Clone expansion of mutation-driven clonal hematopoiesis is associated with aging and metabolic dysfunction in individuals with obesity

Rosanne C. van Deuren[1-4#], Johanna C. Andersson-Assarsson[5#], Felipe M. Kristensson[5,6], Marloes Steehouwer[1,3], Kajsa Sjöholm[5], Per-Arne Svensson[7], Marc Pieterse[1], Christian Gilissen[1,3], Magdalena Taube[5], Peter Jacobson[5], Rosie Perkins[5], Han G. Brunner[1,8-10], Mihai G. Netea[2-4,11], Markku Peltonen[12,13], Björn Carlsson[5,14], Alexander Hoischen[1-4,*], Lena M.S. Carlsson[5,*]

[#]Shared first authorship
[*]Shared last authorship

Affiliations:
[1.] Department of Human Genetics, Radboud University Medical Center, Nijmegen, the Netherlands;
[2.] Department of Internal Medicine, Radboud University Medical Center, Nijmegen, the Netherlands;
[3.] Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, the Netherlands;
[4.] Radboud Expertise Center for Immunodeficiency and Autoinflammation and Radboud Center for Infectious Disease (RCI), Radboud University Medical Center, Nijmegen, the Netherlands;
[5.] Institute of Medicine, Department of Molecular and Clinical Medicine, the Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden;
[6] Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden
[7.] Institute of Health and Care Sciences, the Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
[8.] Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, the Netherlands;
[9.] Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands;
[10.] GROW School of Oncology and developmental biology, and MHeNs School of Mental Health and Neuroscience, Maastricht University, the Netherlands;
[11.] Department of Immunology and Metabolism, Life and Medical Sciences Institute (LIMES), University of Bonn, Bonn, Germany;
[12.] Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Sweden;
[13.] Public Health Promotion Unit, National Institute for Health and Welfare, Helsinki, Finland;
[14.] Early Clinical Development, Research and Early Development, Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

Corresponding authors:
Alexander Hoischen
Departments of Human Genetics and Internal Medicine, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center,
Geert Grooteplein 10, 6525 GA Nijmegen, The Netherlands
e-mail: alexander.hoischen@radboudumc.nl

Lena Carlsson
SOS Secretariat,
Vita Stråket 15,
Sahlgrenska University Hospital,
S-413 45 Gothenburg, Sweden.
email: lena.carlsson@medic.gu.se

ABSTRACT
Aging is associated with clonal hematopoiesis (CH) caused by somatic mutations in blood cell progenitors. Mutations with ≥2% variant allele frequency (VAF), known as clonal hematopoiesis of indeterminate potential (CHIP) mutations, have been linked to risk of hematological malignancies and cardiovascular disease. Using an ultrasensitive single-molecule molecular inversion probe (smMIP) sequencing assay, we identified clonal hematopoiesis driver mutations (CHDMs) in a set of established CH driver genes in individuals with obesity from the Swedish Obese Subjects study. In cross-sectional setting, with samples from 1050 individuals, we identified 273 candidate CHDMs in 216 individuals, with a VAF ranging from 0.01% to 31.15% and with clone sizes and prevalence increasing with age. Longitudinal analysis over 20 years in 40 individuals showed that small clones may grow over time and become CHIP. Greater speed of clone growth was associated with insulin resistance (R=0.04, P=0.025) and low circulating levels of high-density lipoprotein-cholesterol (R=-0.68, P=$1.74E^{-05}$), suggesting that dysfunctional metabolism may accelerate expansion of CH.

INTRODUCTION

Somatic mutations occur at random throughout the course of life, particularly in rapidly dividing cells. Although most somatic mutations have no consequences, some confer a competitive advantage leading to outgrowth of mutant cells, termed clonal expansion[1]. Clonal hematopoietic-driver mutations (CHDM) are somatic mutations that occur in hematopoietic stem cells (HSC) and lead to clones detectable in peripheral blood cells. CHDMs have been described in several genes, including *DNMT3A, TET2, ASXL1,* and *JAK2*[2-4]. The accumulation of somatic mutations is a hallmark of aging[5-8] and the prevalence of CHDMs increases with age[9,10]. Some studies suggest that clonal hematopoiesis may also be influenced by environmental stressors such as inflammation, and the prevalence of CHDMs is increased in people who smoke or have undergone chemo- or radiation therapy[9,11-13]. In addition, it has been suggested that cellular stressors may promote the expansion of mutant clones[14].

Individuals with CHDMs have increased risk of hematological cancers, but many CHDM carriers present without hematological diseases[1]. When the variant allele fraction (VAF) of these mutations in peripheral blood DNA is ≥2%, this state is termed clonal hematopoiesis of indeterminate potential (CHIP)[15]. Unexpectedly, the presence of CHIP has been associated with increased all-cause mortality and risk of several non-hematological diseases, including myocardial infarction, heart failure and type 2 diabetes[10,16-18]. Mechanistic studies in mice with loss of *Tet2* function in hematopoietic cells indicate that the relationship between clonal hematopoiesis and cardiovascular disease is causal and mediated by increased atherosclerosis and inflammation[17,19]. Similarly, a recent study showed that clonal hematopoiesis driven by *Tet2* mutations aggravates insulin resistance in mice[20], suggesting a causal role of clonal hematopoiesis also in the development of type 2 diabetes.

The risk of cardiometabolic diseases increases with age[21], an association that may be strengthened by obesity. Obesity accelerates the normal aging process, reducing life expectancy by 5 to 20 years[22-25]. Recent work suggests that CHDM prevalence in healthy older women is higher in those with obesity compared to those with normal body weight[26], supporting the idea that obesity accelerates the age-related increase in clonal hematopoiesis. The mechanisms that link increased prevalence of CHDMs and obesity are unknown, but metabolic abnormalities (*e.g*. dyslipidemia and glucose dysregulation) that are associated with obesity, cardiovascular disease and CHIP could play a role.

Clone size appears to correlate to risk of disease[17,27], with higher risk of both cardiovascular events and hematological malignancies in those with clones with VAF >10% compared to those with smaller clones[10]. However, due to the detection limit of many assays, the clinical implications of clonal hematopoiesis have so far mainly been studied for CHDMs with >1-2% VAF. Unresolved questions include the health consequences of smaller clones and how they evolve over time. In addition, the drivers of clonal outgrowth over time are largely unknown, *i.e.* do certain CHDMs have a survival advantage and/or are other, potentially environmental,

triggers involved. Furthermore, although it is currently assumed that most clones are stable over time in healthy individuals[28], this may not be the case for individuals with obesity.

Here we used an ultrasensitive assay to analyze clonal hematopoiesis in blood samples from individuals with obesity from the longitudinal Swedish Obese Subjects (SOS) study. Our aims were to determine the prevalence of clonal hematopoiesis driven by clones of variable sizes, and to examine the development of clones over time in relation to age and metabolic dysregulation over up to 20 years in individuals with obesity.

METHODS

Subjects and samples

The present study includes participants from the control group of the SOS study, an ongoing, prospective, controlled intervention study designed to compare outcomes in patients with obesity treated by bariatric surgery and a matched control group given usual care. Inclusion criteria were age between 37 and 60 years and body-mass index (BMI) of ≥34 for men and ≥38 for women; participants were recruited between September 1, 1987, and January 31, 2001, as previously described[29].

Seven regional ethics review boards approved the study protocol and written or oral informed content was obtained. Blood samples were taken after an overnight fast. Biochemical analyses were performed at the Central Laboratory, Sahlgrenska University Hospital, accredited according to International Organization for Standardization/International Electrochemical Commission 15189:2007 standards.

Two datasets were analyzed from the individuals given usual care: a large single-timepoint dataset and a smaller multiple-timepoint dataset with up to 5 timepoints over 20 years. For the single-timepoint dataset, 1746 DNA samples were extracted from blood samples taken at a single timepoint for each patient. Of these, we excluded 668 samples because they did not pass DNA quality control. Thus, 1078 DNA samples of sufficient quality and quantity for sequencing were available. To analyze CHDMs over time, we selected 40 individuals for whom a CHDM had been detected in the single-timepoint dataset and blood samples for DNA extraction were available from baseline and up to four additional timepoints. The resulting multiple-timepoint dataset consisted of 180 DNA samples taken at baseline and at the 2-, 10-, 15-, and/or 20-year examinations (n=40, 38, 40, 38, and 24, respectively).

CHDM detection by single-molecule molecular inversion probe (smMIP) sequencing

CHDMs were analyzed by ultra-sensitive sequencing, as essentially previously described with minor modifications[30] (**Supplementary Figure S1**). Briefly, smMIP sequencing of previously established CH-related hotspots and the entire *DNMT3A* gene was performed on DNA extracted from whole blood in batches of up to 380 samples per run. We processed two PCR and sequencing replicates for each sample from our single- (N=1078) and multiple-timepoint (N=180) datasets. We next applied two independent data-processing strategies, followed by a targeted quality control to ensure reliable somatic variant calls, which resulted in high-quality data for N=1050 (single-timepoint) and N=180 (multiple-timepoint) samples. For a more detailed description of our CHDM detection approach see **Supplementary Information**.

