Evolution of clonal hematopoiesis during cancer treatment and its impact on outcomes

Mona Arabzadeh[1,2], Yi-Han Tang[3,4], Christelle Colin-Leitzinger[3], Sadegh Marzban[5], Daniel Walgenbach[6], Stefania Morganti[7,8], Vaidhyanathan Mahaganapathy[1,#], Erika Harper[9], Mingxiang Teng[4], Jacob K. Kresovich[3], Iman Washington[10], Heather A. Parsons[7,8], Judy E. Garber[7,8], Jeffrey West[5], Shridar Ganesan[1,11,*], Hossein Khiabanian[1,12,^,*], Nancy Gillis[3,13,*]

[1] Center for Systems and Computational Biology, Rutgers Cancer Institute, Rutgers University, New Brunswick, New Jersey

[2] Center for Biomedical Informatics, Bioinformatics Department, Rutgers Cancer Institute, Rutgers University, New Brunswick, New Jersey

[3] Department of Cancer Epidemiology, Moffitt Cancer Center and Research Institute, Tampa, Florida

[4] Department of Biostatistics and Bioinformatics, Moffitt Cancer Center and Research Institute, Tampa, Florida

[5] Department of Integrated Mathematical Oncology, Moffitt Cancer Center and Research Institute, Tampa, Florida

[6] Department of Internal Medicine, University of South Florida, Tampa, Florida

[7] Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts

[8] Harvard Medical School, Boston, Massachusetts

[9] Department of Pathology, Moffitt Cancer Center and Research Institute, Tampa, Florida

[10] Department of Radiation Oncology, Moffitt Cancer Center and Research Institute, Tampa, Florida

[11] Department of Medicine, Center for Molecular Oncology, NYU Perlmutter Cancer Center, New York, New York

[12] Department of Pathology and Laboratory Medicine, Rutgers Robert Wood Johnson Medical School, Rutgers University, New Brunswick, New Jersey

[13] Department of Malignant Hematology, Moffitt Cancer Center and Research Institute, Tampa, Florida

Current address: Ellison Medical Institute, Los Angeles, California

^ Current address: Regeneron Genetics Center, Tarrytown, New York

*** Corresponding authors**

Nancy Gillis, PharmD, PhD

Assistant Member, Department of Cancer Epidemiology

Moffitt Cancer Center and Research Institute, 12902 Magnolia Drive, Tampa, Florida 33612

Phone : +1- 813-745-6442 ; e-mail: nancy.gillis@moffitt.org

Hossein Khiabanian, PhD

Adjunct Professor, Department of Pathology, Robert Wood Johnson School of Medicine

Rutgers Cancer Institute, 195 Little Albany Street, New Brunswick, New Jersey 08901

Phone: +1-401-742-3113 ; e-mail: h.khiabanian@rutgers.edu

Shridar Gensan, MD, PhD

Professor, Department of Medicine, NYU Grossman School of Medicine

Perlmutter Cancer Center, 160 E 34th St, New York, New York 10016

Phone: +1-908-656-1427; e-mail: shridar.ganesan@nyulangone.org

Word count: 148 (Abstract); 9560 (All)

Number of Figures/Tables: 7 Figures, 1 Table; **Supplemental:** 10 Figures, 6 Tables

ABSTRACT

Clonal hematopoiesis (CH) is the age-related expansion of mutated hematopoietic stem cells without hematologic abnormalities. In patients with solid tumors, CH is associated with higher mortality and may evolve to therapy-related myeloid neoplasms; however, the mechanisms by which cancer treatments promote CH dynamics remain largely unknown. Here, we analyzed 392 serial samples from a prospective cohort of patients with breast cancer and showed that cytotoxic treatments led to strong therapeutic bottlenecks, resulting in significant reductions in hematopoietic allelic populations and differential clonal selection. Positively selected CH that expanded through dose-dependent therapeutic bottlenecks harbored mutations in *TP53*, *PPM1D*, *SRCAP*, *DNMT3A*, and *YLPM1*. Patients with positively selected CH during treatment had the shortest progression-free and overall survival compared to patients with unchanging or negatively selected CH across all therapies. These findings, validated in independent breast cancer and pan-cancer cohorts, provide strong evidence for clinical relevance of monitoring CH during cancer treatment.

INTRODUCTION

Clonal hematopoiesis (CH) is caused by the expansion of hematopoietic stem cells that carry somatic alterations in genes recurrently mutated in myeloid neoplasms. CH mutations with variant allele frequencies (VAFs) ≥2%, defined as clonal hematopoiesis of indeterminate potential (CHIP), are associated with increased overall mortality, cardiovascular disease, and progression to hematologic neoplasms(1-4). CH mutations are more common in individuals with solid tumors compared to healthy population-based cohorts(5) and are routinely detected in blood and tumor specimens(6-8).

The growth pattern and Darwinian evolution of mutation-driven CH resemble cancer and are shaped by gene-specific fitness effects, hematopoietic cell-specific rates of mutation, and imposed adaptive pressure on hematopoiesis(9-14). Cross-sectional studies have demonstrated a higher prevalence of CH mutations, particularly at CHIP-defining levels, following exposure to cancer therapy(15). Longitudinal studies in cancer-free populations and patients with solid tumors have further shown mutation-specific CH evolution and progression to myeloid malignancies(16-20). Analyses in patients with solid tumors reveal distinct mutational patterns and clinical associations under cancer treatment(21-25); however, the evolutionary mechanisms driving differential CH clonal dynamics and their relationship to clinical outcomes are largely unknown.

Extrinsic selection pressures imposed by cancer therapy may induce hematopoietic stress that promotes CH clonal expansion or leukemic transformation, or conversely, lead to depletion of mutated hematopoietic populations after treatment for solid tumors(26-28). Restricting therapeutic bottlenecks result in deep reductions in hematopoietic allelic populations, establishing an evolutionary setting in which fitter clones preferentially expand, while less fit or neutral clones may or may not persist due to highly stochastic random drift.

To investigate the impact of cancer treatment on mutation-driven CH, we analyzed 392 serial peripheral blood samples collected before, during, and after treatment from patients with breast cancer. We quantified treatment-specific selective pressures, clonal fitness, and effective hematopoietic population sizes, and evaluated associations between CH dynamics and clinical outcomes. Findings were validated in independent cohorts of patients with breast and pan-cancer tumors.

RESULTS

CH Mutational Spectrum at Breast Cancer Diagnosis

Mapping the mutational landscape of clonal hematopoiesis (CH) mutations in samples collected before treatment initiation (range, 1 day to 7.7 months) revealed 160 protein-changing CH mutations (VAF ≥0.01%) in 55% of patients (94/171) (**Figure 1A-B**). Patients with CH at diagnosis were older than those without CH (median age, 62 vs 52 years; *P*<0.001) (**Table 1**). The prevalence of CH increased with age, reaching 75% among patients aged 65 to 75 years (**Figure 1C**). The median number of CH mutations within an individual was 1 (range, 0–5) and was associated with age (*P*=0.04) (**Figure 1D**).

In *DNMT3A*, *TET2*, and *ASXL1*, the most frequently mutated CH genes in healthy populations(29, 30), we detected missense, loss-of-function nonsense, and frameshift mutations in 60% of patients with CH mutations (60/94) (**Figure 2A**, **Table S2**). To assess evidence of clonal selection(31), we compared observed mutation classes with expectations based on nucleotide composition. In 83% of genes with at least 5 nonsynonymous mutations (5 of 6), nonsense substitutions were enriched 7- to 18-fold relative to neutral expectations (FDR<0.001) (**Figure 2B**). Although the limited number of missense substitutions in the remaining genes did not provide power to assess their clonal selection (**Table S3**), these results corroborate the extent of population-level, positive selection of loss-of-function mutations in CH(9).

CH Mutational Diversity at Breast Cancer Diagnosis

The most frequent co-occurring CH mutations involved *DNMT3A*, which was found co-mutated with *TET2* (n=5; odds ratio (OR)=4.50, *P*=0.14), *YLPM1* (n=4; OR=0.83; *P*=0.73), and *ZNF318* (n=3; OR=1.20; *P*=0.57) (**Figure 2C**), although not reaching statistical significance. At least one CHIP mutation (strictly defined throughout the paper by VAF ≥2%) was detected before treatment in 33% of patients with CH (31/94), most commonly in *DNMT3A* (n=15) and *TET2* (n=3) (**Figure 2D**). Among patients with more than one CH mutation (n=41), clonal diversity measured by the coefficient of variation of VAFs had a median of 0.68 (range, 0.10–1.44) and was significantly higher in patients with at least one CHIP mutation compared with those without CHIP-level mutations (rank-sum *P*=0.001, **Figure 2E**), even after excluding mutations with VAF ≥2% from the analysis (rank-sum *P* = 0.02, **Figure S3A**). These results underscore the notion that as early CH clones grow in abundance, a larger number of CH mutations may accumulate and give rise to increased clonal diversity.

CH Mutational Rise and Fall During Breast Cancer Treatment

To evaluate clonal dynamics under treatment, we analyzed 230 serial peripheral blood samples (1–4 per patient) collected at a median of 5.5 months after treatment initiation (range, 0.01–132.3 months) (**Figure 3A**; **Figure S2**). All CH mutations detected after treatment initiation were present in pretreatment samples, enabling longitudinal tracking of VAF changes (**Figure 3B**).

Short-term Impact of Treatment on CH Allele Frequency. Within the first 18 months after treatment initiation, the maximum normalized monthly change in VAF was highest among patients receiving chemotherapy (0.13±0.04% per month), followed by combination chemotherapy and radiation (0.08±0.02% per month) and radiation alone (0.07±0.02% per month) (**Figure 3C**; **Table S4**). In contrast, patients receiving hormonal therapy alone showed no measurable change in CH VAFs during this period (-0.005±0.02% per month).

The fitness landscape of CH is gene and mutation dependent(9, 12, 14). Gene-specific analyses showed that among patients treated with chemotherapy alone, *TP53*-CH exhibited the greatest average standardized VAF increase (0.40±0.28% per month), followed by *YLPM1* mutations (0.19±0.12% per month) (**Figure 3D**; **Table S4**). Radiation therapy was associated with measurable VAF increases in *ATM*-mutated clones (0.18±0.13% per month). No gene-specific VAF changes were observed among patients receiving hormonal therapy alone, and *DNMT3A* mutations showed significantly smaller VAF changes compared with other treatment modalities (rank-sum $P<0.02$) (**Figure S3B**). All CHIP-level mutations detected before treatment remained at CHIP-level VAFs after treatment (**Figure S3C**).

Long-term Impact of Treatment on CH Allelic Abundance. To assess long-term effects, we analyzed 10 samples from 6 patients collected 22 to 133 months after treatment initiation. CH mutations showed continued changes in VAF over time, with the largest increases observed among patients treated with cytotoxic chemotherapy. In two patients, *DNMT3A* mutations increased by more than 0.1% VAF per year across 4 to 8 years of follow-up (**Table S4**).

CH Evolutionary Dynamics during Breast Cancer Treatment.

Effective Allelic Population Size During Treatment. To quantify selective pressures during treatment, we estimated the effective allelic population size (N_{eff}) using pre- and post-treatment VAFs, adapting the method for estimating pathogen transmission population size(32-34). We interpret N_{eff} as an estimate that captures the magnitude of stochastic drift in mutated hematopoietic populations during treatment, rather than as the absolute number of hematopoietic

stem or progenitor cells. As such, N_{eff} reflects the relative size of the therapeutic bottleneck acting on the evolving hematopoietic compartment, where smaller values indicate stronger drift and greater clonal fluctuation.

