bioanalyzer images from DNA fragment sonification optimization. A – F. Fragment distribution of sonicated DNA. Ladder is on the far left. Titles above indicate the amount of input DNA and how long it was sonicated for.



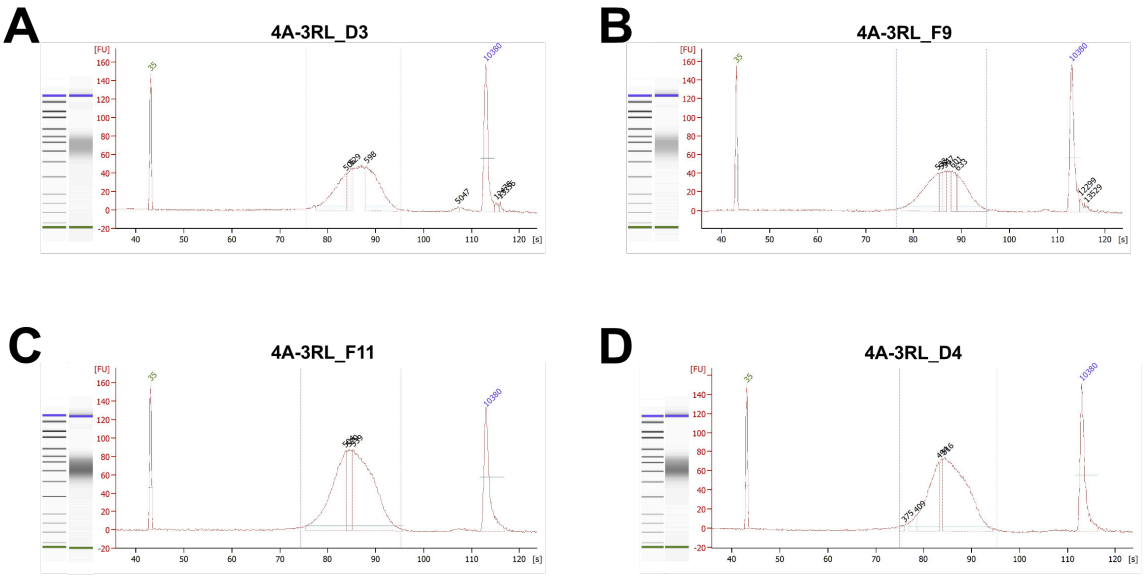
Supp. Figure 5. Bioanalyzer images from sequencing libraries prepared for mouse 9A-1R of the young group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



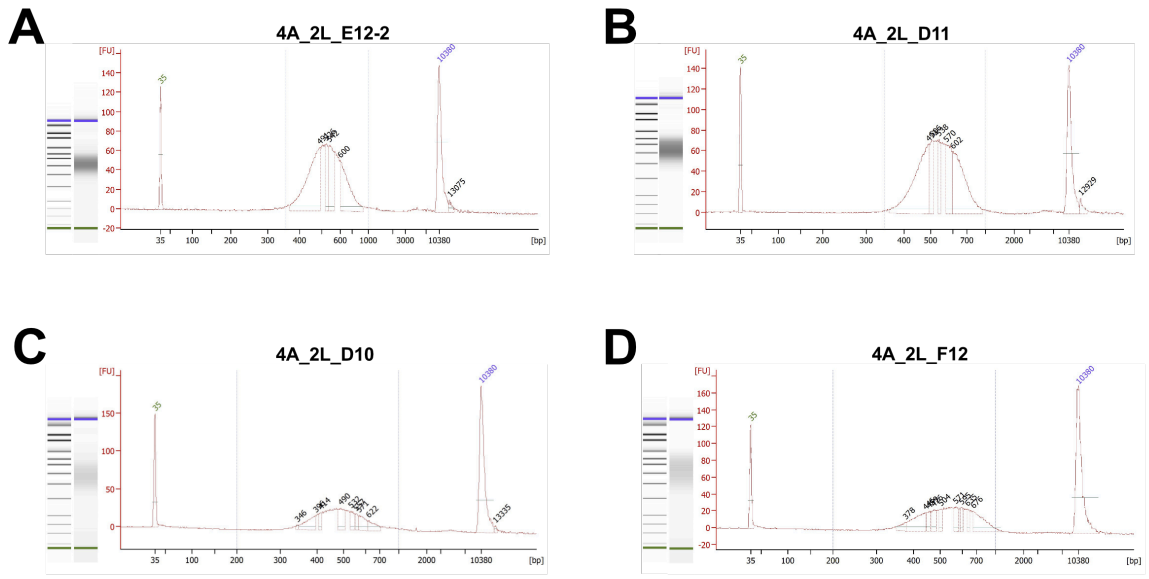
Supp. Figure 6. Bioanalyzer images from sequencing libraries prepared for mouse 9A-2L of the young group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



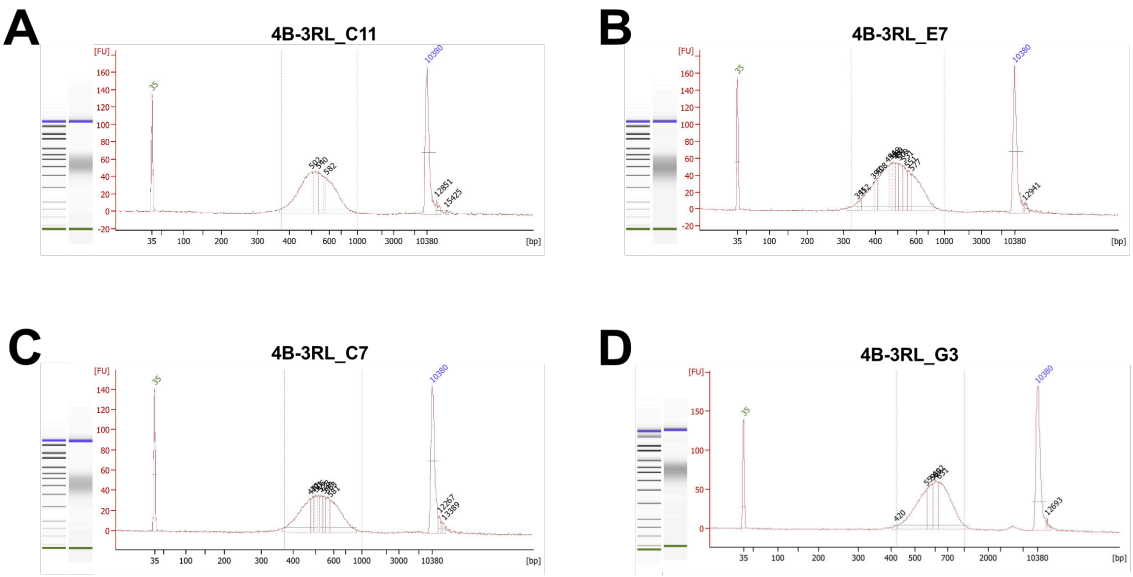
Supp. Figure 7. Bioanalyzer images from sequencing libraries prepared for mouse 9A-4RR of the young group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