CHDM definitions and classifications

CHDMs were subclassified, based on the arbitrary VAF threshold to define CHIP, into small (VAF <2%) and large (≥2%) clones[2].

In our multiple-timepoint dataset, we classified CHDMs based on their evolvement during follow-up. CHDMs present in at least three timepoints for an individual were classified as traceable trajectories, whereas CHDMs only observed at a single or two timepoints were

classified as events. Traceable trajectories were then further subclassified based on clone dynamics: 1) growing trajectories – where the absolute VAF of CHDMs at the final timepoint was at least 0.5% higher than at the first timepoint, 2) shrinking trajectories – where the VAF of CHDMs at the final timepoint was at least 0.5% lower than at the first timepoint, and 3) static trajectories – where the VAF of CHDMs at the final and first timepoints differed by less than 0.5%. As our trajectory definition requires CHDMs to be present at three or more timepoints, we annotated events with a VAF ≥2% at the final timepoint as 'late-appearing clones', as the abrupt appearance of a CHDM with such a high VAF likely suggests a fast-growing clone and therefore has the potential to be a growing trajectory. Finally, we computed relative VAF metrics that define CHDM evolvement over time by comparing the VAF at each timepoint to the first timepoint at which the clone was detected.

CHDMs in literature

To characterize the CHDMs identified here, and differentiate between novel candidate and well-established driver mutations, we compared all CHDMs identified in this study with those previously reported in ten large studies together reporting 7477 coding CHDMs (**Supplementary Table S1a**). The comparison was performed for the single- and multiple-timepoint CHDMs separately.

Statistical analyses

In our cross-sectional, single-timepoint dataset, differences in age were assessed by means of Wilcoxon-rank sum tests. We subsequently used logistic regression models to determine the impact of age on CHDM prevalence (total clones, and large clones only), and a linear regression model for the effect of age on CHDM size (log-transformed VAF).

Our longitudinal, multiple-timepoint dataset enabled us to go beyond clone size and determine the effect of age on clone growth, as dependent measurements allowed for tracing a single clone over time. To this end, we first selected a single CHDM trajectory per individual. We hypothesized that trajectories with higher VAFs are more important than trajectories with lower VAFs, since literature has shown that clone size appears to be correlated to the risk of disease[17,27]. As such, we prioritized growing or shrinking trajectories over static trajectories, and selected for each individual the trajectory with the highest VAF detected at any follow-up timepoint as the 'most important' trajectory. Individuals in which the most important trajectory was shrinking were excluded from these analyses as we cannot exclude the fact that a (potentially cancer-related) mutation outside our target region may have caused the detected CHDM to shrink. We used a mixed linear model with random intercept and random slope, allowing varying effects of age per individual trajectory, to determine the effect of age on growth (VAF at each timepoint). To assess what could underlie these differences in effect size of age on growth (speed of growth), we aimed to evaluate the role of parameters related to glucose metabolism (glucose, insulin, homeostatic model assessment for insulin resistance (HOMA)-index), lipid metabolism (high-density lipoprotein cholesterol (HDL-C)), HSC proliferation (high-sensitivity C-reactive protein (hsCRP)), and obesity in general (systolic blood

pressure (SBP), body mass index (BMI), and waist-hip-ratio). We finally correlated the resulting random Beta estimates for age to clinical parameters averaged over the first three follow-up timepoints, by means of Spearman correlation coefficients, and subsequently expanded the mixed linear model with significant averaged clinical parameters.

All statistical analyses were performed in R version 3.6.1 (R Core Team, URL https://www.R-project.org/), p-values <0.05 were considered statistically significant unless otherwise specified. For more details on specific analyses see **Supplementary Information**.

RESULTS

Subjects with available sequencing data

We obtained high-quality sequencing data for 1050 individuals of our single-timepoint dataset and all 180 individual timepoints from our 20-year multiple-timepoint dataset (**Figure 1a**). The average coverage over the entire CH-smMIP panel for each of two technical replicates per sample was 2840x and 3891x for our single- and multiple-timepoint datasets, respectively (**Supplementary Figure S2**).

The age of individuals in our single-timepoint dataset at the time of DNA sampling ranged from 37.3 to 71.2 years (mean 52.4 years). Women were overrepresented (70.1% vs 29.9% men), but the age distribution was comparable in women and men (**Figure 1b**). Anthropomorphic measurements at baseline reflected the inclusion criteria for this cohort with obesity (**Figure 1b, Supplementary Table S2**).

For the 20-year multiple-timepoint dataset, more than half of the individuals (N=21) had DNA available from five timepoints (baseline/0-, 2-, 10-, 15-, 20 years), 18 individuals had DNA from four timepoints and one had DNA from three timepoints available for analysis (**Supplementary Table S3**). Women were again overrepresented (75.0% vs 25.0% men) (**Figure 1c**). BMI, serum cholesterol, and plasma glucose levels remained virtually stable over 20 years **(Figure 1c)**.

CHDM prevalence and clone size

In our single-timepoint dataset of 1050 individuals, we identified a total of 273 candidate CHDMs, consisting of 135 different mutations, in 216 individuals (**Supplementary Table S4**). A single CHDM was detected in 172 individuals (16.4%), and 44 individuals (4.2%) carried more than one CHDM (**Figure 2a**). The VAF of all 273 candidate CHDMs ranged from 0.01% to 31.15%, with a mean of 2.72% and median of 0.70% (**Figure 2b**). Seventy out of 273 CHDMs had a VAF ≥2%, the arbitrary threshold defining mutations as CHIP. Thus, the sensitivity of our smMIP assay enabled the detection of 203 CHDMs with a VAF <2% (**Figure 2b**). We observed an increasing prevalence of CHDMs in older age categories, and mutations in *DNMT3A* were the most common in all age categories (**Figure 2c**). CHDM carriers were significantly older than non-carriers (54.6 and 51.9 years, respectively, $P=2.43E^{-06}$) (**Figure 2d**). The prevalence of both small (VAF <2%) and large clones (VAF ≥2%) increased with age, and the proportion of large clones was highest in the oldest age category (**Figure 2e**). Accordingly, the mean age was significantly higher in large- versus small-clone carriers (57.2 and 53.4 years, respectively, $P=2.36E^{-04}$) (**Figure 2f**). The association between clone prevalence and age was stronger when only large clones were considered (odds ratio $(OR)_{age}=1.05$, 95% confidence interval (CI) [1.03-1.08] and $OR_{age}=1.11$, 95% CI [1.07-1.16], as determined by logistic regression models for total clones and large clones, respectively) (**Figure 2g**). By means of linear regression, we also observed a significant association of age with clone size (log-transformed VAF) (**Figure 2h**), with an effect estimate of 0.059 ($P=2.58E^{-05}$), which translates to a 6% increase in clone size per year.

Evolution of CHDMs over up to 20 years

A multiple-timepoint dataset was created by selecting 40 individuals in whom we identified a CHDM in the single-timepoint dataset and for whom blood samples for DNA extraction were available at three or more timepoints. We obtained a total of 180 DNA samples collected at up to five timepoints over 20 years (0-, 2-, 10-, 15-, 20-years, see **Supplementary Table S3** for more details). The mean VAF of all detected CHDMs increased with follow-up time (**Figure 3a**), supporting our observation in the single-timepoint dataset that clone size positively correlates with age.

In the multiple-timepoint dataset we identified 115 CHDMs, comprising 53 unique mutations with up to 6 mutations per individual (**Supplementary Table S6**). Out of the 115 CHDMs, 38 were categorized as events and 77 were defined as traceable trajectories, with 22 individuals carrying more than one trajectory (see Methods for details on these definitions) (**Figure 3b**). We then inspected the dynamics of all 77 trajectories, by comparing the VAF at first and last timepoint of detection, and identified 30 growing, 42 static, and 5 shrinking trajectories (**Figure 3c**). The 30 growing trajectories were detected in 21 (52.5%) individuals, carrying mutations in *DNMT3A*, *CBL*, *GNAS*, *JAK2*, *TET2*, *GNB1*, and *IDH2* (**Figure 3c**). For 24 out of the 30 growing trajectories, the CHDM was initially detected as a small clone (VAF <2%); nine of these 24 had become large clones (VAF ≥2%, CHIP) at the last timepoint of detection. For the remaining six growing trajectories, the clone was already above the CHIP level at the first timepoint it was detected. In addition, there were three late-appearing clones in three individuals – CHDMs that appeared at the last measured timepoint as a large clone (VAF ≥2%). The 42 static CHDM trajectories were detected in 24 individuals, carrying mutations in *ASXL1*, *BRCC3*, *DNMT3A*, *GNAS*, *GNB1*, *IDH2*, *NRAS*, *TET2*, and *TP53* (**Figure 3c**). With the exception of one timepoint in one individual, all static CHDM trajectories were below the CHIP level. The five shrinking CHDM trajectories were detected in five individuals, carrying mutations in *DNMT3A*, *GNB1*, and *JAK2* (**Figure 3c**). Three started as large clones, two of which shrunk below CHIP.