In our cohort during the first 18 months of treatment, median N_{eff} was the highest among patients receiving hormonal therapy (median, 3,804 alleles; range, 397–53,206) and radiation (median, 2,765; range, 62–1,000,000). In contrast, N_{eff} was significantly lower among patients receiving chemotherapy (median, 496; range, 16–135 888) similar to combination chemotherapy and radiation (median, 277; range, 42–798,800), pointing to substantially more restrictive therapeutic bottlenecks imposed by cytotoxic treatment of breast cancer (**Figure 3E**).

CH Mutational Fitness During Treatment.

Using patient-specific effective allelic population size (N_{eff}) estimates and longitudinal VAF measurements, we quantified evolutionary fitness of CH mutations relative to the wild-type alleles during treatment. Specifically, we evaluated whether the observed change in VAF after treatment could be explained by neutral genetic drift alone. For each mutation, we calculated the probability of observing the post-treatment VAF given its pre-treatment VAF assuming a neutral sampling process from an allelic population size of N_{eff} and classified the CH mutations as positive or negative selection only when the observed VAF change was unlikely to arise from drift alone. Mutations showing significant expansion beyond drift expectations, with fitness likelihoods exceeding 10,000-fold (>5 orders of magnitude) relative to wild-type alleles, were classified as positively/negatively selected, whereas those showing substantial contraction were classified under the no selection group. The direction of selection was then determined by the sign of the VAF change.

In our cohort, within the first 18 months from treatment start, the highest mutational fitness was observed under chemotherapy. Among patients treated with chemotherapy or combination therapy, 27% (22/81) had positively selected mutated clones (**Table S5**). Under radiation, *DNMT3A* mutations showed the highest relative fitness (**Figure 3F**). The fittest *PPM1D* mutant clones had truncating mutations in exon 6, linked to a chemoresistance phenotype(35), and positively selected *DNMT3A* mutations were in the PWWP/chromatin-binding (exons 8-9) and DNA methylase (exons 16-19) domains (**Figure S4**).

CH Mutational Dynamics Across Treatment Modalities and Dosage Levels. CH has been shown to differentially respond to cancer treatment based on the mechanism of action and dosage. Among patients with pretreatment CH, 38% (36/94) exhibited dynamic VAF changes

within the first 18 months, including 9% (3/35) receiving hormonal therapy, 21% (7/55) radiation, 30% (21/69) chemotherapy, and 42% (5/12) combination therapy. In 3% of patients (3/94), all treated with radiation, both increases and decreases in CH VAFs were observed.

Effective allelic population sizes were smaller among patients with positively or negatively selected CH compared with those with stable CH, with the most pronounced contractions observed among patients receiving chemotherapy or combination therapy, suggesting that therapeutic elimination of wild-type cells may be a major driving factor behind the observed evolutionary fitness of selected CH mutations (**Figures 4, 5A; Figure S5**).

Patients with and without CH had similar levels of chemotherapy exposure. Among patients with CH and treated with chemotherapy (n=36), increasing cumulative cytotoxic exposure was associated with higher frequencies of CH dynamics (chi-square P=0.003). In addition, patients with high or medium cumulative exposure had significantly smaller N_{eff} during cytotoxic therapy relative to patients with low-dosage treatment; rank-sum P=0.025 and P=0.013, respectively (**Figure 5B**). No significant association was observed between radiation dosage and N_{eff}. In fact, it is the dose-dependent severity of bottlenecks imposed by systemic cytotoxic therapy that yields shrinking hematopoietic populations which, together with mutation-specific resistance of CH clones to chemotherapeutic drugs(36), may result in increased CH VAFs during treatment.

CH Mutational Dynamics Validation Across Independent Cancer Cohorts

Analysis of independent cohorts from Dana-Farber Cancer Institute (early stage breast cancer, n=62)(22) and Memorial Sloan Kettering Cancer Center (solid tumors, n=394)(5) confirmed smaller effective allelic populations and stronger selective pressures among patients treated with cytotoxic therapies compared with those receiving hormonal therapy or no cytotoxic treatment (**Figure 5C-D, Figure S6-7**). Moreover, there was a consistent relationship between cytotoxic exposure levels and CH dynamics across all cancers (MKSCC cohort), where positive or negative selection of mutations was significantly associated with higher chemotherapy (chi-squared P=0.001) or radiation dosage (chi-squared P=1e-04) (**Figure S7**). N_{eff} sizes were significantly smaller in patients who received the highest cumulative dosage of chemotherapy (rank-sum P=1e-05, **Figure 5E**) or were exposed to high or medium levels of radiation compared to no cytotoxic treatment (rank-sum P=2e-04 and P=8e-04, respectively, **Figure 5F**). These results confirm the relationship between cytotoxic treatment dosage and quantified reduction in hematopoietic populations(37), pointing to CH clonal dynamics as a biomarker for outcome across cancers.

Association of CH Dynamics With Clinical Outcomes

Overall survival (OS) and progression-free survival (PFS) did not differ between patients with and without CH before treatment (**Figure 6A**). However, among patients with CH, those with positively selected mutations during treatment had significantly worse OS and PFS (**Figures 6B, 7A**). In multivariable Cox regression models adjusted for age, stage, and ER, PR, and HER2 status, positive CH selection was associated with shorter OS (adjusted hazard ratio [HR], 6.6; 95% CI, 1.7–25.9; *P*=0.007) and worse PFS (adjusted HR, 2.86; 95% CI, 1.01–8.1; *P*=0.049) (**Figure 7B-C; Table S6**). These results were further supported by landmark analyses at 12 and 18 months as well as a time-dependent Cox model incorporating patient-specific sample collection time points, all which yielded results consistent with multivariable and univariable Cox models (**Table S6**).

These associations persisted in analyses restricted to patients receiving chemotherapy and in recurrence-free survival analyses (**Figure S8**). When considering all patients who received chemotherapy, there was no association between cumulative chemotherapy dose and OS; however, multivariable analyses assessing the relationship between chemotherapy exposure level and CH dynamics were limited by small samples within the subgroups (**Table S6**). Among patients who received similar treatment regimens, positively selected CH patients tended to have worse outcomes relative to negatively selected or unchanged CH, although statistical significance was not reached for these subgroup analyses (**Figure S9**). Together, these results link CH clonal dynamics during treatment to patient mortality and disease outcome, and for the first time, provide quantified evidence for the observed association between mutation-driven CH and OS(31).

DISCUSSION

In this longitudinal analysis of patients with breast cancer, we demonstrate that CH is present in 55% of patients, evolves dynamically during treatment, and is strongly shaped by therapy-specific selective pressures. By integrating high-sensitivity sequencing (detect >1 mutant in 1,000 wild-type alleles(10)) with evolutionary modeling, we show that cytotoxic therapies impose pronounced therapeutic bottlenecks on hematopoietic populations, leading to clonal expansion of fitter CH mutations. Importantly, we identify treatment-associated CH dynamics, rather than CH presence alone, as a predictor of adverse clinical outcomes. All post-treatment CH mutations were detectable at baseline, supporting the notion that cancer therapy selectively modulated pre-existing CH rather than initiating development of *de novo* CH(26).

The fitness landscape of CH is shaped by multifactorial processes such as aging, inflammation, germline predisposition(38, 39), and DNA-damaging exposures(31, 40). Mutations in the DNA damage response genes, including *TP53* and *PPM1D*, have been implicated in therapy-related myeloid neoplasms(5, 15, 40, 41). *SRCAP* and *YLPM1* have also been recognized as drivers of CH and myelodysplastic neoplasms(9, 42), with *SRCAP* linked to selective advantage under cytotoxic stress(43). Here, *YLPM1* emerged as a CH driver that persisted and expanded under therapy, suggesting a potential role in hematopoietic stress. Among common CH drivers, *DNMT3A* mutations selectively expanded under both chemotherapy and radiation. *DNMT3A* has been linked to chemotherapy resistance in hematologic malignancies(44) and associated with comorbidities in solid tumors(45, 46). *DNMT3A* mutations frequently co-occurred with *TET2*, *YLPM1*, and *ZNF318*, suggesting potential clonal cooperation or co-segregation; however, *TET2* mutations in blood did not show evidence of dynamic selection under therapy. Given that *TET2*-mutant CH are enriched in the tumor microenvironment and associated with tumor progression(7, 8), a possible explanation may be that *TET2* acts as a context-dependent co-driver by remodeling the tumor immune microenvironment, rather than via a direct proliferative advantage. This indirect role may support the fitness of *TET2*-mutant clones even without overt selection under therapy.

To assess whether the observed differences reflect biological selection between treatment groups, we examined VAF changes separately within chemotherapy- and radiation-treated cohorts for *DNMT3A*, *YLPM1*, *TP53*, and *ATM* (**Figure S3**). The observed patterns are consistent with differential selective pressures (chemotherapy favoring *DNMT3A*, *TP53*, and *YLPM1*, and radiation favoring *ATM*), but should be considered hypothesis-generating rather than definitive given the limited statistical power. We observed evidence suggestive of positive selection of *TP53* mutations in chemotherapy-treated patients, consistent with prior reports linking *TP53*-mutant hematopoietic clones to resistance to DNA damage–induced apoptosis(36). In contrast, no clear signal of *TP53* selection was observed in the radiation-treated group. In addition, the enrichment of *ATM* mutations in the radiation-treated group is biologically plausible, given the central role of *ATM* in the DNA damage response to double-strand breaks and radiation-induced genotoxic stress(47), supporting the possibility of treatment-specific selective pressure.

We independently confirmed the associations between CH dynamics, reduction in effective allelic population size, and cumulative cytotoxic exposure in other cohorts, denoting that genotoxic stress promotes relative expansion of clones with higher intrinsic fitness. This enrichment aligns with known links between CH and therapy-related neoplasms and may reflect therapy-specific sensitivity or immune-mediated clearance of less fit clones. The reduced impact of radiation on

281 N_{eff} and CH selection may reflect the limited hematopoietic impact of breast-directed radiation and
282 may not generalize to tumor types where radiation treatment more directly impacts large regions
283 of the hematopoietic stem cell compartment(48) (*e.g.*, pelvic radiation for gynecologic
284 cancers(49)). Consistent with this, pan-cancer analysis supports the likelihood for dose-
285 dependent radiation-induced CH positive selection in other cancer types.

286 CH evolution and growth have been linked with cancer therapy-related adverse sequelae(5, 15,
287 18, 20). Although none of the patients in our cohort developed therapy-related hematological
288 malignancies during the study period, stratification by CH clonal trajectories showed that patients
289 with positively selected CH had increased risk of mortality compared to those with neutral or
290 negatively selected CH, suggesting that the direction of clonal evolution rather than CH presence
291 is prognostically meaningful. This effect persisted when the analysis was limited to only
292 chemotherapy patients and in groups with homogenous treatment.

293 The mechanism linking positive selection of CH clones to worse outcomes remains unclear. In
294 our analyses, CH dynamics were primarily detected in genes with mutations linked to therapeutic
295 resistance and, therefore, their association with poor outcomes might be confounded by the extent
296 of treatment exposure due to progressive disease. Although no overall association was observed
297 between chemotherapy dose and outcomes across our whole cohort, exploratory analyses
298 suggested a differential relationship based on CH selection status. In patients with negative
299 selection of CH, chemotherapy dose had no apparent impact on outcomes, however, in positive
300 selection, higher cumulative doses may be associated with improved outcomes. These
301 observations further suggest that CH selection status may serve as an independent predictor of
302 treatment response and that the positive selection of CH is influenced by factors beyond
303 chemotherapy dose alone. In support of the direct role played by tumor-infiltrating CH clones on
304 tumor growth and treatment response,[8,(50),(51)] the association between positive selection for CH
305 clones in peripheral blood and outcomes also raises the possibility that circulating CH clones have
306 some direct or indirect effects on treatment response.