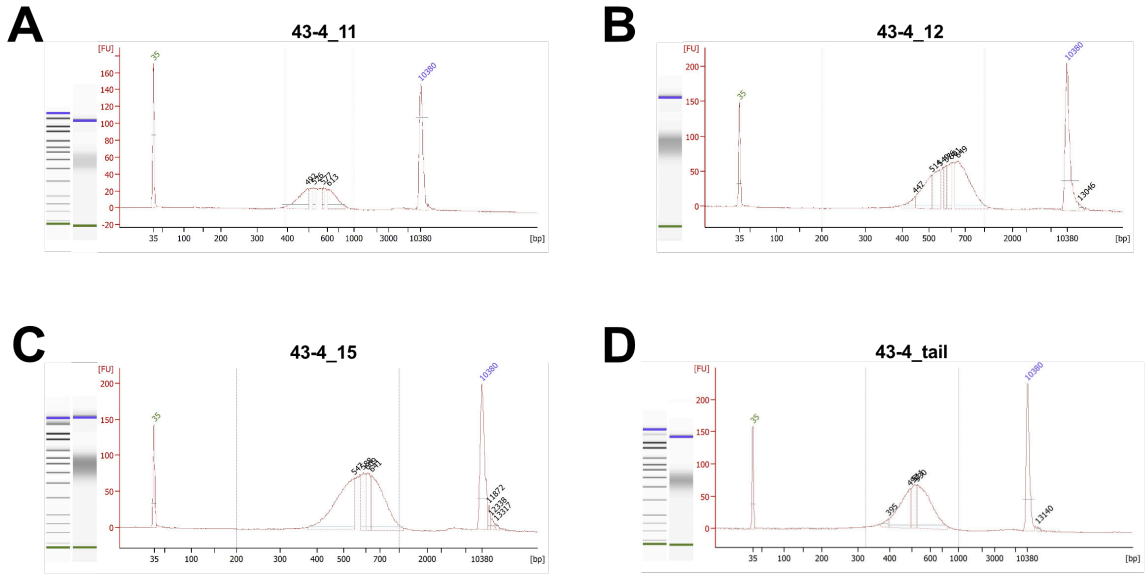
Supp. Figure 8. Bioanalyzer images from sequencing libraries prepared for mouse 4A-3RL of the old group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



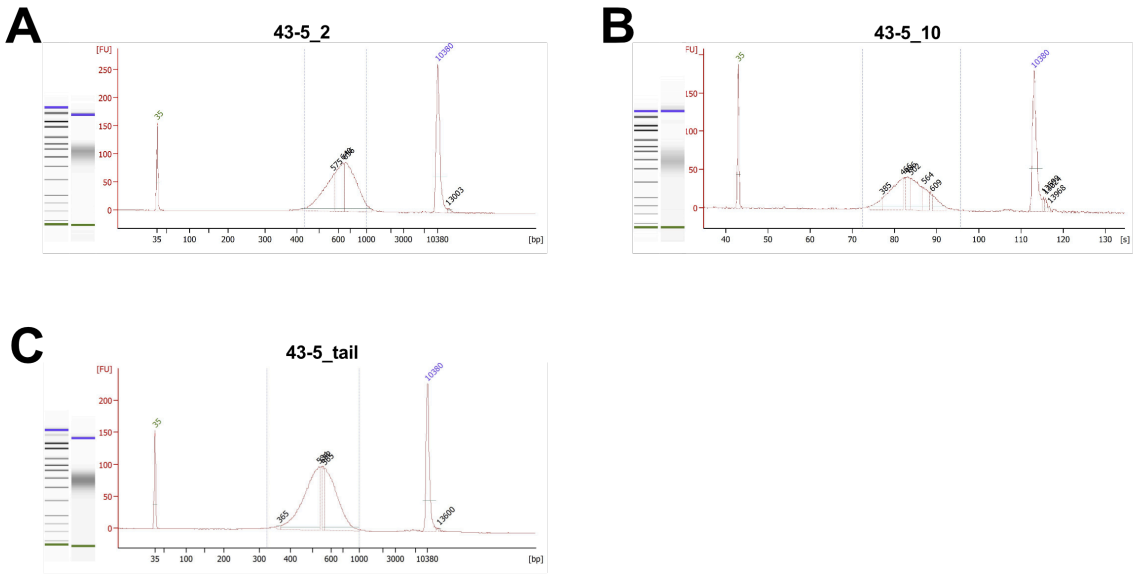
Supp. Figure 9. Bioanalyzer images from sequencing libraries prepared for mouse 4A-2L of the old group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



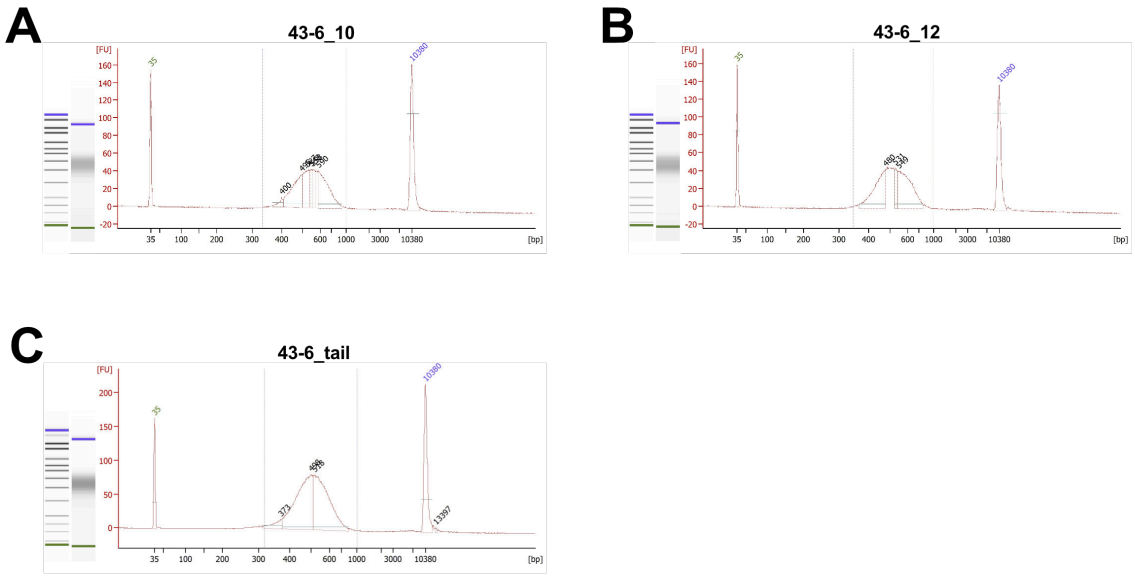
Supp. Figure 10. Bioanalyzer images from sequencing libraries prepared for mouse 4B-3RL of the old group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