By comparing all detected CHDMs per individual and between individuals, we observed that the evolution of CHDMs over time varied substantially both between individuals and mutations (**Supplementary Figure S4**). **Figure 3d** shows examples of CHDM evolution over 20 years for three individuals. Individual 12 had six different clones, none of which seemed to have a clear advantage over the others. In contrast, for individual 19, *GNB1* [p.(Lys57Glu)] appeared to have a growth advantage, as *DNMT3A* [p.(Asp702Glu)] started shrinking when *GNB1* [p.(Lys57Glu)] expanded. In individual 14, we detected four CHDMs, two of which (*TET2* [p.(Glu873*)] and *DNMT3A* [p.(Phe732Ile)]) grew to a VAF of almost 40% after 20 years, but were far below the CHIP level at the first time of detection.

CHDM characteristics and comparison to well-established driver mutations

As expected and in line with previous reports, *DNMT3A* was the most frequently mutated gene in both our single- and multiple-timepoint dataset (**Figure 4**). We also identified several loss of function (LoF) mutations in *TET2* and *ASXL1*, and frequently recurring (*i.e.*

identified each at least five times in the single-timepoint data of our cohort) missense mutations in *GNAS* [p.(Cys843Arg) and p.(Arg844His)], *GNB1* [p.(Lys57Glu)], *TP53* [p.(Arg174Gly)], and *JAK2* [p.(Val617Phe)]. The majority of CHDMs in the single-timepoint dataset involved the same genes as the growing clones from the multiple-timepoint dataset (**Figure 4a and b**). The amino acid residues most frequently affected by CHDMs in the single-timepoint dataset were similarly enriched in trajectories that grew in our multiple-timepoint dataset (**Figure 4c**).

Our full gene coverage of *DNMT3A* allowed us to examine the location of CHDMs in this gene. We observed that *DNMT3A* missense mutations detected in our single-timepoint dataset clustered around the three known protein domains: PWWP, ADD and SAM-dependent MTase C5-type (**Figure 4d**). In our 20-year longitudinal dataset, all LoF mutations in *DNMT3A* were defined as traceable trajectories apart from one late-appearing clone, whereas one third of *DNMT3A* missense mutations were classified as events and two thirds as trajectories (**Figure 4e**). Of all the LoF mutations in *DNMT3A*, five out of six early stops or frameshifts (before amino acid 432) were classified as growing (four trajectories and one late-appearing clone), while none of the six later LoF mutations fit this classification. **Figure 4f** shows that the clustering of CHDMs in *DNMT3A* previously reported in the literature is comparable to that in both of our datasets, substantiating their role as established driver mutations.

The majority of mutations identified in this study have previously been described (**Supplementary Table S1a**). Of all the 273 detected single-timepoint mutations, 226 (82.8%) were identical to previously identified CHDMs from the literature or showed an LoF mutation in *DNTM3A, TET2* or *ASXL1*. Of those 226, 61.9% were well-established drivers (≥5 counts in literature) and 10.3% were new substitutions at previously described amino acid positions; the remaining 7.0%, corresponding to 5 different mutations, were therefore novel candidate CHDMs (**Supplementary Table S1b**).

Of all the 115 mutations identified in the multiple-timepoint measurements, 71 (61.7%) have been reported in the literature or represent LoF mutations in *DNMT3A, TET2* or *ASXL1,* this refers to 39 (73.58%) when only counting different mutations. Amongst the CHDMs classified as events (*i.e.* mutations only seen at one or two of the timepoints), the overlap with the literature was lowest (12/38; 31.6%). In contrast, the overlap was higher for all CHDMs classified as trajectories: static trajectories (27/42; 64.3%) shrinking trajectories (5/5; 100%) and growing trajectories (27/30; 90.0%) (**Supplementary Table S1c**).

Clinical correlations of CHDMs over time

Our multiple-timepoint dataset enabled us to go beyond clone size and examine a potential association between age and actual clone growth. As some individuals carried multiple trajectories over the course of 20 years (**Supplementary Figure S4**), we first selected the most important trajectory per individual, prioritising growing clones and clones with largest VAF at any timepoint (see Methods section for further details). We excluded individuals in whom the most important trajectory was shrinking (n=5), as we aimed to identify factors associated with clone growth. Because we expected variation in growth patterns in our 32

selected trajectories (**Supplementary Figure S5**), we used a mixed linear model including random intercept and random slope, to determine the association between age and clonal growth. This model outperformed alternatives with fixed effect estimates as indicated by the Akaike information criterium (AIC) (**Supplementary Table S6**). We identified an average proportionate increase of 7% in VAF per year, ranging from -4% to 27% (**Figure 5a**), confirming the expected differences in speed of growth per trajectory. These inter-individual differences of clone growth were also observed for five identical mutations in different individuals (**Supplementary Figure S6**). For example, **Figure 5b** shows that trajectories of *DNMT3A* [p.(Arg882Cys)], identified in four individuals, and *DNMT3A* [p.(Arg882His)], identified in two individuals, increased with different speeds. We next examined if clinical characteristics were associated with these different speeds of growth, as expressed by the effect of age, by correlating all individual trajectory effect estimates from our mixed linear model to metabolic clinical parameters averaged over the first three follow-up timepoints (**Figure 5c**). We identified significant positive correlations for speed of growth with insulin (Spearman R=0.42, P=0.018) and with insulin resistance measured by HOMA-index (Spearman R=0.40, P=0.025), whereas HDL-C negatively correlated to the speed of growth (Spearman R=-0.68, P=$1.74E^{-05}$) (**Figure 5d-f**).

DISCUSSION

In this study we generated and analysed single- and multiple-timepoint (longitudinal) data on CHDMs in middle-aged individuals with obesity. Approximately 20% of individuals from our cohort had CHDMs detectable by an ultrasensitive assay designed to target known and novel candidate driver mutations. Both CHDM prevalence and clone size increased with age. Importantly, a significant fraction of clones grew, including several initially small clones. In addition to the age-related increase in clone size, clone growth was further accelerated in individuals with metabolic dysfunction.

The importance of clonal hematopoiesis has mainly been studied for mutations above the CHIP level of 2% because of: (1) methodological difficulties in detecting small clones and (2) the assumption that clones must reach a critical size to be clinically meaningful[2]. CHIP mutations are associated with a wide range of diseases, including hematological malignancies and cardiovascular disease[27]. However, the health consequences of smaller clones are largely unknown and it is unclear how mutations with low VAF evolve over time. Our ultrasensitive assay allowed us to detect small clones, and almost 75% of all CHDMs in the single-timepoint data were below the CHIP level definition. Approximately 80% of all CHDMs (both total and small clones) that we found overlapped with CHDMs that have been described in the literature, indicating that the higher sensitivity of our assay did not result in increased false positive findings. In addition, we observed an age-related increase in the prevalence of both small and large clones, and repeated measurements over two decades showed that small clones may grow beyond the CHIP level. This indicates that small clones are also linked to aging and possibly age-related phenotypes, and that with time small clones may become large enough to become clinically relevant. The importance of small clones is supported by a recent study in which VAFs significantly lower than the CHIP level for *DNMT3A* and *TET2* mutations were associated with impaired survival in patients with chronic ischemic heart failure[31].

CHDM prevalence comparisons are heavily dependent on the technology applied[32]. The ultra-sensitive assay used in this study detects smaller clones as well as additional mutations in the full coding sequence of *DNMT3.* Our previous work using an assay similar to that used in the current study indicated an overall prevalence of CHDMs of 9.6% in 2,000 individuals from the general Dutch population with an average age of 45 years, while a subgroup of 400 individuals with age similar to the current cohort had a prevalence of 13.2%[30]. The 20% CHDM prevalence in the single-timepoint dataset, with an average age of 53 years, is high compared to most other reports[2,32]. This suggests that clonal hematopoiesis is increased in individuals with obesity, which is in line with one recent study[26]. Obesity is a complex phenotype, often associated with a wide range of metabolic abnormalities and diseases, accelerated aging and shortened lifespan. Accumulating evidence shows that there must be factors beyond traditional risk factors that contribute to the link between obesity and cardiovascular disease[1,2]. It is therefore tempting to speculate that clonal hematopoiesis could in part explain increased cardiovascular disease and decreased life expectancy in people with obesity.

For the CHDMs identified in the single-timepoint dataset, the clone size was on average 6% larger for each year of age. However, a unique feature of our study was the collection of up to five samples taken over 20 years from 40 well-characterized individuals, which allowed measurement of clonal evolution over time and determination of actual clone growth. When we assessed growth of specific clones within individuals, the average annual growth was 7%. However, there were large inter-individual differences in the speed of clone growth, also in individuals sharing the same CHDM, suggesting that factors other than the mutation itself influence clone growth.