307 Our study provides a distinct advance over previous large retrospective CH analyses by shifting
308 from a static, cross-sectional framework to a prospective, longitudinal evaluation of clonal
309 hematopoiesis dynamics during cancer treatment. Although our cohort is smaller than earlier
310 retrospective studies(5),(15), the design enables direct measurement of how CH clones evolve
311 under toxic therapy, supported by a hormone-therapy control group, showing that observed
312 changes reflect both natural time-dependent drift and treatment-specific selective pressures.

High-depth sequencing further enhances this resolution by detecting subtle shifts in clonal architecture that would be missed at lower coverage.

Despite the modest sample size, our analyses within more homogeneous treatment groups revealed consistent patterns across CH-dynamic-stratified groups, reinforcing the robustness of these observations. By tracking clonal trajectories rather than cataloguing the presence of CH, we demonstrate that therapy-driven clonal expansion is associated with earlier tumor progression –an association not captured by prior work focused predominantly on enrichment patterns and leukemia risk. We further show that therapy-dependent reductions in hematopoietic allelic population size shape clonal fitness across treatment modalities. Notably, CH selection may not necessarily represent a causal driver of adverse outcomes but may instead reflect underlying physiological vulnerability or conditions promoting clonal expansion that influence infection risk, treatment tolerance, and survival.

Our results suggest that monitoring dynamic changes in CH during therapy, rather than baseline CH alone, could help identify patients at higher risk of adverse outcomes. Serial sampling during treatment may allow detection of expanding clones, particularly in genes such as *TP53* and *YLPM1* that show evidence of positive selection. Tracking this expansion during therapy may provide an early signal of emerging risk. Patients with CH mutations may therefore benefit from careful monitoring during cytotoxic therapies, particularly chemotherapy, as measurable clonal expansion can occur within the ~18-month timeframe captured by our analysis. These findings suggest a potential role for longitudinal CH monitoring as a complementary biomarker for guiding patient stratification during treatment. As tumor-informed minimal/measurable residual disease assays are now entering clinical practice, the presence of increasing CH could also be explored in prospective studies as a trigger to initiate or increase frequency of plasma minimal/measurable residual disease testing to detect subclinical tumor recurrence.

Beyond cell-intrinsic genetic advantages, the presence and expansion of CH clones may also reflect interactions with the broader hematopoietic and inflammatory microenvironment. Emerging evidence indicates that CH-associated mutations can alter inflammatory signaling and reshape the bone marrow niche, generating feedback loops that further promote clonal expansion. Thus, therapy-related stressors such as chemotherapy or radiation may interact not only with the genetic properties of CH clones, but also with the surrounding inflammatory and hematopoietic microenvironment to shape clonal dynamics(8, 50, 52). Expanding CH clones seen in peripheral blood may also reflect increasing presence of CH cells in micro-metastatic tumor

microenvironment. As tumor-associated CH has been linked with worse outcomes in lung cancer(8), it raises the hypothesis that increasing peripheral CH may reflect enrichment of CH cells in the tumor microenvironment surrounding dormant cancer cells, which then may contribute to future outgrowth/activation and metastatic progression. Taken together, these observations suggest that CH may function both as a biomarker of hematopoietic stress and as a biologically active process influencing systemic inflammation and tissue environments. Longitudinal monitoring of CH during therapy could therefore provide clinically relevant insight into treatment-associated selective pressures and patient-specific risk. Work is ongoing to define the optimal timing, frequency, and actionable thresholds for such monitoring.

In conclusion, these findings demonstrate that both treatment type and dynamic changes in CH during therapy have important clinical implications. The evolutionary responses of blood clones to cancer treatment have direct translational relevance for predicting patient outcomes. While CH presence alone does not universally confer adverse risk, longitudinal changes in CH allelic burden during treatment may signal increased susceptibility to therapy-related complications across solid tumors. Further validation in larger, clinically homogeneous cohorts is required to substantiate these observations.

METHODS

Sex as a biological variable. This study exclusively included breast tumor samples from female patients (n = 171), consistent with the predominance of breast cancer in females. As such, sex was not evaluated as a biological variable, and the generalizability of these findings to male breast cancer remains unknown.

Patient population, inclusion criteria, and clinical parameters. Clinical data and samples used for this study were collected between January 1, 1994, and July 31, 2021. Patients were eligible for inclusion if they had a primary breast cancer diagnosis and serial peripheral blood samples available in Moffitt Cancer Center's institutional biorepository that could be used for DNA extraction and CH detection, including one sample collected before the start of any cancer treatment and one sample collected after the first therapy given for breast cancer. Specifically, the pre-treatment sample needed to be collected within one year prior to treatment start or a maximum of five days after treatment start for patients treated with chemotherapy or hormone therapy. Patients treated with radiation needed to have a sample collected within a year before radiation start date. For patients treated with chemotherapy, the first sequential sample needed to be collected after the chemotherapy stop date, within a year from the start date, and before any

radiation treatment. For patients treated with radiation, the first sequential sample needed to be at least 100 days after radiation, within a year from radiation start, and before any chemotherapy treatment. For patients treated with hormone therapy, the first sequential sample had to be collected within a year, but at least 85 days (the average duration of chemotherapy), after hormone therapy start date and before any exposure to chemotherapy or radiation. For patients meeting these inclusion criteria, all other serial sampling timepoints available in the institutional biorepository, including during treatment and after the post treatment sampling times, were included.

Patients' cumulative exposure to chemotherapy drugs and radiation during the observation period was calculated following the approach by Bolton *et al*.[15], derived from the Late Effects Study Group1(53). For each patient, cumulative dose per kilogram for each drug received during the observation period was calculated, converted into tertile-based scores, and summed within drug classes. Radiation dose tertiles were calculated using the cumulative radiation dose during the observation period in 2-Gy per fraction (EQD_2), using an α/β of 4 Gy(54). These class-level scores were then divided into tertiles to categorize overall exposure across the cohort. Of note, cumulative radiation exposure was largely similar across the population (EQD_2 median, IQR).

DNA extraction and hybrid-capture, error-corrected sequencing. DNA was extracted using the Autopure LS automated DNA extractor (QIAGEN, Hilden, Germany) and quantified using Qubit (Invitrogen, Waltham, MA). Molecular data generation was done in collaboration with Moffitt Molecular Genomics Core. DNA libraries were constructed using a custom designed SureSelect^{XTHS2} kit (Agilent, Santa Clara, CA), which optimally captured whole exons of 81 hematopoietic disorder-related genes, including single nucleotide mutations and small indels. Libraries were sequenced on a NextSeq 2000 sequencer (Illumina, San Diego, CA) per manufacture's recommendations with a goal of achieving an average depth of coverage >5,000x.

Bioinformatics analysis and somatic variant calling. Sequencing reads were aligned to a corrected human genome (GRCh38) reference(55) using the BWA-MEM algorithm(56). Consensus variant calling from reads with the same unique molecular identifiers was done using fgbio v2.2.1(57). Somatic variant calling was performed using Genome Analysis Toolkit best practices and Mutect2(58). Additional statistical support for detection of variants in individual samples was obtained using a beta-binomial model for allele-specific sequencing noise (FDR<1e-4)(59). Variants were filtered and annotated using BCFtools(60). Insertions and deletions were called by two indel callers, Mutect2 and Lofreq(61), and were only retained if the VAF was ≥0.02.

Variants that occurred in greater than a tenth of the samples were considered sequencing artifacts and were removed. Quality control filters included strand odds ratio <3, quality score >6.3, observation of each variant at least once on the forward and reverse reads, and masking of the repetitive regions of the genome as defined by the DUST algorithm(62) plus a region with low-complexity in *KMT2D* (c. 1659–2544) linked to unreliable calling. All analyses in this study were performed using the full set of detected CH mutations (VAF >0.01%); the term CHIP (VAF ≥2%) is used only in analyses specifically restricted to clinically defined CHIP-level mutations.

Germline variants were removed using publicly available reference populations (*i.e.*, variants observed in non-cancer populations at a prevalence >0.005 and with a VAF >0.25(63, 64)). The high sequencing depth in this study yielded very small confidence intervals of less than 0.5% for measuring VAF; therefore, germline variants that are expected to be present at 50% allele frequency were robustly identified(65); however, to ensure removal of private germline variants, we excluded variants with VAF >40% detected in each patient's serial samples. To filter for likely functional somatic variants, only mutations or indels located in exonic regions were considered CH mutations. Germline single nucleotide polymorphism loci heatmaps were used to confirm correct sample pairing within patients. Chart review confirmed non-tumor cell origin of *TP53* CH mutations and distinction from *TP53* mutations detected by prior sequencing of tumor specimens.

External validation cohorts. The mutational and clinical data for the DFCI and MSKCC cohorts were obtained through Data Transfer Agreements between the institutions or from their respective publications(5, 22) and were reviewed to match the patient selection and variant calling criteria in the discovery cohort. In particular, downstream analyses was limited to cases with CH VAFs measured before and after cancer treatment and those with exonic variants curated by Vlasschaert *et al.*(66).

Clonal selection of CH driver genes. Genes with driver CH mutations were expected to show a higher number of loss-of-function nonsense or hotspot missense mutations compared to expected distribution under neutral selection. Chi-squared statistics were used to compare the observed to expected mutations counts, and lower and upper 95% binomial-approximated confidence intervals were calculated to assess uncertainty for excess of nonsense, missense, and synonymous mutations.

Calculating standardized CH VAF change per time. Variant allele frequencies (VAFs) were quantified across serial samples, and z statistics were used to measure temporal mutation- and gene-specific VAF change. Changes in CH allele frequencies for mutation *j* in patient *i* was

normalized and defined by $s_j^i = z_j^i/(t_2 - t_1)$, where t_1 and t_2 were collection time points, and z_j^i was two-proportion z-test. Across all analyses, the mean and standard error in the distributions of VAF changes were used to calculate a standardized VAF change per unit time and its confidence intervals for specific treatments and genes (per treatment modality).

In more detail, a two-proportion z-test was used to compare CH VAFs between time points, accounting for differences in sequencing depths. The z statistics is calculated based on the VAFs (f_1 and f_2) and total sequencing depths (D_1 and D_2) measured in serial samples, following $z_j^i = \frac{f_2 - f_1}{\sqrt{f\,(1-f)\,(\frac{1}{D_1}+\frac{1}{D_2})}}$ with $f = \frac{f_1\,D_1 + f_2\,D_2}{D_1 + D_2}$. The pooled allele frequency f is bounded within [0,1], and represents a weighted average of two proportions. In this context, the null hypothesis is that standardized CH VAFs do not change during treatment, and as such, $H_0: f_1 = f_2$. Since sequencing depths differ in each VAF measurement, this z-statistic reflects the changes in VAFs relative to sampling noise beyond just the raw measurement. For example, a small VAF change at very high depth leads to a large z; and a large VAF change at low depth results in a small z, driven by statistical certainty and not the VAF change magnitude alone.