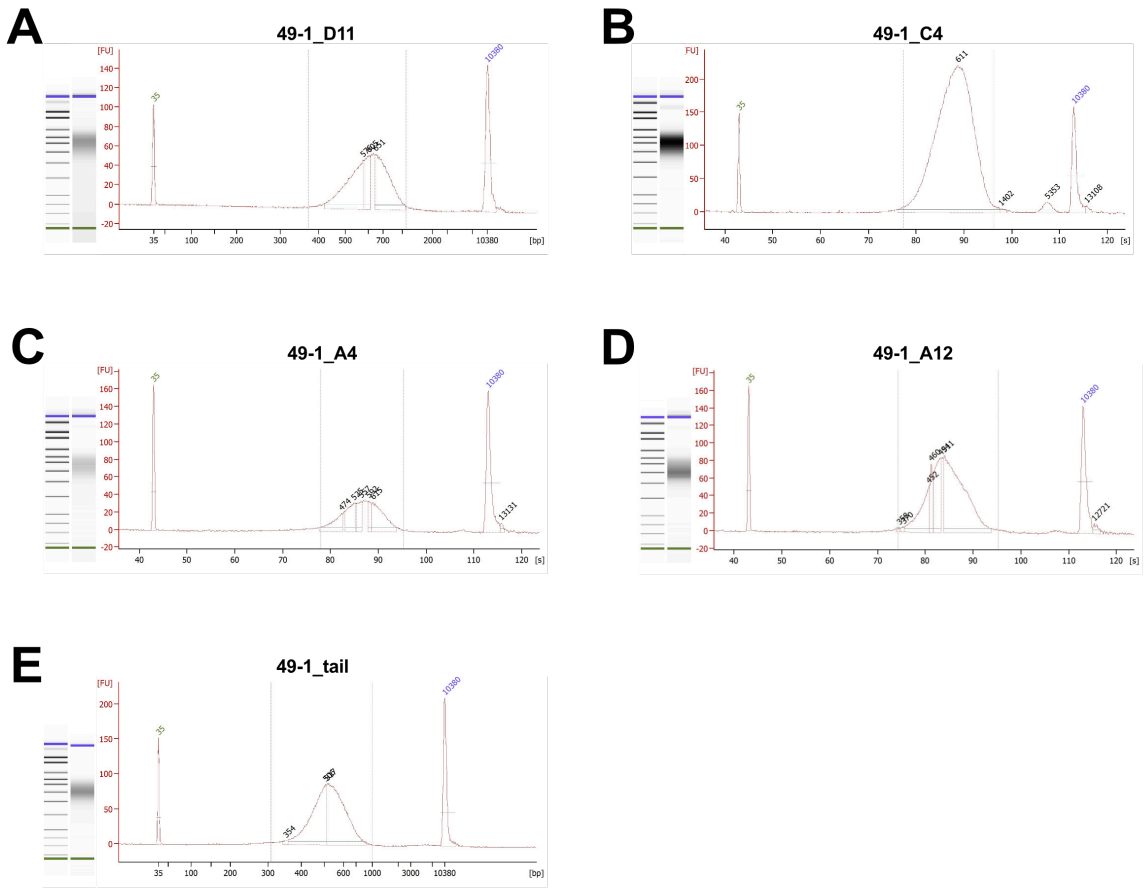
Supp. Figure 11. Bioanalyzer images from sequencing libraries prepared for mouse 43-4 of the pIpC group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



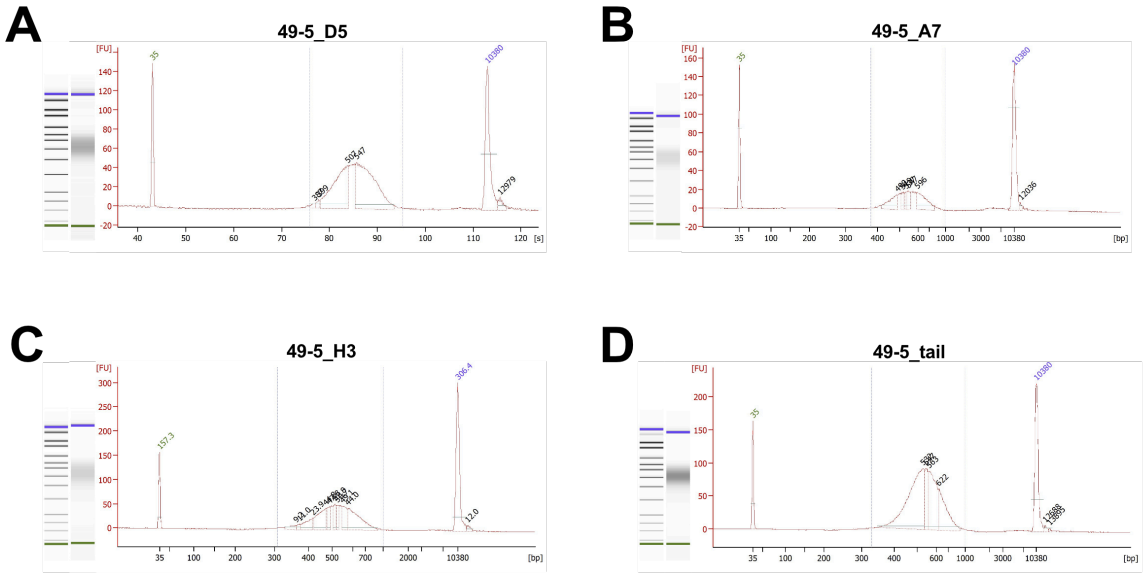
Supp. Figure 12. Bioanalyzer images from sequencing libraries prepared for mouse 43-5 of the pIpC group. A – C. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