More than half of the individuals in the multiple-timepoint dataset, who were selected on the basis that clones were present, had growing clones, which is an unexpectedly high proportion[28,33,34]. This could be explained by the long timeframe of tracing and/or by the severe obesity and obesity-related metabolic abnormalities in our cohort, in contrast to previous longitudinal studies of mainly healthy individuals. Importantly, nine out of the 30 growing trajectories identified in our study were initially small clones that grew over the sampling period to beyond the CHIP level. Our longitudinal data thus showed that clones can be detected many years before they reach the CHIP level.

Our longitudinal observations indicated that increased speed of clone growth was associated not only with age but also with insulin resistance and low HDL-C, both well-established risk factors for cardiovascular disease[35]. These findings are of interest in light of the recent observation that humans with atherosclerosis have increased HSC division rates[36]. Increased HSC proliferation may promote clonal hematopoiesis both by increasing the risk of acquiring CHDMs and by facilitating the expansion of mutant clones[36]. Animal studies indicate that HDL-C and the adenosine triphosphate–binding cassette (ABC) cholesterol efflux transporter ABCG1, which promotes cholesterol efflux from macrophages to HDL-C, inhibit HSC proliferation and thereby reduce circulating numbers of leukocytes[37]. A recent Mendelian randomization study in humans supports a causal inverse relationship between HDL-C and leukocyte counts[38]. Similarly, insulin resistance has been associated with high white blood cell count in humans[39-41]. It is therefore possible that the observed association between metabolic dysfunction and clone growth is linked to increased HSC proliferation. A combination of our results and earlier studies showing that loss of *Tet2* function in HSCs leads to increased atherosclerosis[17,19] and insulin resistance[20] suggest that a vicious cycle might exist, whereby dysfunctional metabolism increases the risk of developing pro-inflammatory CHIP that, once formed, further increases the risk of worsening insulin resistance and atherosclerosis.

The increased risk associated with CHIP for several severe diseases and overall mortality raises the question of how this risk could be mitigated. In individuals with established CHIP, interventions may target the pro-inflammatory activity of CHIP[18]. Mice with experimental CHIP show increased pro-inflammatory drive of the NLRP3/IL-1β/IL-6 axis and atherosclerosis[2,42],

and humans carrying a genetic deficiency in *IL6* appear to have reduced cardiovascular risk related to CHIP[43]. In addition, treatments that increase HDL-C and/or reduce insulin resistance may potentially target the risk associated with CHIP by preventing expansion of clones and formation of pro-inflammatory CHIP.

Limitations of this study include the use of a targeted assay which may miss mutations in other genes or outside targeted hotspots. Furthermore, our study is limited by the lack of individuals with normal body weight for comparison. Strengths of this study are the ability to detect small clones down to VAF 0.01%, and the longitudinal measurements in general - which allowed us to examine how CHDMs develop over up to 20 years, and to examine the association with clinical characteristics.

In conclusion, this study showed that approximately one in five middle-aged individuals with obesity in our cohort had CHDMs, and that a proportion of these clones expanded to become CHIP, mutations that have been linked to hematological cancers and cardiovascular and metabolic diseases. We also demonstrated that accelerated growth of clones was associated with insulin resistance and low HDL-C, suggesting biological mechanisms that may open up opportunities for strategies to reduce the risk of clones expanding to CHIP level.

DATA AVAILABILITY STATEMENT

The data is subject to legal restrictions according to national legislation. Confidentiality regarding personal information in studies is regulated in the Public Access to Information and Secrecy Act (SFS 2009:400), OSL. A request to get access to public documents can be rejected or granted with reservations by the University of Gothenburg. If University of Gothenburg refuses to disclose the documents the applicant is entitled to get a written decision that can be appealed to the administrative court of appeal.

ACKNOWLEDGEMENTS

Supported by grants from the Swedish Research Council (2017-01707), the Swedish state under an agreement between the Swedish government and the county councils, the ALF (Avtal om Läkarutbildning och Forskning) agreement (ALFGBG-717881 and ALFGBG-717891), the Swedish Diabetes Foundation (2019- 417), the Swedish Heart-Lung Foundation (20180410), and the Novo Nordisk Foundation (NNF 19OC0057184).
Ms. M. Steehouwer and Dr H.G. Brunner, and Dr A. Hoischen are supported by the Solve-RD project. The Solve-RD project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 779257. This research was part of the Netherlands X-omics Initiative and partially funded by NWO (The Netherlands Organization for Scientific Research; project 184.034.019). Dr M.G. Netea was supported by an European Research Council Advanced Grant (No. 833247) and a Spinoza Grant of the

Netherlands Organization for Scientific Research. We thank the Radboud Genomics Technology Center, Radboud University Medical Center Nijmegen, for their technical assistance.

AUTHOR CONTRIBUTION

BC, AH and LMSC conceptualised the project, RCvD, JCA-A, FMK, MS, KS, P-AS, MPi, CG, MT, PJ, RP, HGB, MGN, MPe, BC, AH, LMSC contributed intellectually to the study and the preparation of the manuscript. RCvD, JCA-A, FMK, AH and LMSC wrote the first draft of the manuscript and all authors contributed to data interpretation, discussion and revised the manuscript. JCA-A, KS, P-AS, MT, PJ and LMSC were responsible for collection of data and samples. RCvD, MS and AH generated and analysed sequencing data. RCvD and MPe did the statistical analyses. All authors approved the final manuscript.

CONFLICTS OF INTEREST

Dr B. Carlsson is employed by and owns stock in AstraZeneca. Dr L. Carlsson has received consulting fees from Johnson & Johnson. Dr M.G. Netea reported being a scientific founder of Trained Therapeutic Discovery and receiving grants from ViiV HealthCare outside the submitted work. No other potential conflict of interest relevant to this article was reported.

REFERENCES

1 Fuster, J. J. & Walsh, K. Somatic Mutations and Clonal Hematopoiesis: Unexpected Potential New Drivers of Age-Related Cardiovascular Disease. *Circ Res* **122**, 523-532, doi:10.1161/CIRCRESAHA.117.312115 (2018).

2 Jaiswal, S. & Ebert, B. L. Clonal hematopoiesis in human aging and disease. *Science* **366**, doi:10.1126/science.aan4673 (2019).

3 Steensma, D. P. Clinical consequences of clonal hematopoiesis of indeterminate potential. *Blood Adv* **2**, 3404-3410, doi:10.1182/bloodadvances.2018020222 (2018).

4 Steensma, D. P. & Ebert, B. L. Clonal hematopoiesis as a model for premalignant changes during aging. *Exp Hematol* **83**, 48-56, doi:10.1016/j.exphem.2019.12.001 (2020).

5 Acuna-Hidalgo, R., Veltman, J. A. & Hoischen, A. New insights into the generation and role of de novo mutations in health and disease. *Genome Biol* **17**, 241, doi:10.1186/s13059-016-1110-1 (2016).

6 Forsberg, L. A., Gisselsson, D. & Dumanski, J. P. Mosaicism in health and disease - clones picking up speed. *Nat Rev Genet* **18**, 128-142, doi:10.1038/nrg.2016.145 (2017).

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

8 Vijg, J. Somatic mutations, genome mosaicism, cancer and aging. *Curr Opin Genet Dev* **26**, 141-149, doi:10.1016/j.gde.2014.04.002 (2014).

9 Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* **371**, 2477-2487, doi:10.1056/NEJMoa1409405 (2014).

10 Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* **371**, 2488-2498, doi:10.1056/NEJMoa1408617 (2014).

11 Bick, A. G. *et al.* Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763-768, doi:10.1038/s41586-020-2819-2 (2020).

12 Coombs, C. C. *et al.* Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell* **21**, 374-382 e374, doi:10.1016/j.stem.2017.07.010 (2017).

13 Zink, F. *et al.* Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood* **130**, 742-752, doi:10.1182/blood-2017-02-769869 (2017).

14 Wong, T. N. *et al.* Cellular stressors contribute to the expansion of hematopoietic clones of varying leukemic potential. *Nat Commun* **9**, 455, doi:10.1038/s41467-018-02858-0 (2018).

15 Steensma, D. P. *et al.* Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* **126**, 9-16, doi:10.1182/blood-2015-03-631747 (2015).

16 Bonnefond, A. *et al.* Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. *Nat Genet* **45**, 1040-1043, doi:10.1038/ng.2700 (2013).

17 Jaiswal, S. *et al.* Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med* **377**, 111-121, doi:10.1056/NEJMoa1701719 (2017).

18 Jaiswal, S. & Libby, P. Clonal haematopoiesis: connecting ageing and inflammation in cardiovascular disease. *Nat Rev Cardiol* **17**, 137-144, doi:10.1038/s41569-019-0247-5 (2020).