The 95% confidence intervals for standardized VAF change estimates were calculated using confidence intervals for z_j^i at α=1.96. The standard deviation of the z-statistics is given by $\delta = \sqrt{f(1-f)\left(\frac{1}{D_1}+\frac{1}{D_2}\right)}$, which for fixed allele frequencies f_1 and f_2, converges to zero as sequencing depths $D_1$and D_2 increase and the term $\left(\frac{1}{D_1}+\frac{1}{D_2}\right)$ converges to zero. Finally, since $z = \frac{f_2 - f_1}{\delta}$, any non-zero allele frequency difference yields increasingly large z values with increasing depths. Thus, this metric reflects statistical confidence in VAF change and is depth-aware by definition.

Therefore, the resulting z-statistic, normalized by elapsed time, was used to define a metric to calculate statistical confidence for changes in CH VAFs while considering sampling variance. As detailed above, this z-statistic scales with sequencing depth, and therefore, $\frac{z}{\Delta t}$ represents a significance-weighted, time-normalized statistical measure of change in VAF that incorporates uncertainty arising from finite sampling. It should be noted that this metric cannot be interpreted as a direct estimate of biological growth rate or selective advantage.

Therapeutic bottleneck and evolutionary fitness likelihood calculation. A binomial sampling model was applied to estimate effective hematopoietic population size and CH fitness likelihood during treatment. To estimate the effective allelic population size, it was assumed that CH

mutations were clonally independent and VAFs were representative of single CH clones. If n_i mutated alleles harbored a specific variant at time point 1, the probability of observing m_i mutated alleles with the same variant at time point 2 would be described by binomial sampling, $p(m_i|N,n_i) = \binom{N}{m_i}\left(\frac{n_i}{N}\right)^{m_i}\left(1-\frac{n_i}{N}\right)^{N-m_i}$, with bottleneck size, N. For s number of CH variants shared between any two time points, the maximum likelihood estimate for N, describing the lower bound on effective bottleneck population size was equal to $\widehat{N} = \frac{s-1}{2\,\mathrm{KL}(f_2|f_1)} < N_{\mathrm{eff}}$, with KL representing the Kullback-Leibler divergence, and f_1 and f_2 as measured VAFs at time points 1 and 2, respectively. The maximum-likelihood-estimated variance of N_{eff} was then equal to $\sigma_N = s/N_{\mathrm{eff}}^2$.

Using patient-specific effective allelic population size (N_{eff}) estimates, the probability of observing a post-treatment VAF (f_2) given an initial pre-treatment VAF (f_1) under neutral drift was modeled using a binomial sampling process: $P(f_2 \mid f_1, N_{eff}) = \binom{N_{eff}}{f_2 N_{eff}}(f_1)^{f_2 N_{eff}}(1-f_1)^{(1-f_2)N_{eff}}$. This represents the likelihood that the observed allelic frequency change can be explained by stochastic genetic drift alone following a treatment-induced population bottleneck. Evolutionary fitness likelihood for each mutation was then defined as the negative base-10 logarithm of this probability: $L = -\log_{10}\big(P(f_2 \mid f_1, N_{eff})\big)$. Operational classification of selection was defined as follows: Neutral: $L \leq 5$ (VAF change consistent with drift); Positive Selection: $L > 5$ and $f_2 > f_1$; Negative Selection: $L > 5$ and $f_2 < f_1$. Thus, selection reflects statistically unlikely deviation from neutral drift expectations rather than raw VAF change alone.

Depth sensitivity analysis. The observed VAF is derived from finite read depth and thus subject to sampling noise, the high sequencing depth in our dataset (mean >5,000× per mutation) ensures that the binomial sampling variance is small relative to the observed differences in allele frequency. Therefore, treating the observed VAF as an exact probability provides a close approximation to the true underlying frequency. Any residual sampling noise would tend to bias effective population size (N_{eff}) estimates downward, making our estimates conservative. Furthermore, since read depths are comparable across mutations and patients, the relative differences in N_{eff} between clones and individuals remain valid, supporting the conclusions drawn from our analysis. To quantify these expectations, we calculated the confidence interval of N_{eff} based on the propagation of uncertainty in measuring VAFs as a function of sequencing depth.

Sampling Variance for CH Variant Allele Frequencies. Observed VAFs are affected by sequencing sampling noise. For a sequencing depth of D, the sampling variance of an allele frequencies p_i

and q_i are approximated by a binomial model: $\sigma_{p_i}^2 = \frac{p_i(1-p_i)}{D}, \sigma_{q_i}^2 = \frac{q_i(1-q_i)}{D}$. Since the frequency of the wild-type clone is equal to one minus the sum of CH VAFs, the variance in wild-type frequency is equal to the sum of variances in CH VAFs. Thus, the uncertainty in N_{eff} depends on sequencing depth through the uncertainty in measured allele frequencies.

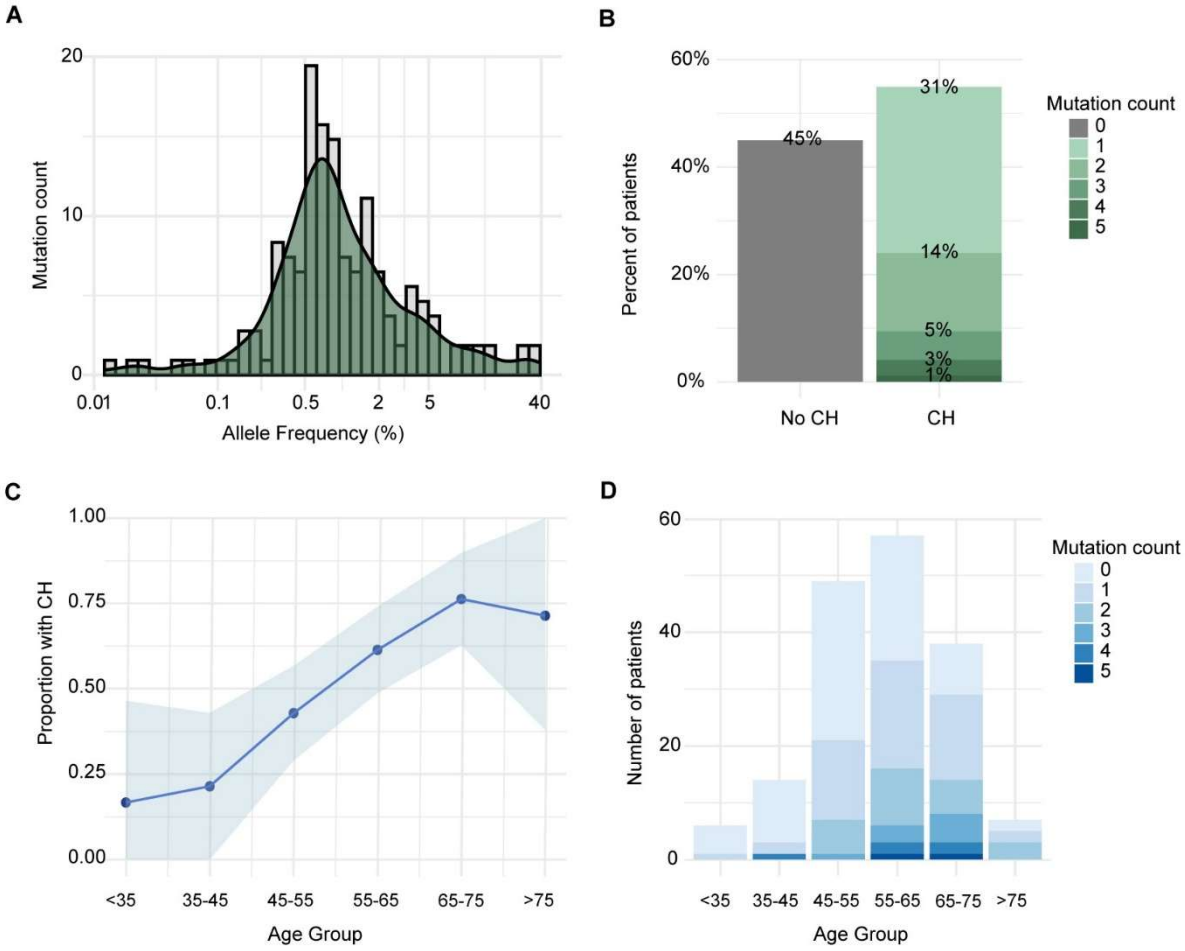
Propagation of Uncertainty for N_{eff}. Effective population size N_{eff} can be estimated with $N_{\text{eff}} = \frac{s-1}{2\,KL(p|q)}$ where $KL(p|q) = \sum_{i=1}^{S} p_i \log\left(\frac{p_i}{q_i}\right)$, where s is as the number of CH and wild-type clones, and q and p are their pre- and post-treatment allele frequencies, respectively. To estimate the uncertainty of N_{eff}, a first-order error propagation was applied where the variance of N_{eff} is approximately equal to $\sigma_{N_{\text{eff}}}^2 = \sum_{i=1}^{S}\left(\frac{\partial N_{\text{eff}}}{\partial p_i}\right)^2 \sigma_{p_i}^2 + \sum_{i=1}^{S}\left(\frac{\partial N_{\text{eff}}}{\partial q_i}\right)^2 \sigma_{q_i}^2$. The derivatives follow from the dependence of N_{eff} on the KL divergence: $N_{\text{eff}} = \frac{S-1}{2KL}$, so that $\frac{\partial N_{\text{eff}}}{\partial x} = -\frac{S-1}{2KL^2}\frac{\partial KL}{\partial x}$ for $x \in \{p_i, q_i\}$. The derivatives of the KL divergence are $\frac{\partial K}{\partial p_i} = \log\left(\frac{p_i}{q_i}\right) + 1$, $\frac{\partial KL}{\partial q_i} = -\frac{p_i}{q_i}$. Substituting these derivatives into the propagation formula yields an analytical approximation of the variance of N_{eff} as a function of sequencing depth. Confidence Intervals (CI) for N_{eff} can then be approximated using $\sigma_{N_{\text{eff}}}$ where 95% CI equal to $N_{\text{eff}} \pm 1.96\,\sigma_{N_{\text{eff}}}$.

To evaluate the effect of sequencing depth on the precision of the N_{eff} estimator, the above propagation was evaluated across a range of sequencing depths 200x to 6,000x. For each patient in our cohort, we down-sampled the sequencing depths and for each depth, calculated the sampling variance for measured CH VAFs according to the model above, yielding depth-dependent confidence intervals for N_{eff}. As expected, the coefficient of variance calculated across the cohort for all patients showed reduction of error in estimating N_{eff} at higher depths (**Figure S10**).

Clonal independence sensitivity analysis. In estimating N_{eff}, it was assumed that each CH mutation marked an independent clone. While this assumption is supported by the low biological plausibility of sequential CH-driving mutations arising within the same cell as discuss in Marzban et al. (2026)(67), nested or co-occurring clones may occur as predicted by Watson et al. (2024)(10). To assess the impact of nested clones on N_{eff} estimation, all analyses were repeated assuming that all detected CH variants were nested in the largest clone, representing the most restrictive scenario. This analysis showed that the differences in N_{eff} across treatment groups, as well as CH mutational dynamics, remained consistent with our original analyses (**Figure S10**).