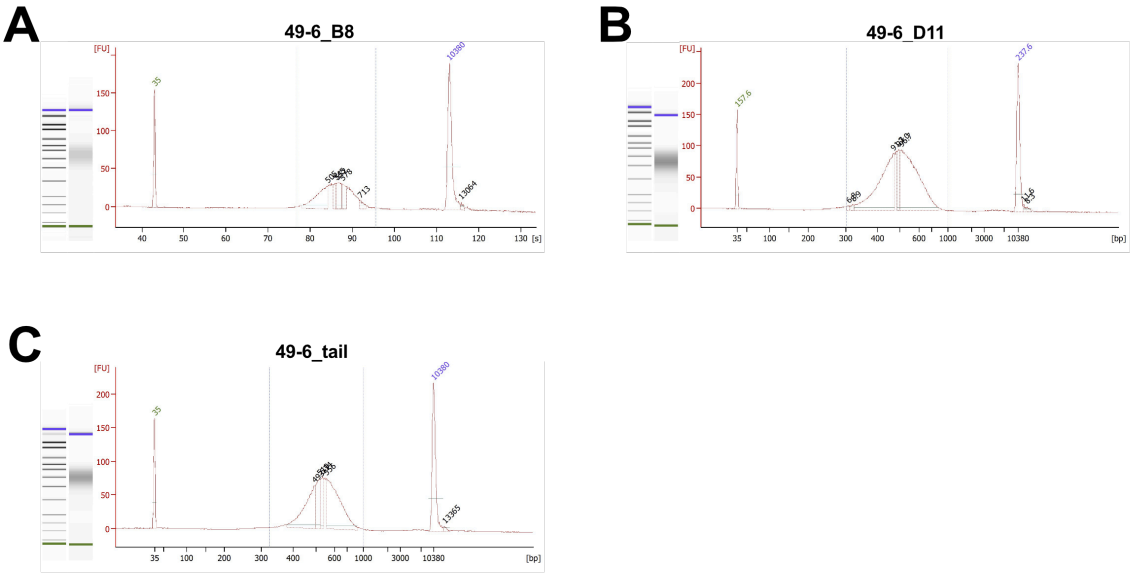
Supp. Figure 13. Bioanalyzer images from sequencing libraries prepared for mouse 43-6 of the pIpC group. A – C. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



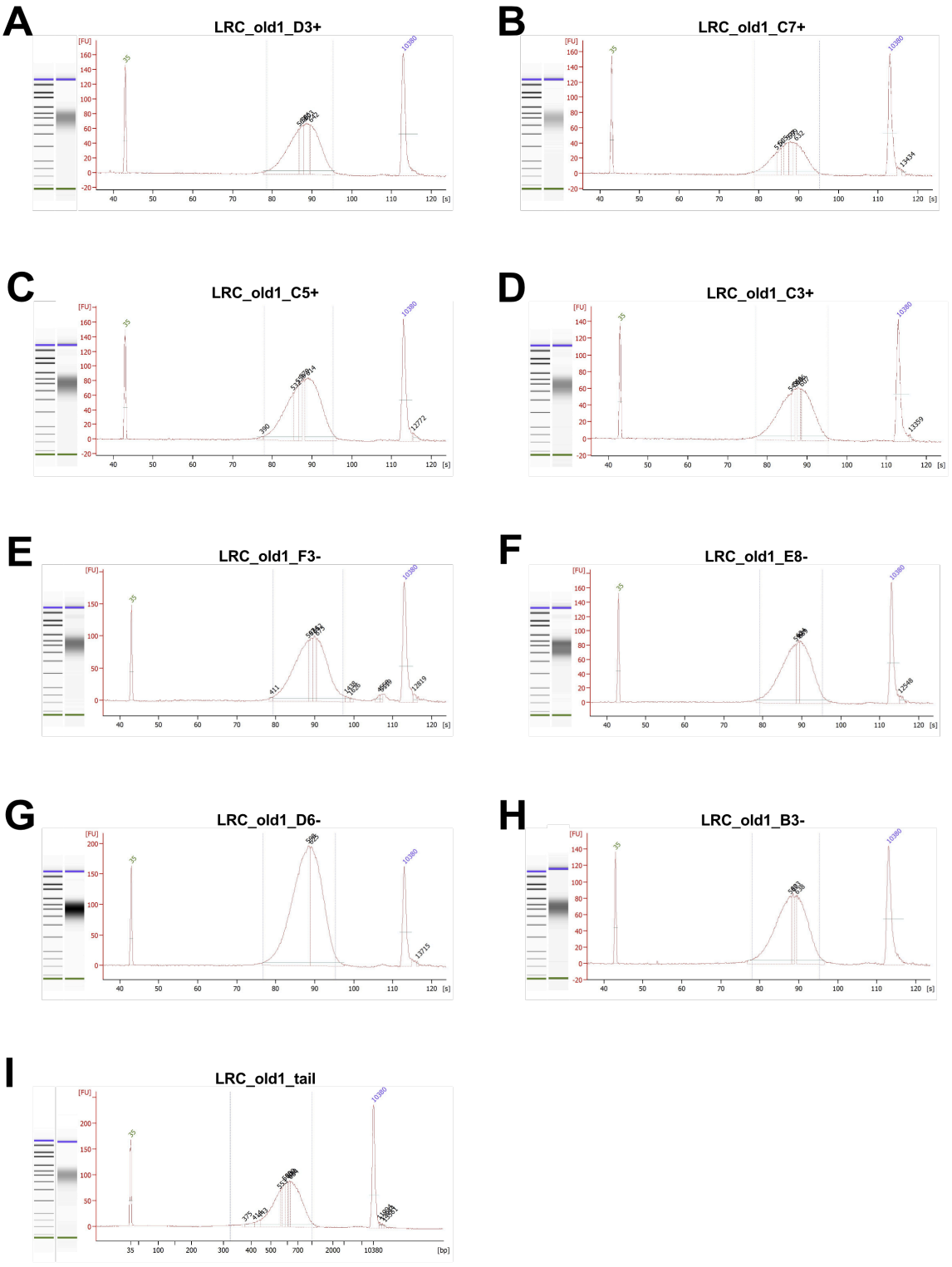
Supp. Figure 14. Bioanalyzer images from sequencing libraries prepared for mouse 49-1 of the TPO group. A – E. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