19 Wang, Y. *et al.* Tet2-mediated clonal hematopoiesis in nonconditioned mice accelerates age-associated cardiac dysfunction. *JCI Insight* **5**, doi:10.1172/jci.insight.135204 (2020).
20 Fuster, J. J. *et al.* TET2-Loss-of-Function-Driven Clonal Hematopoiesis Exacerbates Experimental Insulin Resistance in Aging and Obesity. *Cell Rep* **33**, 108326, doi:10.1016/j.celrep.2020.108326 (2020).
21 Tzanetakou, I. P., Katsilambros, N. L., Benetos, A., Mikhailidis, D. P. & Perrea, D. N. "Is obesity linked to aging?": adipose tissue and the role of telomeres. *Ageing Res Rev* **11**, 220-229, doi:10.1016/j.arr.2011.12.003 (2012).
22 Bhaskaran, K., Dos-Santos-Silva, I., Leon, D. A., Douglas, I. J. & Smeeth, L. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3.6 million adults in the UK. *Lancet Diabetes Endocrinol* **6**, 944-953, doi:10.1016/S2213-8587(18)30288-2 (2018).
23 Carlsson, L. M. S. *et al.* Life Expectancy after Bariatric Surgery in the Swedish Obese Subjects Study. *N Engl J Med* **383**, 1535-1543, doi:10.1056/NEJMoa2002449 (2020).
24 Prospective Studies, C. *et al.* Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* **373**, 1083-1096, doi:10.1016/S0140-6736(09)60318-4 (2009).
25 Valdes, A. M. *et al.* Obesity, cigarette smoking, and telomere length in women. *Lancet* **366**, 662-664, doi:10.1016/S0140-6736(05)66630-5 (2005).
26 Haring, B. *et al.* Healthy Lifestyle and Clonal Hematopoiesis of Indeterminate Potential: Results From the Women's Health Initiative. *J Am Heart Assoc* **10**, e018789, doi:10.1161/JAHA.120.018789 (2021).
27 Silver, A. J. & Jaiswal, S. Clonal hematopoiesis: Pre-cancer PLUS. *Adv Cancer Res* **141**, 85-128, doi:10.1016/bs.acr.2018.12.003 (2019).
28 Young, A. L., Challen, G. A., Birmann, B. M. & Druley, T. E. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun* **7**, 12484, doi:10.1038/ncomms12484 (2016).
29 Sjostrom, L. *et al.* Effects of bariatric surgery on mortality in Swedish obese subjects. *N Engl J Med* **357**, 741-752, doi:10.1056/NEJMoa066254 (2007).
30 Acuna-Hidalgo, R. *et al.* Ultra-sensitive Sequencing Identifies High Prevalence of Clonal Hematopoiesis-Associated Mutations throughout Adult Life. *Am J Hum Genet* **101**, 50-64, doi:10.1016/j.ajhg.2017.05.013 (2017).
31 Assmus, B. *et al.* Clonal haematopoiesis in chronic ischaemic heart failure: prognostic role of clone size for DNMT3A- and TET2-driver gene mutations. *Eur Heart J* **42**, 257-265, doi:10.1093/eurheartj/ehaa845 (2021).
32 Watson, C. J. *et al.* The evolutionary dynamics and fitness landscape of clonal hematopoiesis. *Science* **367**, 1449-1454, doi:10.1126/science.aay9333 (2020).
33 van den Akker, E. B. *et al.* Dynamic clonal hematopoiesis and functional T-cell immunity in a supercentenarian. *Leukemia*, doi:10.1038/s41375-020-01086-0 (2020).
34 Young, A. L., Tong, R. S., Birmann, B. M. & Druley, T. E. Clonal hematopoiesis and risk of acute myeloid leukemia. *Haematologica* **104**, 2410-2417, doi:10.3324/haematol.2018.215269 (2019).
35 Kopin, L. & Lowenstein, C. Dyslipidemia. *Ann Intern Med* **167**, ITC81-ITC96, doi:10.7326/AITC201712050 (2017).
36 Heyde, A. *et al.* Increased stem cell proliferation in atherosclerosis accelerates clonal hematopoiesis. *Cell* **184**, 1348-1361 e1322, doi:10.1016/j.cell.2021.01.049 (2021).

37 Yvan-Charvet, L. *et al.* ATP-binding cassette transporters and HDL suppress hematopoietic stem cell proliferation. *Science* **328**, 1689-1693, doi:10.1126/science.1189731 (2010).
38 Harslof, M., Pedersen, K. M., Nordestgaard, B. G. & Afzal, S. Low HDL (High-Density Lipoprotein) Cholesterol and High White Blood Cell Counts: A Mendelian Randomization Study. *Arterioscler Thromb Vasc Biol*, ATVBAHA120314983, doi:10.1161/ATVBAHA.120.314983 (2020).
39 Facchini, F., Hollenbeck, C. B., Chen, Y. N., Chen, Y. D. & Reaven, G. M. Demonstration of a relationship between white blood cell count, insulin resistance, and several risk factors for coronary heart disease in women. *J Intern Med* **232**, 267-272, doi:10.1111/j.1365-2796.1992.tb00582.x (1992).
40 Kuo, T. Y. *et al.* Relationships between white blood cell count and insulin resistance, glucose effectiveness, and first- and second-phase insulin secretion in young adults. *Medicine (Baltimore)* **99**, e22215, doi:10.1097/MD.0000000000022215 (2020).
41 Piedrola, G., Novo, E., Escobar, F. & Garcia-Robles, R. White blood cell count and insulin resistance in patients with coronary artery disease. *Ann Endocrinol (Paris)* **62**, 7-10 (2001).
42 Fuster, J. J. *et al.* Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science* **355**, 842-847, doi:10.1126/science.aag1381 (2017).
43 Bick, A. G. *et al.* Genetic Interleukin 6 Signaling Deficiency Attenuates Cardiovascular Risk in Clonal Hematopoiesis. *Circulation* **141**, 124-131, doi:10.1161/CIRCULATIONAHA.119.044362 (2020).

MAIN FIGURE LEGENDS

Figure 1. Flowchart of CHDM identification in single- and multiple-timepoint datasets. (**a**) Single-molecule molecular inversion probes (smMIPs) were used for targeted sequencing of 24 CH-related genes/hotspots to enable reliable detection of CHDMs. For more details, see **Supplementary Figure S1**. (**b**) Our single-timepoint dataset consisted of 1050 individuals (70.1% women and 29.9% men) with an average age of 52.4 years and average BMI of 40.4 kg/m^2. (**c**) Our multiple-timepoint dataset consisted of 40 individuals (75.0% women and 25.0% men) for whom a CHDM had been detected in the single-timepoint dataset in combination with the availability of sufficient follow-up blood samples. The average age was 48.6 years at baseline and 67.5 at the last follow-up. BMI, serum cholesterol and plasma glucose remained virtually unchanged during follow-up.

Figure 2. Single-timepoint CHDM prevalence and clone size in relation to age. (**a**) Percentage of individuals in our single-timepoint dataset (total n = 1050) with no, a single, two, three or four CHDMs. (**b**) The variant allele frequency (VAF) of identified small (VAF <2%) and large (VAF ≥2%) CHDMs in our single-timepoint dataset. (**c**) Total CHDM prevalence and the affected genes per age category. (**d**) Age of non-carriers and CHDM carriers. The black bars and annotated numbers show the mean values; the Wilcoxon-rank-sum test P-value is shown at the top of the plot. (**e**) Prevalence of small and large clones per age category. (**f**) Age of small-clone carriers and large-clone carriers. The black bars and annotated numbers show the mean values; the Wilcoxon-rank-sum test P-value is shown at the top of the plot. (**g**) Visualization of probability curve of logistic regression models using all clones (total clone in green) or large clones only (in dark blue) as dependent variable, and age as predictor for both models. The odds ratios (ORs) with accompanying 95% confidence intervals for both models are annotated in the plot. For details on model parameters, see **Supplementary Table S5a and S5b**, and model fit, see **Supplementary Figure S3.** (**h**) Visualization of linear regression model using clone size of all CHDMs (log-transformed VAF) as dependent variable and age as predictor. For details on model parameters, see **Supplementary Table S5c**.

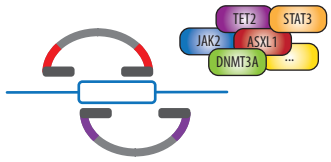
Figure 3. Clone size of CHDMs at multiple time points over 20 years in individuals with obesity. (**a**) The variant allele frequency (VAF) of all detected CHDMs per follow-up timepoint, with mean and median VAF indicated. (**b**) Relative changes in clone size of CHDMs over 20 years and according to chronological age. Trajectories are connected by red, green and blue lines indicating growing, static and shrinking trajectories, respectively. Single- or multiple-timepoint events are shown in yellow, a dashed circle indicates late-appearing clones. (**c**) Categorization of CHDMs into events (single or multiple) and trajectories. Trajectories are further categorized into growing, static and shrinking, and their VAF as well as the proportion of genes in which these trajectories were detected are shown. (**d**) Examples of CHDM evolution with follow-up time in three individuals with obesity.