Statistical and clinical association analyses. Statistical significance between groups was assessed using exact tests, chi-squared tests, log-rank tests, or nonparametric rank-sum tests, as indicated. The Benjamini-Hochberg false discovery rate (FDR) method was applied to adjust for multiple hypotheses testing. Clinical associations were evaluated using multivariable Cox proportional hazards regression models assessing overall survival and progression-free survival. Models included age, cancer stage, and estrogen receptor (ER), progesterone receptor (PR), and HER2 status as covariates. Because CH selection status was defined using post-baseline measurements (with baseline defined as the treatment start date), potential concerns about immortal time bias were addressed by landmark analyses at 12 and 18 months, as well as a time-dependent Cox model incorporating patient-specific sample collection time points. Given the limited number of covariates relative to the number of events, the models fell within commonly accepted ranges for events per variable (≥10 events per variable), reducing the likelihood of substantial overfitting. Statistical analysis and data visualizations were performed using following packages in R: *surminer*, *survival* (survival analyses), and *ggplot2* (plotting). *Maftools(68)* was used for generating oncoplot and lollipop figures, and *EvoFreq(69)* was used to visualize the evolutionary models.

Study approval. Patients for this study were consented to the Moffitt Cancer Center's Total Cancer Care Protocol, an Institutional Review Board-approved institutional biorepository (MCC#14690; Advarra institutional review board [IRB] Pro00014441)(70). Use of biobanked patient samples for genetic data generation for this study was approved under a release protocol (MCC#21545, Advarra IRB Pro00058968).



553

554 **FIGURE 1. Prevalence and mutational distribution of CH prior to treatment for breast cancer.**
555 A) Density distribution of CH mutation variant allele frequencies (VAFs). B) Percent of patients with or
556 without CH, stratified based on the number of detected mutations per individual. C) Proportion of patients
557 with CH across age groups. D) Number of patients with CH across age groups, stratified based on the
558 number of detected mutations per individual.

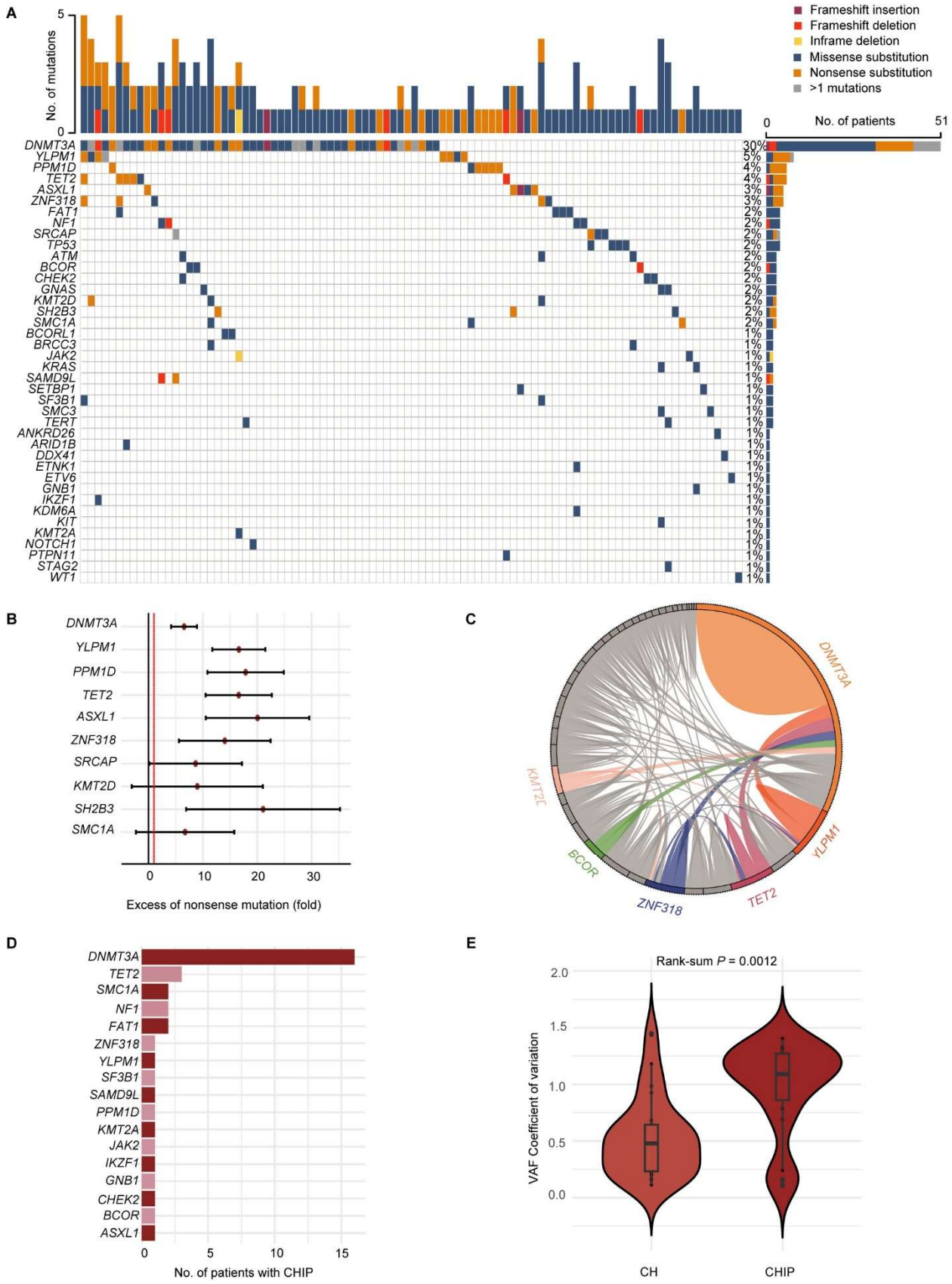
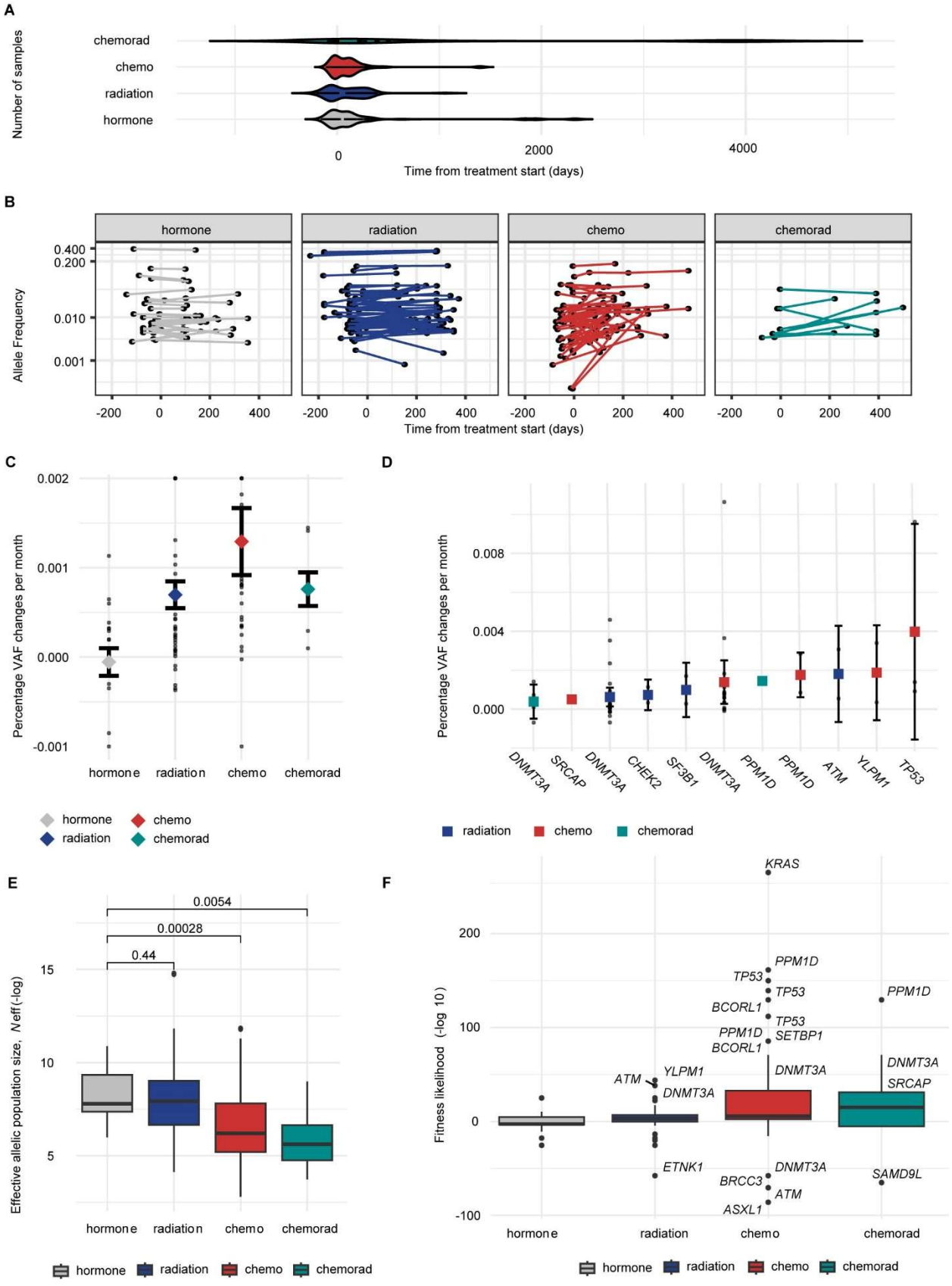

A
No. of mutations
Frameshift insertion
Frameshift deletion
Inframe deletion
Missense substitution
Nonsense substitution
>1 mutations
No. of patients
DNMT3A
YLPM1
PPM1D
TET2
ASXL1
ZNF318
FAT1
NF1
SRCAP
TP53
ATM
BCOR
CHEK2
GNAS
KMT2D
SH2B3
SMC1A
BCORL1
BRCC3
JAK2
KRAS
SAMD9L
SETBP1
SF3B1
SMC3
TERT
ANKRD26
ARID1B
DDX41
ETNK1
ETV6
GNB1
IKZF1
KDM6A
KIT
KMT2A
NOTCH1
PTPN11
STAG2
WT1
B
Excess of nonsense mutation (fold)
C
D
No. of patients with CHIP
E
Rank-sum P = 0.0012
VAF Coefficient of variation
CH
CHIP

FIGURE 2. Mutational spectrum and diversity of CH prior to treatment for breast cancer. A) Oncoplot of CH mutations detected across a panel of 81 genes in patients with breast cancer. Columns represent individual patients and colors correspond to the mutation type, including frameshift insertion and deletions, in-frame deletions, and missense and nonsense single nucleotide variants. B) Fold excess of nonsense mutations in each gene relative to the expected distribution from the genes' nucleotide content; chi-squared statistics shown with 95% confidence intervals for genes with ≥3 protein-changing substitutions and ≥1 nonsense mutation. C) CH mutation co-occurrence illustrated by a circos plot. Colored lines highlight the most frequently mutated genes and co-mutations. D) Number of patients with CHIP mutations (VAF ≥2%) across genes. E) Divergence in CH allelic frequency, defined by mean CH VAF per patient, compared between patients with CHIP mutations (VAF ≥2%) versus patients with CH mutations (VAF <2%).