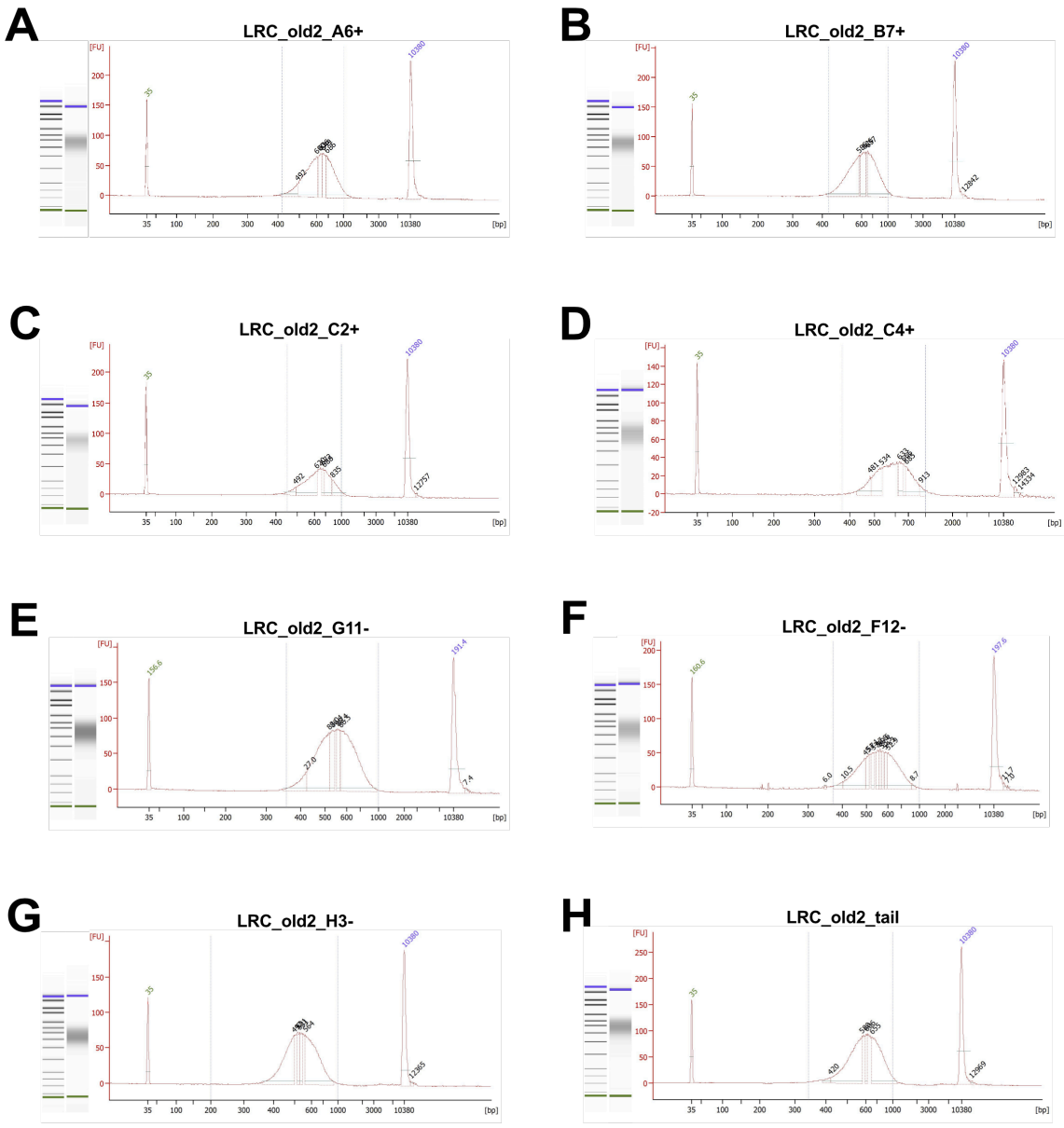
Supp. Figure 15. Bioanalyzer images from sequencing libraries prepared for mouse 49-5 of the TPO group. A – D. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



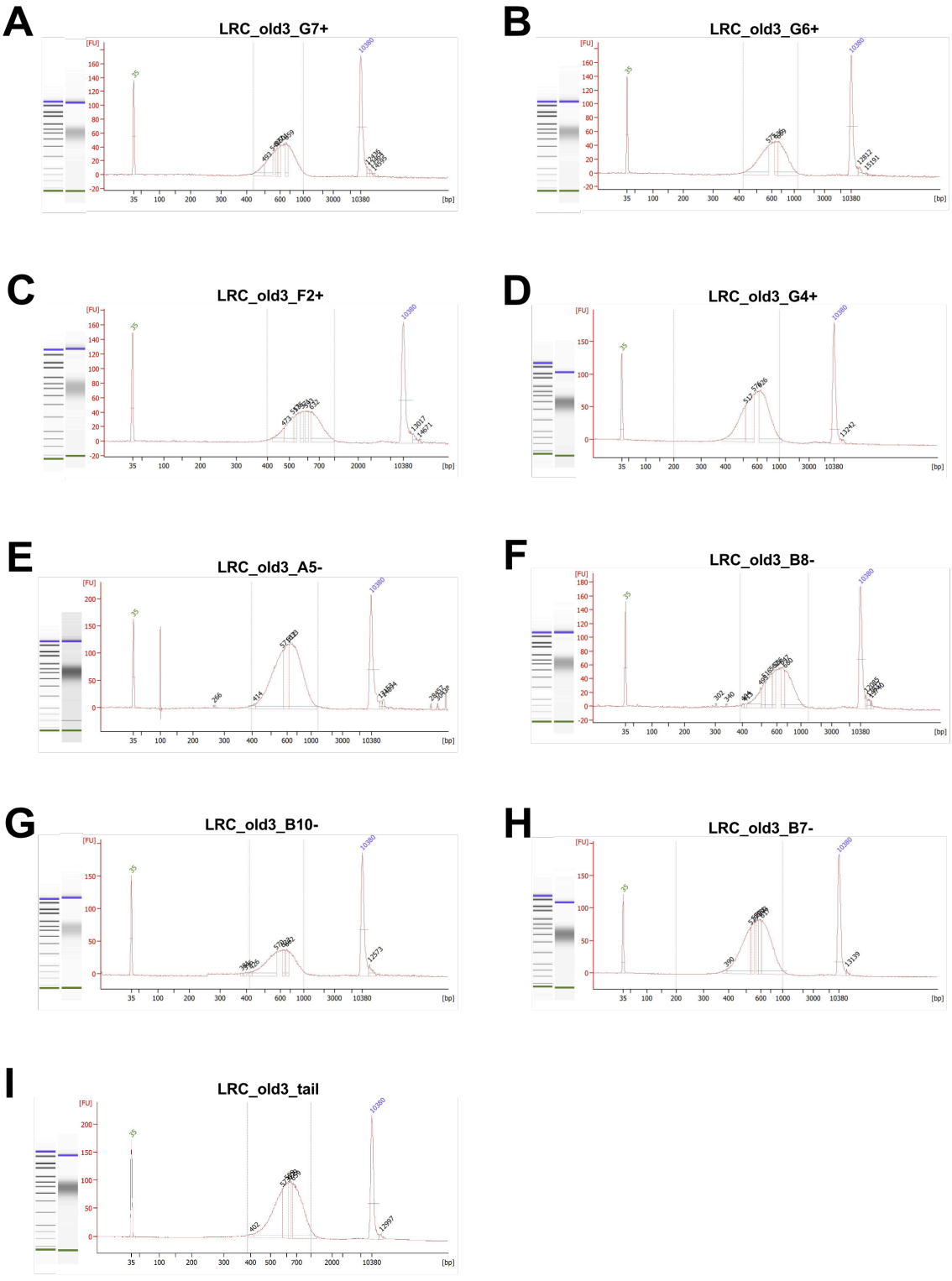
Supp. Figure 16. Bioanalyzer images from sequencing libraries prepared for mouse 49-6 of the TPO group. A – C. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