Figure 4. CHDMs identified in the single- and multiple timepoint datasets and comparison to literature. (**a**) The gene distribution of identified CHDMs in the single-timepoint dataset (total number of CHDMs=273). (**b**) The gene distribution of growing CHDM trajectories in the multiple-timepoint dataset (total number of growing trajectories=30). (**c**) The number of CHDMs resulting in specific amino acid changes in the single-timepoint dataset at the protein level. Labels indicate CHDMs that have been defined as growing trajectories in our multiple-timepoint dataset. (**d**) Amino acid changes caused by *DNMT3A* mutations in our single-timepoint dataset annotated on a schematic visualization of the DNMT3A protein. Mutations affecting the same amino acid residue five or more times have been annotated and the label specifies the reference amino acid, position and number of mutations in superscript. Labels: R326$^{(6)}$ = R326H (n=3), R326C (n=3); R635$^{(5)}$ = R635W (n=3), R635Q (n=1), R635L (n=1); R729$^{(12)}$ = R729W (n=9), R729Q (n=1), R729G (n=1), R729L (n=1); Y735$^{(10)}$ = Y735C (n=9), Y735S (n=1); R736$^{(12)}$ = R736C (n=5), R736H (n=5), R736G (n=2); D768$^{(10)}$ = D768G (n=9), D768V (n=1); R771$^{(6)}$ = R771stop (n=6); R882$^{(21)}$ = R882H (n=11), R882C (n=8), R882S (n=2). (**e**) Amino acid changes caused by *DNMT3A* mutations in our multiple-timepoint dataset annotated on a schematic visualization of DNMT3A protein. CHDMs that have been categorized as trajectories have been given a coloured dot: red for growing (with a dashed yellow circle for 'late-appearing clone'), green for static and blue for shrinking. Labels constitute the reference amino acid in single letter code, position, and alternative amino acid, and are provided only for growing and shrinking trajectories (as represented by the number in superscript), and not for static trajectories or events (with the exception of 'late-appearing clone'). (**f**) *DNMT3A* mutations (amino acid changes) observed at least five times in literature annotated on a schematic visualization of DNMT3A protein. Missense mutations observed >25 times, and LoF mutations observed >10 times have been annotated and the label specifies the reference amino acid, position and number of calls in superscript. For a complete table of all CHDMs observed in literature, see **Supplementary Table S1a**. DNMT3A protein annotation source: https://www.uniprot.org/uniprot/Q9Y6K1.

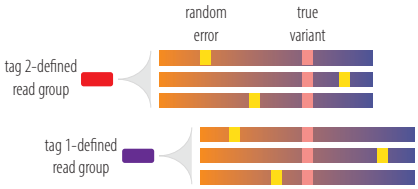
Figure 5. Speed of CHDM growth and clinical associations. (**a**) VAF in relation to age for the 32 most important trajectories. The average mixed linear model (MLM) regression line and equation is shown in black. All individual MLM regression lines are shown in various colours, in purple the trajectory with maximum slope and accompanying equation, in orange the trajectory with minimum slope and accompanying equation. For details on model parameters see **Supplementary Table S7a**. (**b**) VAFs and slopes for CHDMs in *DNMT3A* [p.(Arg822Cys)] in 4 individuals and *DNMT3A* [p.(Arg882His)] in 2 individuals. (**c**) Heatmap of Spearman R correlations between all individual trajectory effect estimates from our MLM to averaged (based on the first three follow-up timepoints) metabolic clinical parameters. (**d**) Positive Spearman R correlation between individual trajectory effect estimates and insulin. (**e**) Positive Spearman R correlation between individual trajectory effect estimates and HOMA-index. (**f**) Negative Spearman R correlation between individual trajectory effect estimates and HDL-C.

a.

single-molecule Molecular Inversion Probe (smMIP) sequencing

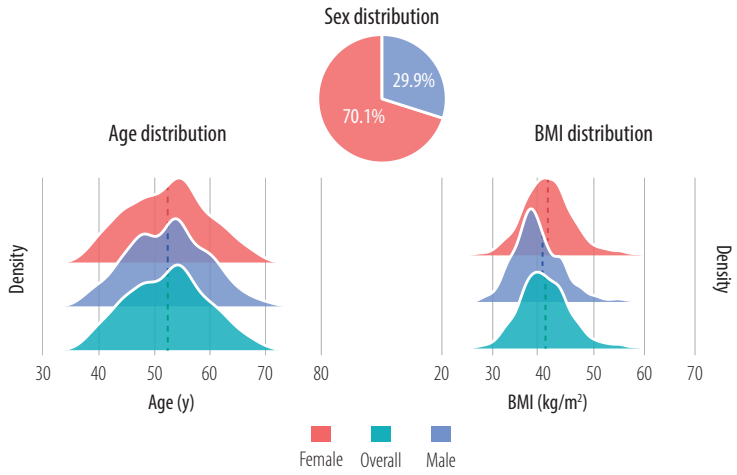


smMIP sequencing of CH-related genes/hotspots



smMIP consensus calling

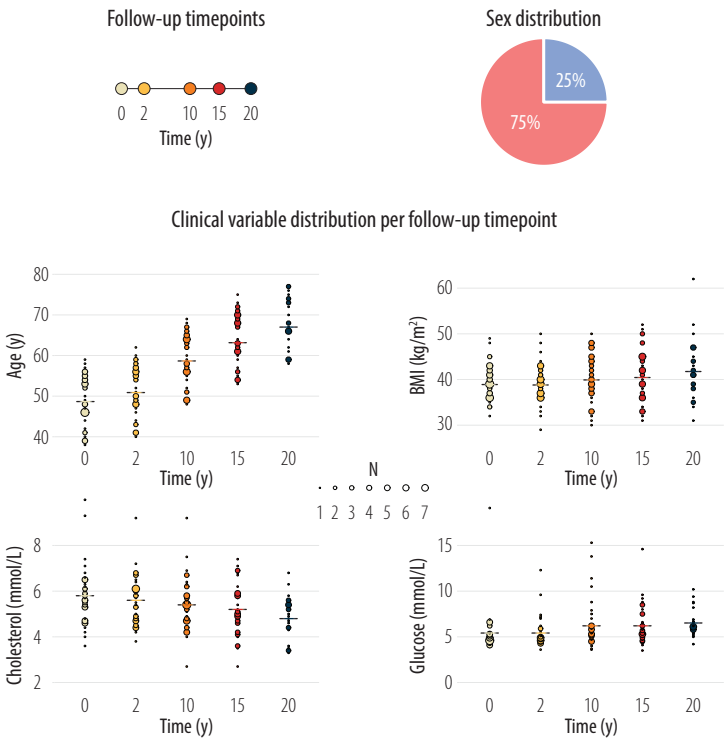
b.