A
Number of samples
chemorad
chemo
radiation
hormone
0
2000
4000
Time from treatment start (days)
B
hormone
radiation
chemo
chemorad
Allele Frequency
0.400
0.200
0.010
0.001
-200
0
200
400
Time from treatment start (days)
C
Percentage VAF changes per month
0.002
0.001
0.000
-0.001
hormone
radiation
chemo
chemorad
D
Percentage VAF changes per month
0.008
0.004
0.000
DNMT3A
SRCAP
DNMT3A
CHEK2
SF3B1
DNMT3A
PPM1D
PPM1D
ATM
YLPM1
TP53
radiation
chemo
chemorad
E
0.0054
0.00028
0.44
Effective allelic population size, Neff (-log)
15
10
5
hormone
radiation
chemo
chemorad
F
Fitness likelihood (-log 10)
KRAS
200
100
0
-100
PPM1D
TP53
TP53
BCORL1
TP53
PPM1D
SETBP1
BCORL1
DNMT3A
ATM
YLPM1
DNMT3A
ETNK1
DNMT3A
BRCC3
ATM
ASXL1
PPM1D
DNMT3A
SRCAP
SAMD9L
hormone
radiation
chemo
chemorad

570

FIGURE 3. CH mutational dynamics during breast cancer treatment. A) Number and temporal spread of longitudinal samples collected by treatment modality; times are presented relative to treatment start (time=0). B) Measured variant allele frequency (VAF) for all detected CH mutations during the study by treatment modality. C) Standardized percent change in VAF per month for all detected mutations by treatment modality. D) Standardized percent change in VAF per month for gene-specific mutations showing significant allelic increase per treatment modality. E) Effective allelic population size (N_{eff}) across treatment modalities (rank-sum test). F) Fitness likelihood, defined by the minus log of probability of observing post-treatment VAF of a CH mutation given the measured effective allelic population size per patient, across treatment modalities; genes with mutations showing the highest CH clonal change relative to the wild-type population are indicated.

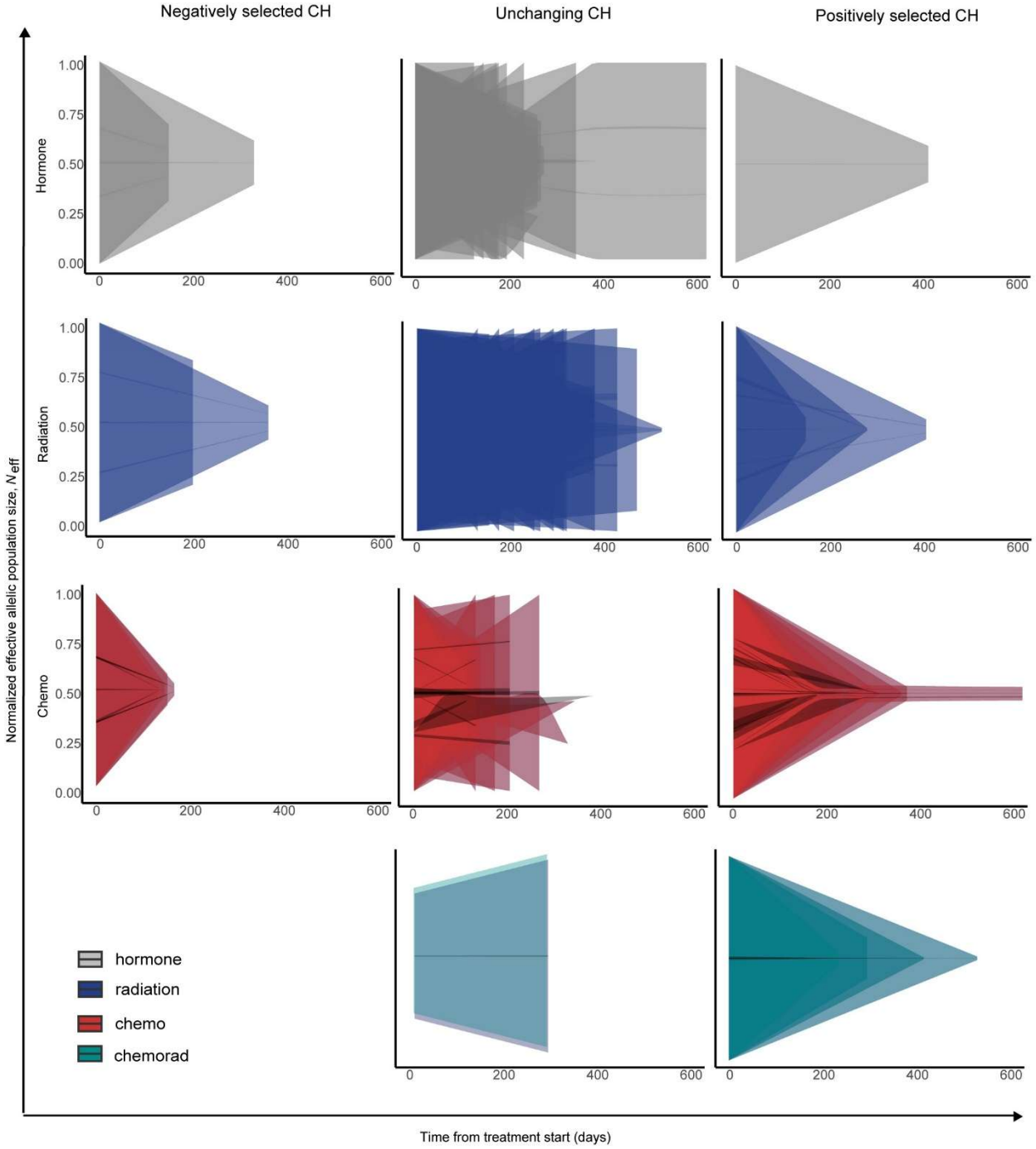
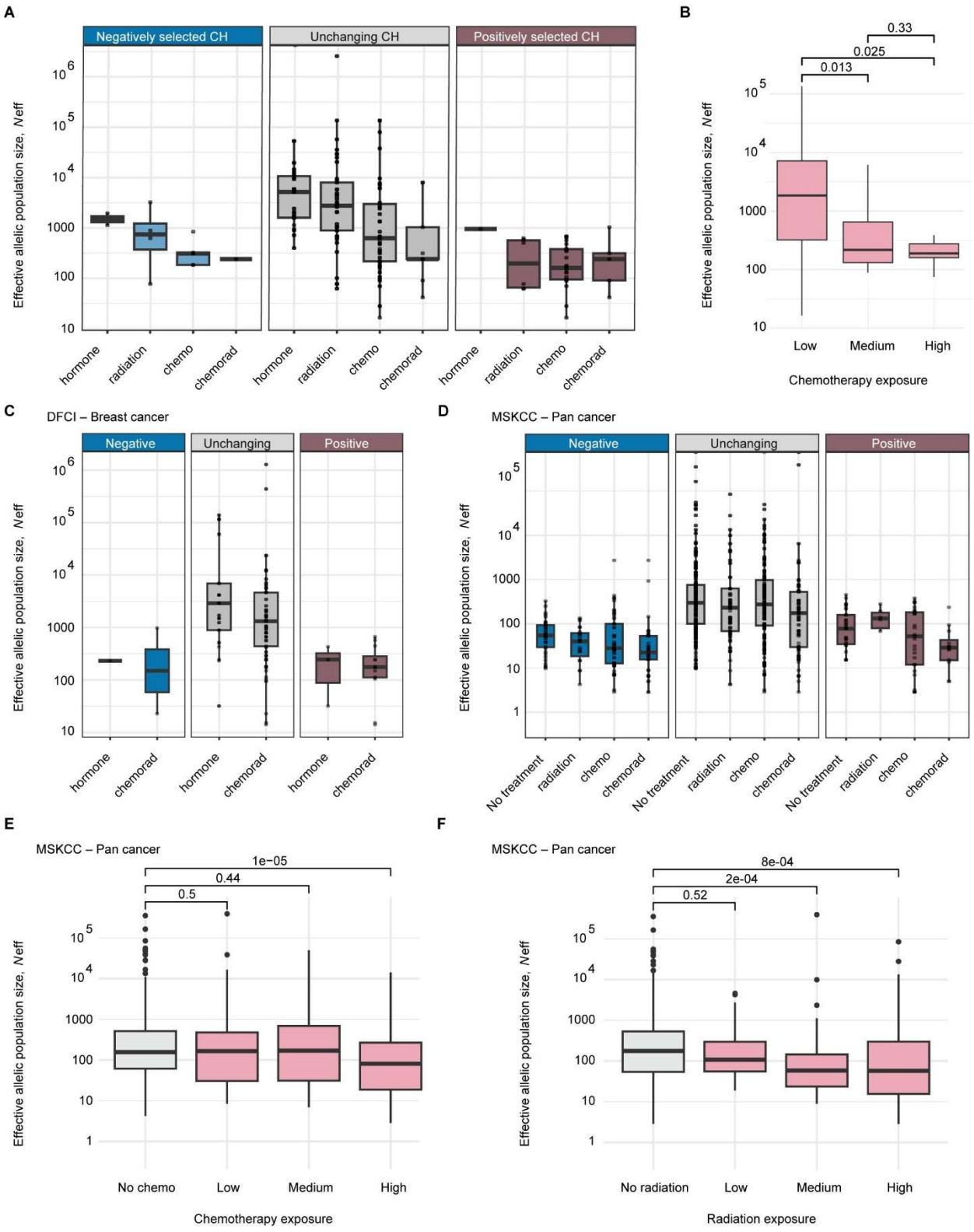


FIGURE 4. Allelic population bottlenecks imposed by cancer treatment. Schematics showing effective allelic population size (N_{eff}) across treatment modalities, normalized by mean N_{eff} in patients treated with hormonal therapy only,.



587

588 **FIGURE 5. Effective allelic populations associate with treatment modality and dosage.** A) Effective
589 allelic population sizes (N_{eff}) by treatment modalities, grouped based on CH mutational dynamics, including

positively selected, unchanging, and negatively selected CH. B) N_{eff} distributions across cumulative cytotoxic chemotherapy exposure levels (rank-sum test). C) N_{eff} in patients with breast cancer treated at the Dana Farber Cancer Institute (DFCI), by treatment modalities, and grouped based on CH mutational dynamics, including positively selected, unchanging, and negatively selected CH. D) N_{eff} in patients with solid tumors treated at Memorial Sloan Kettering Cancer Center (MSKCC), by treatment modalities, and grouped based on CH mutational dynamics, including positively selected, unchanging, and negatively selected CH. E) N_{eff} distributions across cumulative chemotherapy exposure levels in solid tumors treated at MSKCC (rank-sum test). F) N_{eff} distributions across cumulative radiotherapy exposure levels in solid tumors treated at MSKCC (rank-sum test).