Supp. Figure 17. Bioanalyzer images from sequencing libraries prepared for mouse LRC old 1 of the LRC and non-LRC groups. A – I. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



Supp. Figure 18. Bioanalyzer images from sequencing libraries prepared for mouse LRC old 2 of the LRC and non-LRC groups. A – H. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.



Supp. Figure 19. Bioanalyzer images from sequencing libraries prepared for mouse LRC old 3 of the LRC and non-LRC groups. A – I. Fragment distribution of the single libraries. Ladder is on the far left. Titles above indicate library name.

2. Supplementary tables

Supp. Table 1. Overview of sequencing parameters for each library.

Sample	Group	Coverage (X)	Duplicates (%)	PCR cycles	DNA input (ng)
RB17-6 single lib. 1	Benchmark colony	89	N/A	10	4
RB17-6 single lib. 2	Benchmark colony		N/A	10	4
RB17-6 single lib. 3	Benchmark colony		N/A	10	4
RB17-6 single lib. 4	Benchmark colony		N/A	10	4
RB17-6 single lib. 5	Benchmark colony		N/A	10	4
RB17-6 quad. lib. 1	Benchmark colony		N/A	10	4
RB17-6 quad. lib. 2	Benchmark colony		N/A	10	4
RB17-6 quad. lib. 3	Benchmark colony		N/A	10	4
RB17-6 quad. lib. 4	Benchmark colony		N/A	10	4
RB17-6 tail single lib. 1	Benchmark tail	98	N/A	10	4
RB17-6 tail single lib. 2	Benchmark tail		N/A	10	4
RB17-6 tail single lib. 3	Benchmark tail		N/A	10	4
RB17-6 tail single lib. 4	Benchmark tail		N/A	10	4
RB17-6 tail single lib. 5	Benchmark tail		N/A	10	4
RB17-6 tail quad. lib. 1	Benchmark tail		N/A	10	4
RB17-6 tail quad. lib. 2	Benchmark tail		N/A	10	4
RB17-6 tail quad. lib. 3	Benchmark tail		N/A	10	4
RB17-6 tail quad. lib. 4	Benchmark tail		N/A	10	4
9A-2L_G3	Young	39.45	19.79	8	100
9A-2L_E7	Young	39.92	14.46	8	100
9A-2L_G4	Young	39.71	14.6	8	100
9A-2L_G8	Young	43.46	15.46	8	100
9A-2L_tail	Young	41.5	N/A	N/A	N/A
9A-1R_G4	Young	37.75	22.35	8	85.5
9A-1R_F5	Young	43.06	18.08	8	100
9A-1R_G12	Young	41.87	15.71	9	53
9A-1R_H3	Young	40.47	14.91	9	54
9A-1R_tail	Young	39.6	N/A	N/A	N/A
9A-4RR_F12	Young	43.73	14.6	9	92
9A-4RR_E12	Young	44.13	14.34	8	100
9A-4RR_E3	Young	36.75	14.94	10	63
9A-4RR_F6	Young	36.26	18.67	10	43
9A-4RR_tail	Young	45.44	N/A	N/A	N/A
4B-3RL_C11	Old	38.96	16.62	8	88
4B-3RL_E7	Old	40.93	13.93	8	73

