Competition and interference in clonal haematopoiesis

Blood cell production is characterized by a tree-like differentiation hierarchy, at its root sourced by a collection of stem cells generally located in the bone marrow. These continuously provide offspring that develop into the diverse cast of differentiated immune cells and erythrocytes, and as such must remain viable over the course of a lifetime to ensure healthy functioning of the blood. Nevertheless, with age they acquire somatic mutations [REF] which can ultimately affect the operation of their differentiated offspring. If a mutation conveys some fitness advantage, it can expand within the HSC pool, resulting in a high prevalence in the blood. The presence of a large mutant clone – accompanied by a drastic reduction in genetic heterogeneity of blood cells – is referred to as \textit{clonal haematopoiesis} (CH), and is associated with numerous haematological disorders [REF].

Current evidence suggests that CH is rare in young individuals, however, from around 50 to 60 years of age the detectable presence of expanded mutant clones (larger than 1\% of all blood) surges drastically [REF]. In fact, multiple distinct clones may exist simultaneously in a single individual, as shown in a recent study by Mitchell et al. [REF]: In their cohort of 10 patients aged 29 to 81, the four oldest (75 to 81) presented between 12 and 18 concurrent clones comprising up to 61\% of the patient’s total blood. In another recent study, Fabre et al. showed that many clones of known driver genes presented stable logistic growth when tracked over a decade’s time, consistent with mathematical models of a fit mutant in a size-limited population [REF].

Taken together, these findings suggest clonal haematopoiesis may be the natural outcome of a surprisingly simple model of cellular dynamics, in which mutants are born in the HSC population at random times, and may expand at a rate determined by their genetically prescribed fitness. However, there remain some observations do not fit this idealized mould. First, it has been found that clones with similar or identical driver mutations may present different dynamic behaviour across individuals [REFs]. This muddles the supposition that a particular genetic alteration would cause a fixed change in reproductive fitness, and instead introduces the concept of a patient-dependent fitness. And second, Fabre et al. found that, by combining their time-resolved measurements with phylogenetic analysis, the expanding clones which originated early in life eventually slowed their growth at old age – more so than expected from a logistic model of evolution [REF]. In fact, more than a handful of the driver clones they followed decreased in size over the follow-up period, despite being initially present at a size unlikely to be reached by neutral drift alone.

While different explanations for these observations can be hypothesized, such as varying environmental factors, age-related changes in stem cell niches, or more specific cell-intrinsic changes, in this work we propose a simple sophistication of this cellular dynamics model, from which these unexplained phenomena arise naturally. The key insight is that we do not take expanding clones to act independently of one another, which is the implicit assumption of the logistic model. Instead, we consider the case where new fit clones alter the current fitness landscape simply by their presence. This introduces competition between existing clones, making the observed growth rates highly individual-specific. The concept of distinct fit clones competing within a population is in fact not novel, and has been referred to as \textit{clonal interference} in the context of evolution and maintenance of sex [REFs!].

The model we investigate here hinges on the assumption that, once it has reached maturity, the total stem cell compartment does not increase in size, even as fit clones arise and expand within it.

# Abstract

The production of blood cells is driven by a small group of haematopoietic stem cells (HSCs) whose offspring develop into the diverse cast of differentiated immune cells and erythrocytes found in the blood. With age these HSCs acquire mutations \cite{Abascal2021} which, if conveying some selective advantage, can expand within the HSC pool, resulting in a high prevalence in the blood. This phenomenon, known as \textit{clonal haematopoiesis} (CH), is pervasive amongst the elderly, and is associated with various haematological disorders \cite{Zink2017}.

Current evidence hints CH might be the natural outcome of a surprisingly simple model of cellular dynamics: One where mutants are born in the population at random times and expand according to a genetically determined fitness. Indeed, recent studies have shown that multiple distinct clones often exist simultaneously \cite{Mitchell2022}, and many clones of known cancer genes present stable sigmoidal growth when tracked over time \cite{Fabre2022}. However, two important observations appear to contradict this picture: Firstly, identical mutations do not confer the same fitness across individuals; and secondly, clonal fitness appears to decrease later in life \cite{Fabre2022}.

While numerous biological explanations have been offered for this discrepancy, we propose that these phenomena can in fact arise naturally from a more sophisticated population genetics model. Taking into account that distinct clones in a fixed-size population cannot act independently, we arrive at a picture of competing mutants whose effective fitness decreases as the competition grows, a phenomenon known as \textit{clonal interference} \cite{Gerrish1998}. We show that allowing for such competition, the model recapitulates many of the important observations made in recent experimental work, including those under scrutiny: a sudden surge of CH prevalence from 50-60 years of age, the existence of multiple clones in older individuals, a variable patient-dependent fitness landscape, and a deceleration of clonal expansion late in life.

# Plots

* Decreasing expansion:
  + logistic fits show decreasing fitness; Fabre data + simulations
  + summary statistics showing decreasing fitness:
    - mean fitness (can we calculate this analytically?)
    - Spearman correlation of mean fitness?
* Individual trajectories?
* Clone size distributions:
  + Size dist across all individuals: data vs sims
  + Number of expanded clones per individual

# Paper structure v2:

* Introduce CH
* Describe current knowledge concerning it:
  + Prevalence in elderly
  + Relation to CV afflictions
  + Understanding of clone dynamics
* Describe open questions relating to dynamics
  + Why is there so much inter-individual variability?
  + Do the dynamics change with age?
* Introduce model of clone dynamics:
  + Clones arise stochastically
  + Total population size is fixed -> clones compete
* Maybe put equation? Refer to methods for derivation, etc.!

## Clone expansion rate decreases with age under clonal competition

* Explain that the mathematical prediction for an advantaged clone in a fixed-size population describes a logistic growth function: fast (exponential) initial growth with a deceleration as the clone approaches fixation.
* Note that in individuals expansion appears to decelerate:
  + Show Fabre data with decreasing logistic fitness
* Show how competition decelerates clones compared to logistic growth:
  + Illustrate with equation
  + Show plot of logistic fits to decelerating clones

## Randomness in arrival times causes inter-patient variability

* Explain how individual clone fitness (and therefore growth trajectory) depends on environment, in particular on existence and size of other clones. Stochasticity of clone arrival times thus causes distinct environments for different individuals. This leads to high variability, even for identical mutations.
* Illustrate this with figure of:
  + Individual trajectories
  + Comparison of inter-patient growth rates? Fitness distribution?
* Talk about effect of fitness distribution: Causes clones decreasing clones as well.
* (perhaps worth noting that such inter-patient variability can also arise from more purely biological sources, e.g. chronic inflammation, etc.

## Clone size distributions

* Size distribution is heavily tailed: Large abundance of “low-fitness” clones, much fewer high-fitness clones.
* Because high-fitness clones have a lower chance of occurring, we tend to see them later in life.
* By late life: large number of low-fitness “invisible clones” contribute strongly to competition induced slowdown effect.