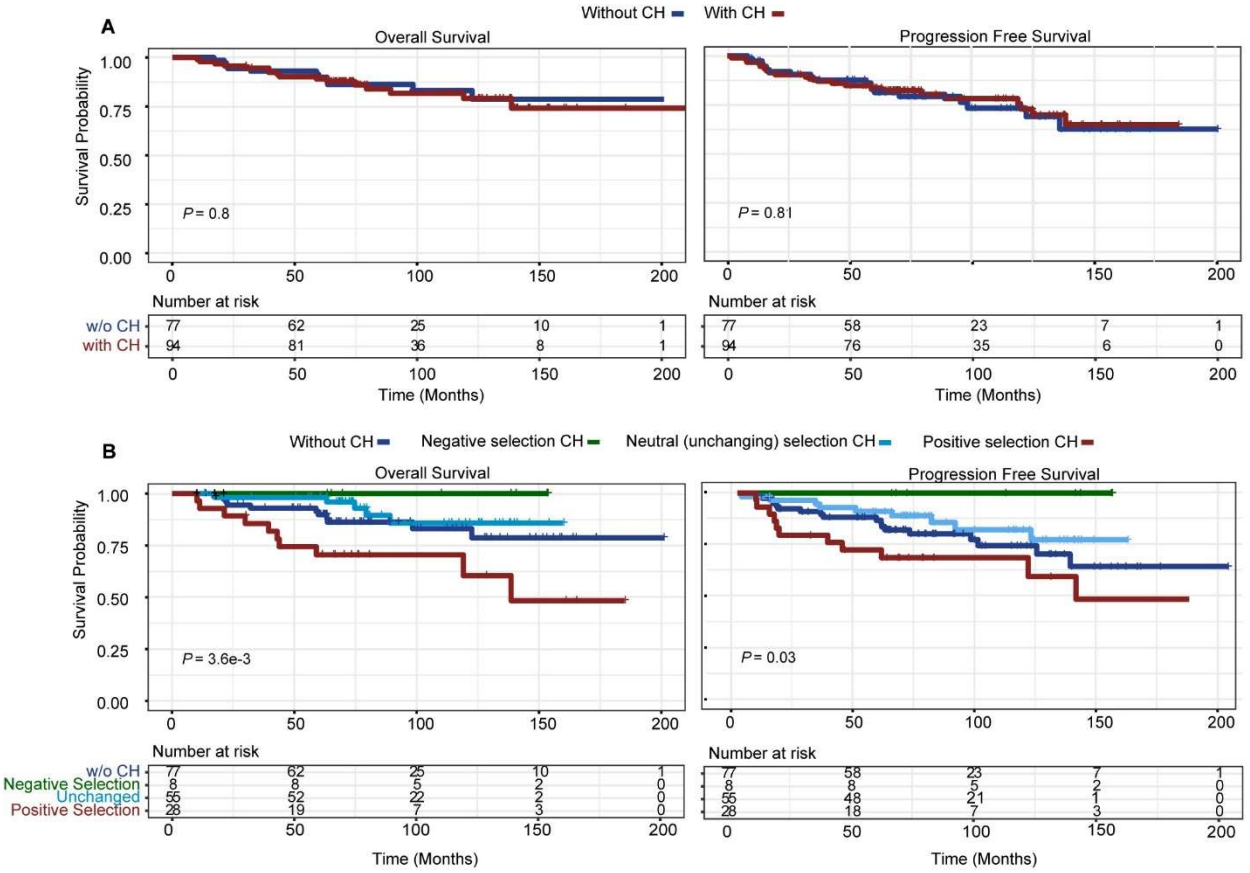


FIGURE 6. CH mutational dynamics and patient survival. A) Difference in overall survival (OS) and progression-free survival (PFS) between patients with and without CH (log-rank test). B) Differences in OS and PFS between patients with positively selected, negatively selected, unchanging, and no CH (log-rank test).

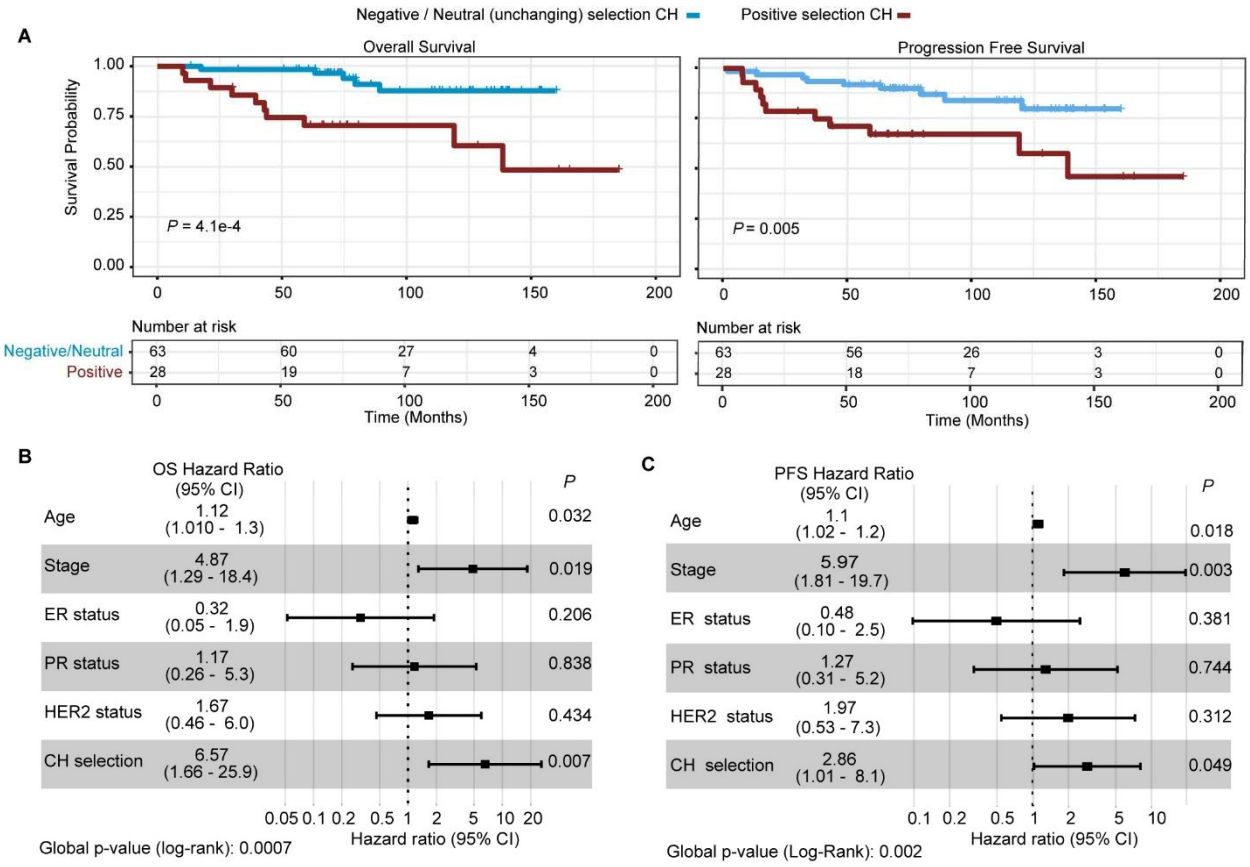


FIGURE 7. Positive selection of CH impacts patient survival. A) Differences in OS and PFS between patients with positively selected CH and those with negatively selected or unchanging CH (log-rank test). B) The impact of positively selected CH on OS using a multivariable Cox regression model including patient age, tumor stage and ER, PR, and HER2 hormonal status. C) The impact of positively selected CH on PFS using a multivariable Cox regression model including patient age, tumor stage and ER, PR, and HER2 hormonal status. Kaplan-Meier plots with log-rank *P* statistics are shown.

DATA AVAILABILITY

Raw sequencing data of peripheral blood from the discovery cohort (n=171 patients) were deposited in the dbGaP repository (phs004283.v1.p1). All data used in the figures and throughout the manuscript are provided in the Supporting Data Values and Supplementary Tables.

CODE AVAILABILITY

All code to reproduce and calculate the evolutionary fitness and selection pressure strength are available on GitHub at https://github.com/marabzadeh/EchoCH.

FUNDING

This manuscript is the result of funding in whole or in part by the National Institutes of Health (NIH). It is subject to the NIH Public Access Policy. Through acceptance of this federal funding, NIH has been given a right to make this manuscript publicly available in PubMed Central upon the Official Date of Publication, as defined by NIH. This work was supported by the National Cancer Institute grant R01-CA233662 (HK, SG, and NG), the Total Cancer Care Protocol and, in part, by the Collaborative Data Services Core Facility, Tissue Core Facility, and Molecular Genomics Core Facility at the H. Lee Moffitt Cancer Center & Research Institute (P30-CA076292, NG), Breast Cancer Research Foundation, and Rutgers Cancer Institute Biomedical Informatics Shared Resource (P30-CA072720-5917, HK and SG). MA was supported by the New Jersey Commission on Cancer Research (COCR24PDF008). SMo was supported by the American-Italian Cancer Foundation, Fondazione Gianni Bonadonna and Associazione Italiana per la Ricerca contro il Cancro, Saverin Family Fund.

AUTHOR CONTRIBUTION

NG, HK, and SG conceived and supervised the study. MA performed evolutionary modeling; MA and YT conducted sequencing analyses, variant calling, statistical association analyses, and visualization of the results with help from VM and MT. CCL and JK assisted with identifying the patient cohort. CCL extracted and analyzed clinical data with help from DW and EH. SMa visualized therapeutic bottleneck results with supervision from JW. SMo, HAP, and JEG generated and analyzed the Dana-Farber Cancer Institute cohort. All authors drafted the manuscript. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

SG has been a consultant for KayoThera, Lunit, Ipsen, Roche, Merck, Foghorn Therapeutics, and EQRX, and has received research funding from Gandeeva and M2GEN. SMo reports consulting or advisory roles at Daiichi-Sankyo, institutional research funding from Precede Biosciences and Merck. HK is a full-time employee of Regeneron Pharmaceuticals. All other authors declare no conflicts of interest.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Gregory Riedlinger, Alexandra Jacunski, and Daniel Herranz for the careful reading of the manuscript and for providing constructive feedback.

TABLE 1. Clinical characteristics of patients with breast cancer in the discovery cohort

	Overall, N = 171[1]	CH, N = 94[1]	No CH, N = 77[1]
Age	58 (50, 65)	62 (55, 67)	52 (47, 60)
Sex			
Female	171 (100%)	94 (100%)	77 (100%)
Race			
Asian	3 (1.8%)	3 (3.2%)	0 (0%)
Black	10 (5.8%)	4 (4.3%)	6 (7.8%)
White	151 (88%)	85 (90%)	66 (86%)
Other race[2]	7 (4.1%)	2 (2.1%)	5 (6.5%)
Ethnicity			
Hispanic	23 (13%)	7 (7.4%)	16 (21%)
Non-Hispanic	148 (87%)	87 (93%)	61 (79%)
Smoking status			
Current	9 (5.3%)	7 (7.4%)	2 (2.6%)
Former	61 (36%)	36 (38%)	25 (32%)
Never	101 (59%)	51 (54%)	50 (65%)
Histology			
Invasive ductal carcinoma	114 (67%)	60 (64%)	54 (70%)
Invasive lobular carcinoma	21 (12%)	14 (15%)	7 (9.1%)
Ductal carcinoma in situ	15 (8.8%)	10 (11%)	5 (6.5%)
Mixed invasive carcinoma	11 (6.4%)	5 (5.3%)	6 (7.8%)
Special subtypes of invasive carcinoma[3]	10 (5.8%)	5 (5.3%)	5 (6.5%)
HR/HER2 status[4]			
HR+/HER2-	118 (69%)	61 (65%)	57 (74%)
HR+/HER2+	15 (8.8%)	7 (7.4%)	8 (10%)
HR-/HER2+	4 (2.3%)	4 (4.3%)	0 (0%)
TNBC	23 (13%)	14 (15%)	9 (12%)
HR+/HER2 untested	11 (6.4%)	8 (8.5%)	3 (3.9%)
Tumor stage at diagnosis, Node, Metastasis stage			
0	15 (8.8%)	10 (11%)	5 (6.5%)
1	81 (47%)	42 (45%)	39 (51%)
2	55 (32%)	32 (34%)	23 (30%)
3	13 (7.6%)	7 (7.4%)	6 (7.8%)
4	7 (4.1%)	3 (3.2%)	4 (5.2%)
Treatment between samples			
Chemotherapy (± hormone therapy)	69 (40%)	33 (35%)	36 (47%)
Radiation therapy (± hormone therapy)	55 (32%)	33 (35%)	22 (29%)
Chemotherapy + radiation (± hormone therapy)	12 (7.0%)	8 (8.5%)	4 (5.2%)
Hormone therapy only	35 (20%)	20 (21%)	15 (19%)
Days between treatment start and first sample	-41 (-66, -14)	-44 (-69, -15)	-41 (-62, -14)

[1] Median (IQR); n (%)
[2] Other races include patient-reported American Indian or Alaska Native (n=1), Laotian (n=1), or “Other race” (n=5)
[3] Tubular carcinoma, mucinous carcinoma, solid papillary carcinoma with invasion, or metaplastic carcinoma
[4] HR: hormone receptor, + indicates positivity for estrogen and/or progesterone receptor (cutoff ≥1%); TNBC: triple-negative breast cancer

SUPPLEMETARY FIGURE LEGENDS

SUPPLEMETARY FIGURE S1. Schematics showing the study design, sampling timeline criteria, and evolutionary models considered for CH during breast cancer treatment.