4B-3RL_C7	Old	38.63	17.91	8	71
4B-3RL_G3	Old	35.84	24.24	10	41
4B-3RL_tail	Old	35.83	N/A	N/A	N/A
4A-2L_E12-2	Old	36.85	18.18	10	50
4A-2L_D11	Old	36.78	18.89	10	48
4A-2L_D10	Old	30.26	24.67	10	47.9
4A-2L_F12	Old	42.72	16.11	8	66.8
4A-2L_tail	Old	36.02	N/A	N/A	N/A
4A-3RL_D3	Old	38.83	18.69	8	115
4A-3RL_F9	Old	39.69	19.68	8	65
4A-3RL_F11	Old	40.49	17.96	8	110
4A-3RL_D4	Old	38.86	19.51	8	110
4A-3RL_tail	Old	31.56	N/A	N/A	N/A
43-4_15	pIpC	35.04	16.3	10	46
43-4_12	pIpC	42.06	13.58	8	79
43-4_11	pIpC	35.33	28.65	9	51
43-4_tail	pIpC	36.68	15.23	8	100
43-5_2	pIpC	41.31	16.39	8	57.2
43-5_10	pIpC	35.42	23.95	8	27
43-5_tail	pIpC	40.46	14.09	8	100
43-6_12	pIpC	32.17	35.98	10	42
43-6_10	pIpC	42.69	18.91	8	72
43-6_tail	pIpC	42.57	13.35	8	100
49-1_D11	TPO	41.7	19.57	8	74.4
49-1_A4	TPO	29.08	33.17	10	30.6
49-1_A12	TPO	37.02	20.67	8	46
49-1_C4	TPO	33.97	20.75	10	39
49-1_tail	TPO	41.56	14.11	8	100
49-5_A7	TPO	23.13	44.95	10	33
49-5_D5	TPO	35.73	24.44	8	32
49-5_H3	TPO	40.06	23.92	9	51.48
49-5_tail	TPO	41.09	14.49	8	100
49-6_B8	TPO	31.85	32.1	8	28
49-6_D11	TPO	42.6	21.57	9	56.1
49-6_tail	TPO	36.92	15.34	8	100
LRC_old1_C7_pos	Old LRC	38.22	18.98	8	85
LRC_old1_D3_pos	Old LRC	34.11	15.24	8	89.5
LRC_old1_C5_pos	Old LRC	41.47	15.51	8	113
LRC_old1_C3_pos	Old LRC	40.66	16.8	8	109

LRC_old1_F3_neg	Old non-LRC	34.38	22.11	10	111.5
LRC_old1_E8_neg	Old non-LRC	38.76	16.38	8	112
LRC_old1_D6_neg	Old non-LRC	38.58	15.68	8	132
LRC_old1_B3_neg	Old non-LRC	42.1	15.86	8	94
LRC_old1_tail	Old LRC and non-LRC	43.08	17.22	8	100
LRC_old2_C2_pos	Old LRC	39.14	18.95	8	100
LRC_old2_A6_pos	Old LRC	42.6	18.09	8	100
LRC_old2_B7_pos	Old LRC	42.09	15.04	8	100
LRC_old2_C4_pos	Old LRC	37.76	18.98	8	100
LRC_old2_G11_neg	Old non-LRC	40.36	12.75	9	50
LRC_old2_F12_neg	Old non-LRC	38.19	16.69	9	50
LRC_old2_H3_neg	Old non-LRC	37.38	22.12	9	54
LRC_old2_tail	Old LRC and non-LRC	45.06	14.22	8	100
LRC_old3_G6_pos	Old LRC	34.95	19.41	9	100
LRC_old3_G7_pos	Old LRC	37.42	17.37	8	43.6
LRC_old3_F2_pos	Old LRC	37.91	19.29	9	26.7
LRC_old3_G4_pos	Old LRC	38.37	16.48	9	75.8
LRC_old3_A5_neg	Old non-LRC	42.14	13.81	8	70.9
LRC_old3_B10_neg	Old non-LRC	34.88	21.9	9	57.4
LRC_old3_B8_neg	Old non-LRC	41.02	15.38	8	100
LRC_old3_B7_neg	Old non-LRC	38.57	18.55	9	54
LRC_old3_tail	Old LRC and non-LRC	46.57	12.57	8	100

Supp. Table 2. Overview of HSC colony and mouse identifiers used for mutational signature analysis.

Colony	Colony identifier	Group	Mouse	Mouse identifier
9A-2L_G3	91241	Young	9A-2L	10001
9A-2L_E7	91242	Young		
9A-2L_G4	91243	Young		
9A-2L_G8	91244	Young		
9A-1R_G4	91111	Young	9A-1R	10002
9A-1R_F5	91112	Young		
9A-1R_G12	91113	Young		
9A-1R_H3	91114	Young		
9A-4RR_F12	91441	Young	9A-4RR	10003
9A-4RR_E12	91442	Young		
9A-4RR_E3	91443	Young		
9A-4RR_F6	91444	Young		
4B-3RL_C11	42331	Old	4B-3RL	20001

4B-3RL_E7	42332	Old		
4B-3RL_C7	42333	Old		
4B-3RL_G3	42334	Old		
4A-2L_E12-2	41221	Old	4A-2L	20002
4A-2L_D11	41222	Old		
4A-2L_D10	41223	Old		
4A-2L_F12	41224	Old		
4A-3RL_D3	41331	Old	4A-3RL	20003
4A-3RL_F9	41332	Old		
4A-3RL_F11	41333	Old		
4A-3RL_D4	41334	Old		
43-4_15	43415	pIpC	43-4	30001
43-4_12	43412	pIpC		
43-4_11	43411	pIpC		
43-5_2	43502	pIpC	43-5	30002
43-5_10	43510	pIpC		
43-6_12	43612	pIpC	43-6	30003
43-6_10	43610	pIpC		
49-1_D11	49101	TPO	49-1	40001
49-1_A4	49102	TPO		
49-1_A12	49103	TPO		
49-1_C4	49104	TPO		
49-5_A7	49501	TPO	49-5	40002
49-5_D5	49502	TPO		
49-5_H3	49503	TPO		
49-6_B8	49601	TPO	49-6	40003
49-6_D11	49602	TPO		
LRC_old1_C7_pos	10101	Old LRC	LRC old1	50001
LRC_old1_D3_pos	10102	Old LRC		
LRC_old1_C5_pos	10103	Old LRC		
LRC_old1_C3_pos	10104	Old LRC		
LRC_old1_F3_neg	10201	Old non-LRC	LRC old1	50011
LRC_old1_E8_neg	10202	Old non-LRC		
LRC_old1_D6_neg	10203	Old non-LRC		
LRC_old1_B3_neg	10204	Old non-LRC		
LRC_old2_C2_pos	20101	Old LRC	LRC old2	50002
LRC_old2_A6_pos	20102	Old LRC		
LRC_old2_B7_pos	20103	Old LRC		
LRC_old2_C4_pos	20104	Old LRC		

LRC_old2_G11_neg	20201	Old non-LRC	LRC old2	50022
LRC_old2_F12_neg	20202	Old non-LRC		
LRC_old2_H3_neg	20203	Old non-LRC		
LRC_old3_G6_pos	30101	Old LRC	LRC old3	50003
LRC_old3_G7_pos	30102	Old LRC		
LRC_old3_F2_pos	30103	Old LRC		
LRC_old3_G4_pos	30104	Old LRC		
LRC_old3_A5_neg	30201	Old non-LRC	LRC old3	50033
LRC_old3_B10_neg	30202	Old non-LRC		
LRC_old3_B7_neg	30203	Old non-LRC		
LRC_old3_B8_neg	30204	Old non-LRC		