Single-timepoint data

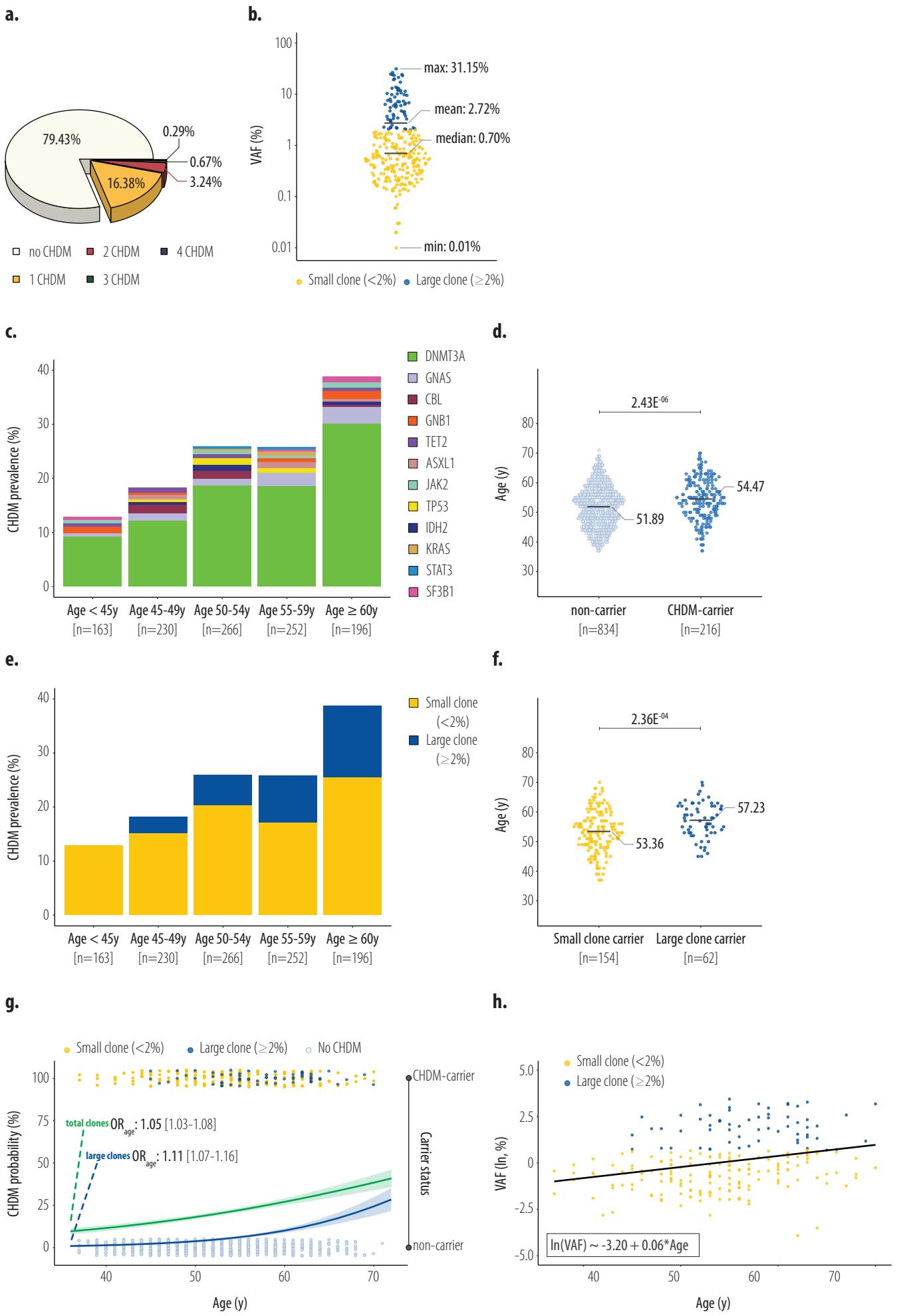
N individuals = 1050

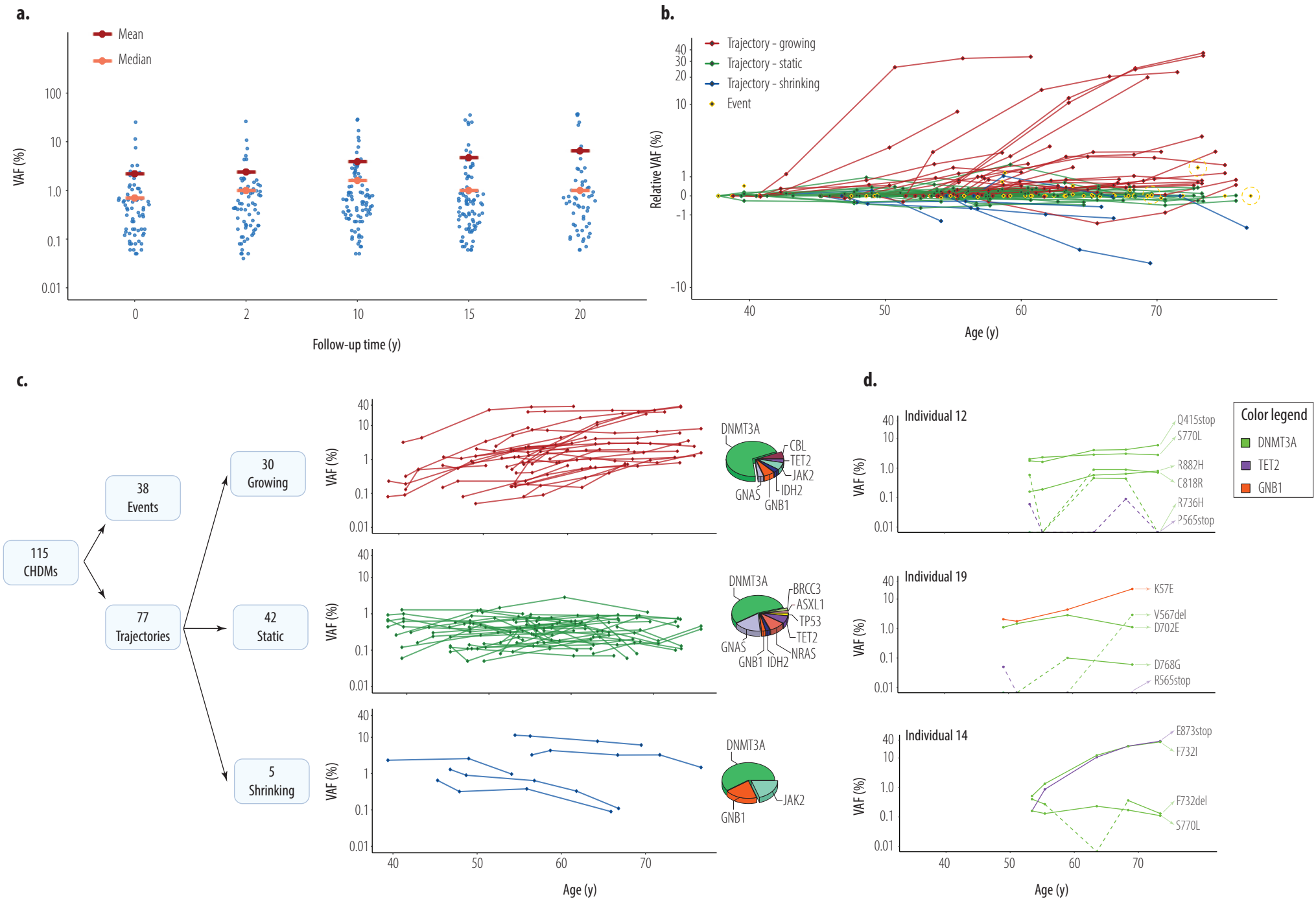
c.