3. Abbreviations and acronyms

A
ANOVA = analysis of variance
ASC = adult stem cell
B
BER = base excision repair
C
CO_2 = carbon dioxide
CHIP – clonal hematopoiesis of indeterminate potential
CN = copy number
D
DBS = doublet base substitutions
DNA = deoxyribonucleic acid
E
EPO = Erythropoietin
F
Fig = Figure
H
HR = homologous recombination
HSC = hematopoietic stem cell
HSPC = hematopoietic stem and progenitor cell
I
INDEL = insertions and deletions
IL = interleukin
IU = international unit
L
Lib = library
L-glut = L Glutamine

Lin^- = lineage negative
LRC = label-retaining cell
LT = Long-term

M

Min = minute
MMR = mismatch repair
mtDNA = mitochondrial DNA

N

N/A = not applicable
Neg = negative
NER = nucleotide excision repair
ng = nanogram
NGS = next generation sequencing
NHEJ = non-homologous end joining
Norm = normalized

O

O_2 = oxygen

P

PCR = polymerase chain reaction
pen/strep = Penicillin-Streptomycin
pIpC = polyinosinic:polycytidylic acid
Pos = positive

Q

Quad = quadriplex

R

Rm = recombinant murine
RNA = ribonucleic acid

S

SBS = single base substitution
SCL = stem cell leukemia
scRNA-seq = single-cell RNA sequencing
Sec = second
Sig = signature
SNV = single nucleotide variant
Supp = supplementary
ST = short-term

T

TPO = thrombopoietin
TLS = translesion synthesis

V

VAF = variant allele frequency

W

WGS = whole genome sequencing

4. List of figures

(B) Reprinted by permission from **Springer Nature Customer Service Centre GmbH**: Springer Nature, **Nature**, Lee-Six, H. *et al.* Population dynamics of normal human blood inferred from somatic mutations. 561, 473–478, **Springer Nature and Copyright Clearance Center** (2018).

Figure 8. Clonal expansion of single HSCs to determine somatic mutation burden.
Reprinted with permission from Osorio, F. G. *et al.* Somatic Mutations Reveal Lineage Relationships and Age-Related Mutagenesis in Human Hematopoiesis. ***Cell Rep***. 25, 2308-2316.e4 (2018) [DOI: 10.1016/j.celrep.2018.11.014] through the **Creative Commons user license**: https://creativecommons.org/licenses/by-nc-nd/4.0/.

Figure 9. Model of variant allele frequency (VAF) distribution of heterozygous variants within single-cell derived cultures.
(B) Reprinted by permission from **Springer Nature Customer Service Centre GmbH**: Springer Nature, **Nature Protocols**, Jager, M. *et al.* Measuring mutation accumulation in single human adult stem cells by whole-genome sequencing of organoid cultures. 13, 59–78, **Springer Nature and Copyright Clearance Center** (2018).

Figure 10. Schematic of experimental overview for WGS of single HSC colonies.

Figure 11. Variant allele frequency distributions for pre- and post-filtered MuTect and CaVEMan SNVs.

Figure 12. Effect of filtering parameters and coverage on the ability to call variants in HSC colonies using WGS.

Figure 13. Assessment of results from individual SNV callers, MuTect and CaVEMan, compared to the overlapping results from both.

Figure 14. Effect of genomic DNA input and alignment quality control measures on mutation burden.

Figure 15. Age-associated increase in mutation burden within HSC colonies.

Figure 16. Comparison of SNV burden between mice from young and old groups.

Figure 17. Comparison of INDEL burden between mice from young and old groups.

Figure 18. Replication-associated increase in mutation burden within HSC colonies.

Figure 19. Comparison of SNV burden between mice from LRC and non-LRC groups reveals an increase in SNVs with replication.

Figure 20. Comparison of SNV burden between mice from LRC and non-LRC groups.

Supp. Figure 2. Bioanalyzer images from sequencing libraries prepared for the benchmark colony dataset.

Supp. Figure 3. Bioanalyzer images from sequencing libraries prepared for the benchmark tail dataset.

Supp. Figure 4. Bioanalyzer images from DNA fragment sonification optimization.

Supp. Figure 5. Bioanalyzer images from sequencing libraries prepared for mouse 9A-1R of the young group.

Supp. Figure 6. Bioanalyzer images from sequencing libraries prepared for mouse 9A-2L of the young group.

Supp. Figure 7. Bioanalyzer images from sequencing libraries prepared for mouse 9A-4RR of the young group.

Supp. Figure 8. Bioanalyzer images from sequencing libraries prepared for mouse 4A-3RL of the old group.

Supp. Figure 9. Bioanalyzer images from sequencing libraries prepared for mouse 4A-2L of the old group.

Supp. Figure 10. Bioanalyzer images from sequencing libraries prepared for mouse 4B-3RL of the old group.

Supp. Figure 11. Bioanalyzer images from sequencing libraries prepared for mouse 43-4 of the pIpC group.

Supp. Figure 12. Bioanalyzer images from sequencing libraries prepared for mouse 43-5 of the pIpC group.

Supp. Figure 13. Bioanalyzer images from sequencing libraries prepared for mouse 43-6 of the pIpC group.

Supp. Figure 14. Bioanalyzer images from sequencing libraries prepared for mouse 49-1 of the TPO group. A

Supp. Figure 15. Bioanalyzer images from sequencing libraries prepared for mouse 49-5 of the TPO group.

Supp. Figure 16. Bioanalyzer images from sequencing libraries prepared for mouse 49-6 of the TPO group.

5. List of tables

IX. Contributions

The work presented in this thesis would not have been possible without the contribution from several people:

Dr. Michael Milsom and Prof. Dr. Benedikt Brors for supervision of this study.

Dr. Inigo Martincorena and Dr. Alex Cagan at the Wellcome Trust Sanger Institute.

The DKFZ-MOST research grant for funding my PhD.

Dr. Charles Imbusch

Julia Knoch

Dr. Ruzhica Bogeska

Marleen Büchler-Schäff, M.Sc.

Dr. Susanne Lux

The Genomics and Proteomics DKFZ Core Facility

The Omics IT and Data Management DKFZ Core Facility

The Flow Cytometry DKFZ Core Facility

The Light Microscopy DKFZ Core Facility