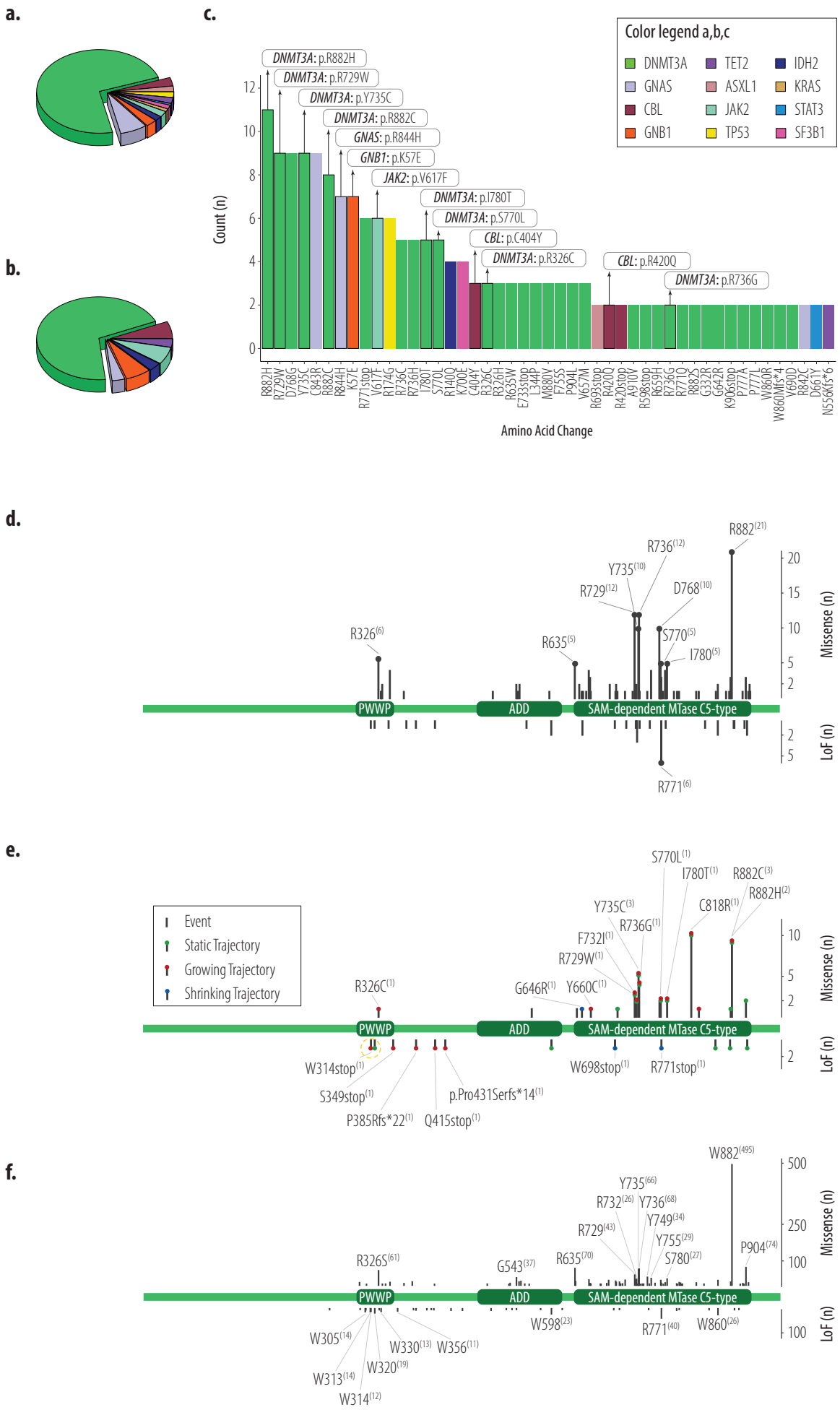


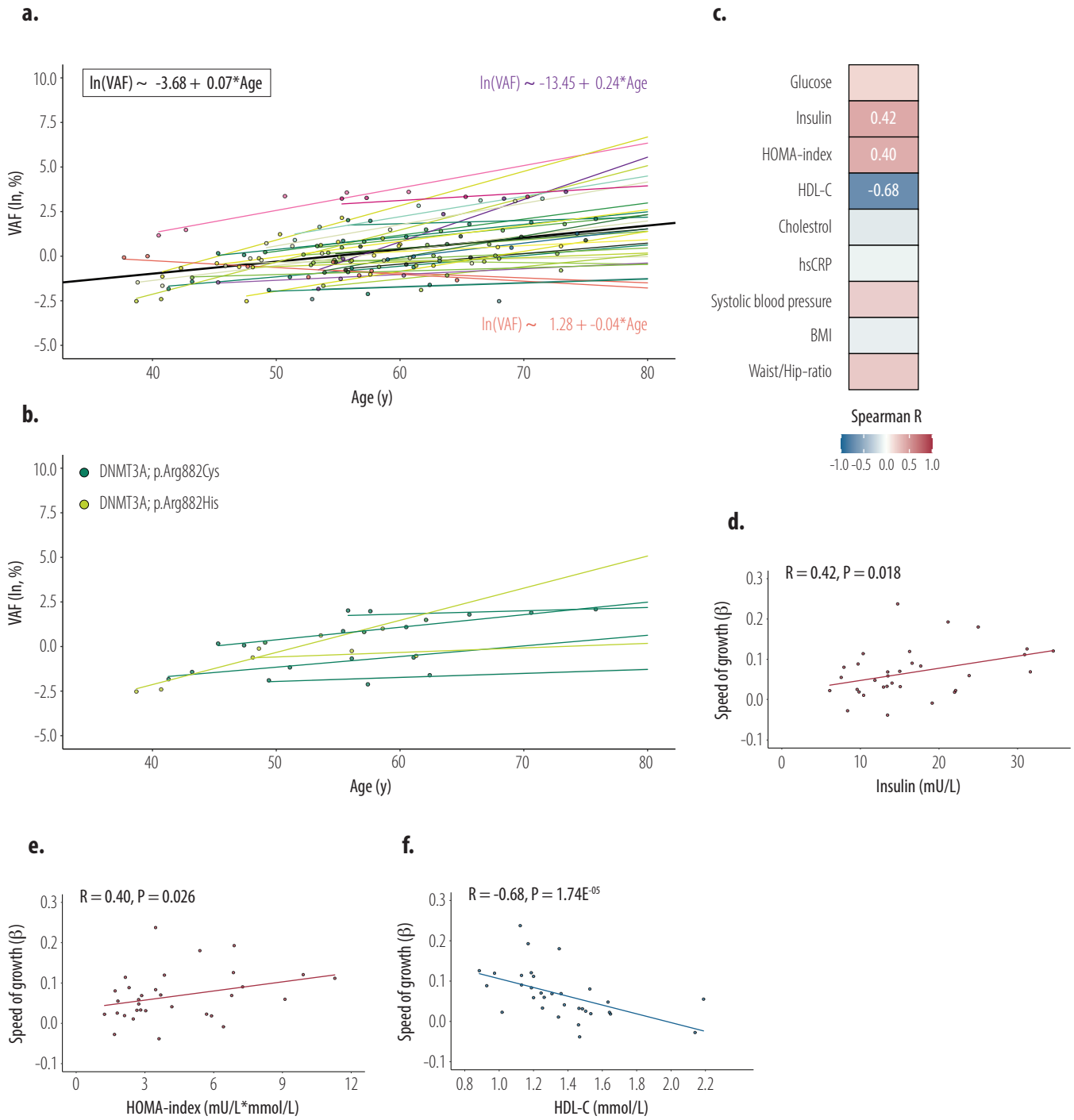
Multiple-timepoint data

N individuals = 40


a.
79.43%
0.29%
0.67%
3.24%
16.38%
no CHDM
2 CHDM
4 CHDM
1 CHDM
3 CHDM
b.
VAF (%)
max: 31.15%
mean: 2.72%
median: 0.70%
min: 0.01%
Small clone (<2%)
Large clone (≥2%)
c.
CHDM prevalence (%)
DNMT3A
GNAS
CBL
GNB1
TET2
ASXL1
JAK2
TP53
IDH2
KRAS
STAT3
SF3B1
Age < 45y [n=163]
Age 45-49y [n=230]
Age 50-54y [n=266]
Age 55-59y [n=252]
Age ≥ 60y [n=196]
d.
Age (y)
2.43E-06
51.89
54.47
non-carrier [n=834]
CHDM-carrier [n=216]
e.
CHDM prevalence (%)
Small clone (<2%)
Large clone (≥2%)
Age < 45y [n=163]
Age 45-49y [n=230]
Age 50-54y [n=266]
Age 55-59y [n=252]
Age ≥ 60y [n=196]
f.
Age (y)
2.36E-04
53.36
57.23
Small clone carrier [n=154]
Large clone carrier [n=62]
g.
Small clone (<2%)
Large clone (≥2%)
No CHDM
CHDM probability (%)
total clones OR age: 1.05 [1.03-1.08]
large clones OR age: 1.11 [1.07-1.16]
CHDM-carrier
non-carrier
Carrier status
Age (y)
h.
Small clone (<2%)
Large clone (≥2%)
VAF (ln, %)
ln(VAF) ~ -3.20 + 0.06*Age
Age (y)


a.
Mean
Median
VAF (%)
100
10
1.0
0.1
0.01
0
2
10
15
20
Follow-up time (y)
b.
Trajectory - growing
Trajectory - static
Trajectory - shrinking
Event
Relative VAF (%)
40
30
20
10
1
0
-1
-10
40
50
60
70
Age (y)
c.
115 CHDMs
38 Events
77 Trajectories
30 Growing
42 Static
5 Shrinking
VAF (%)
DNMT3A
CBL
TET2
JAK2
IDH2
GNB1
GNAS
BRCC3
ASXL1
TP53
NRAS
Age (y)
d.
Individual 12
Q415stop
S770L
R882H
C818R
R736H
P565stop
Individual 19
K57E
V567del
D702E
D768G
R565stop
Individual 14
E873stop
F732I
F732del
S770L
Color legend
DNMT3A
TET2
GNB1


a.
b.
c.
Color legend a,b,c
DNMT3A
TET2
IDH2
GNAS
ASXL1
KRAS
CBL
JAK2
STAT3
GNB1
TP53
SF3B1
DNMT3A: p.R882H
DNMT3A: p.R729W
DNMT3A: p.Y735C
DNMT3A: p.R882C
GNAS: p.R844H
GNB1: p.K57E
JAK2: p.V617F
DNMT3A: p.I780T
DNMT3A: p.S770L
CBL: p.C404Y
DNMT3A: p.R326C
CBL: p.R420Q
DNMT3A: p.R736G
Count (n)
Amino Acid Change
d.
R882(21)
R736(12)
Y735(10)
R729(12)
D768(10)
R326(6)
R635(5)
S770(5)
I780(5)
R771(6)
PWWP
ADD
SAM-dependent MTase C5-type
Missense (n)
LoF (n)
e.
Event
Static Trajectory
Growing Trajectory
Shrinking Trajectory
S770L(1)
I780T(1)
R882C(3)
R882H(2)
C818R(1)
Y735C(3)
R736G(1)
F732I(1)
R729W(1)
R326C(1)
G646R(1)
Y660C(1)
W314stop(1)
S349stop(1)
P385Rfs*22(1)
Q415stop(1)
p.Pro431Serfs*14(1)
W698stop(1)
R771stop(1)
f.
W882(495)
Y735(66)
R732(26)
Y736(68)
Y749(34)
R729(43)
Y755(29)
S780(27)
P904(74)
R326S(61)
G543(37)
R635(70)
W598(23)
R771(40)
W860(26)
W305(14)
W330(13)
W356(11)
W320(19)
W313(14)
W314(12)


a.
ln(VAF) ~ -3.68 + 0.07*Age
ln(VAF) ~ -13.45 + 0.24*Age
ln(VAF) ~ 1.28 + -0.04*Age
VAF (ln, %)
Age (y)
b.
DNMT3A; p.Arg882Cys
DNMT3A; p.Arg882His
VAF (ln, %)
Age (y)
c.
Glucose
Insulin
HOMA-index
HDL-C
Cholestrol
hsCRP
Systolic blood pressure
BMI
Waist/Hip-ratio
0.42
0.40
-0.68
Spearman R
-1.0 -0.5 0.0 0.5 1.0
d.
R = 0.42, P = 0.018
Speed of growth (β)
Insulin (mU/L)
e.
R = 0.40, P = 0.026
Speed of growth (β)
HOMA-index (mU/L*mmol/L)
f.
R = -0.68, P = 1.74E-05
Speed of growth (β)
HDL-C (mmol/L)