SUPPLEMETARY FIGURE S2. Swimmer plots showing the sampling timeline relative to breast cancer diagnosis and treatment schedule per patient. A) Patients with overall survival (OS) ≤ 60 months are included. B) Patients with overall survival (OS) 60-120 months are included. C) Patients with overall survival (OS) > 120 months are included.

SUPPLEMETARY FIGURE S3. Clonal diversity and CH mutation-specific changes under treatment for breast cancer. A) Clonal diversity after excluding mutations with VAF ≥ 2% for both with and without CHIP-defined mutations groups. B) Standardized percent change in variant allele frequency (VAF) per month for *DNMT3A*, *YLPM1*, *TP53* and *ATM* mutations by treatment modality. C) Change in VAF for CH mutations that grow to CHIP-defining VAF during treatment.

SUPPLEMETARY FIGURE S4. CH dynamics spectrum. A) The number of variants with positive (Pos) or negative (Neg) selection of CH mutations across treatment modalities and genes. B) Mutational domain spectrum of DNMT3A in the 3 groups of negatively selected, positively selected, or unchanging CH.

SUPPLEMETARY FIGURE S5. Schematics showing effective allelic population size (N_{eff}) during treatment for patients with A) unchanging CH, B) positively selected CH, or C) negatively selected CH normalized by mean N_{eff} in cases treated with hormonal therapy only, across treatment modalities. D) Percentage and number of patients across cumulative chemotherapy exposure levels stratified based on CH mutational dynamics. E) Change in VAF for CH mutations in patients with both positive and negative selection.

SUPPLEMETARY FIGURE S6. CH mutational dynamics in the Dana Farber Cancer Institute early-stage breast cancer cohort. A) Effective allelic population size (N_{eff}) across treatment modalities. B) Percent change in variant allele frequency (VAF) per month for CH mutations by treatment modality. C) REMARK diagram for the DFCI cohort.

SUPPLEMETARY FIGURE S7. CH mutational dynamics in the Memorial Sloan Kettering Cancer Center pan-cancer cohort. A) Effective allelic population size (N_{eff}) across treatment modalities. A) Percent change in variant allele frequency (VAF) per month for CH mutations by treatment modality. C) Number and percentage of patients across cumulative chemotherapy exposure

levels stratified by CH mutational dynamics. D) Number and percentage of patients across cumulative radiotherapy exposure levels stratified by CH mutational dynamics.

SUPPLEMETARY FIGURE S8. Difference in overall survival (OS) and progression free survival (PFS) between patients with positively selected CH and those with either negatively selected or unchanging CH. A) Including only patients with early-stage disease. B) Excluding patients with *TP53*-mutated CH. C) Including patients treated only with chemotherapy during examined period. Kaplan-Meier plots with log-rank *P* statistics are shown.

SUPPLEMETARY FIGURE S9. Difference in overall survival (OS) and progression free survival (PFS) between patients with positively selected CH and those with negatively selected or unchanging CH within homogenous treatment subsets of the cohort. Subsets were selected based on the most common chemotherapy regimens received. A) Including patients receiving cyclophosphamide, doxorubicin, and paclitaxel (n=14). B) Including patients receiving cyclophosphamide and docetaxel (n=11). C) Including patients receiving low levels of cyclophosphamide and docetaxel (n=8). Kaplan-Meier plots with log-rank *P* statistics are shown.

SUPPLEMETARY FIGURE S10. Sensitivity analyses for allelic population size (N_{eff}) assumptions. A) Coefficient of variation for N_{eff} calculated across the cohort at different sequencing depths. B) N_{eff} stratified by treatment modality and, C) dynamic selection groups; with the assumption of clonal dependency and considering only the largest clone in each sample.

SUPPLEMETARY TABLES

TABLE S1. List of 81 genes included in targeted sequencing

TABLE S2. List of detected CH mutations in the cohort

TABLE S3. Statistics for evaluation of excess of nonsense CH mutations per gene

TABLE S4. Summary of gene level change in CH VAF per treatment

TABLE S5. Statistics for evaluating CH mutation-specific fitness relative to wild-type

TABLE S6. Summary of results for survival analyses using multivariable Cox regression models

REFERENCES

1. Alaggio R, Amador C, Anagnostopoulos I, Attygalle AD, Araujo IBdO, Berti E, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: lymphoid neoplasms. Leukemia. 2022;36(7):1720-48.
2. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. New England Journal of Medicine. 2014;371(26):2488-98.
3. Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. New England Journal of Medicine. 2014;371(26):2477-87.
4. Ahmad H, Jahn N, Jaiswal S. Clonal hematopoiesis and its impact on human health. Annual Review of Medicine. 2023;74(1):249-60.
5. Coombs CC, Zehir A, Devlin SM, Kishtagari A, Syed A, Jonsson P, et al. Therapy-related clonal hematopoiesis in patients with non-hematologic cancers is common and associated with adverse clinical outcomes. Cell stem cell. 2017;21(3):374-82. e4.
6. Coombs CC, Gillis NK, Tan X, Berg JS, Ball M, Balasis ME, et al. Identification of clonal hematopoiesis mutations in solid tumor patients undergoing unpaired next-generation sequencing assays. Clinical Cancer Research. 2018;24(23):5918-24.
7. Severson EA, Riedlinger GM, Connelly CF, Vergilio J-A, Goldfinger M, Ramkissoon S, et al. Detection of clonal hematopoiesis of indeterminate potential in clinical sequencing of solid tumor specimens. Blood, The Journal of the American Society of Hematology. 2018;131(22):2501-5.
8. Pich O, Bernard E, Zagorulya M, Rowan A, Pospori C, Slama R, et al. Tumor-infiltrating clonal hematopoiesis. New England Journal of Medicine. 2025;392(16):1594-608.
9. Bernstein N, Spencer Chapman M, Nyamondo K, Chen Z, Williams N, Mitchell E, et al. Analysis of somatic mutations in whole blood from 200,618 individuals identifies pervasive positive selection and novel drivers of clonal hematopoiesis. Nature Genetics. 2024;56(6):1147-55.
10. Watson CJ, Papula A, Poon GY, Wong WH, Young AL, Druley TE, et al. The evolutionary dynamics and fitness landscape of clonal hematopoiesis. Science. 2020;367(6485):1449-54.
11. Mitchell E, Spencer Chapman M, Williams N, Dawson KJ, Mende N, Calderbank EF, et al. Clonal dynamics of haematopoiesis across the human lifespan. Nature. 2022;606(7913):343-50.
12. Robertson NA, Latorre-Crespo E, Terradas-Terradas M, Lemos-Portela J, Purcell AC, Livesey BJ, et al. Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness effects. Nature medicine. 2022;28(7):1439-46.
13. Uddin MM, Zhou Y, Bick AG, Burugula BB, Jaiswal S, Desai P, et al. Longitudinal profiling of clonal hematopoiesis provides insight into clonal dynamics. Immunity & Ageing. 2022;19(1):23.
14. van Zeventer IA, de Graaf AO, Salzbrunn JB, Nolte IM, Kamphuis P, Dinmohamed A, et al. Evolutionary landscape of clonal hematopoiesis in 3,359 individuals from the general population. Cancer cell. 2023;41(6):1017-31. e4.
15. Bolton KL, Ptashkin RN, Gao T, Braunstein L, Devlin SM, Kelly D, et al. Cancer therapy shapes the fitness landscape of clonal hematopoiesis. Nature genetics. 2020;52(11):1219-26.
16. Abelson S, Collord G, Ng SW, Weissbrod O, Mendelson Cohen N, Niemeyer E, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. Nature. 2018;559(7714):400-4.
17. Fabre MA, de Almeida JG, Fiorillo E, Mitchell E, Damaskou A, Rak J, et al. The longitudinal dynamics and natural history of clonal haematopoiesis. Nature. 2022;606(7913):335-42.
18. Gillis NK, Ball M, Zhang Q, Ma Z, Zhao Y, Yoder SJ, et al. Clonal haemopoiesis and therapy-related myeloid malignancies in elderly patients: a proof-of-concept, case-control study. The lancet oncology. 2017;18(1):112-21.
19. Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nature medicine. 2014;20(12):1472-8.
20. Takahashi K, Wang F, Kantarjian H, Doss D, Khanna K, Thompson E, et al. Preleukaemic clonal haemopoiesis and risk of therapy-related myeloid neoplasms: a case-control study. The lancet oncology. 2017;18(1):100-11.
21. Takahashi K, Nakada D, Goodell M. Distinct landscape and clinical implications of therapy-related clonal hematopoiesis. The Journal of Clinical Investigation. 2024;134(19).

22. Morganti S, Gibson CJ, Jin Q, Santos K, Patel A, Wilson A, et al. Prevalence, dynamics, and prognostic role of clonal hematopoiesis of indeterminate potential in patients with breast cancer. Journal of Clinical Oncology. 2024;42(31):3666-79.
23. Nead KT, Kim T, Joo L, McDowell TL, Wong JW, Chan IC, et al. Impact of cancer therapy on clonal hematopoiesis mutations and subsequent clinical outcomes. Blood Advances. 2024;8(19):5215-24.
24. Mayerhofer C, Sedrak MS, Hopkins JO, Li T, Tayob N, Faggen MG, et al. Clonal hematopoiesis in older patients with breast cancer receiving chemotherapy. JNCI: Journal of the National Cancer Institute. 2023;115(8):981-8.
25. Arends CM, Kopp K, Hablesreiter R, Estrada N, Christen F, Moll UM, et al. Dynamics of clonal hematopoiesis under DNA-damaging treatment in patients with ovarian cancer. Leukemia. 2024;38(6):1378-89.
26. Bowman RL, Busque L, Levine RL. Clonal hematopoiesis and evolution to hematopoietic malignancies. Cell stem cell. 2018;22(2):157-70.
27. Pich O, Cortes-Bullich A, Muiños F, Pratcorona M, Gonzalez-Perez A, Lopez-Bigas N. The evolution of hematopoietic cells under cancer therapy. Nature communications. 2021;12(1):4803.
28. Park SJ, Bejar R. Clonal hematopoiesis in cancer. Experimental hematology. 2020;83:105-12.
29. Pich O, Reyes-Salazar I, Gonzalez-Perez A, Lopez-Bigas N. Discovering the drivers of clonal hematopoiesis. Nature communications. 2022;13(1):4267.
30. Kar SP, Quiros PM, Gu M, Jiang T, Mitchell J, Langdon R, et al. Genome-wide analyses of 200,453 individuals yield new insights into the causes and consequences of clonal hematopoiesis. Nature Genetics. 2022;54(8):1155-66.
31. Weeks LD, Ebert BL. Causes and consequences of clonal hematopoiesis. Blood. 2023;142(26):2235-46.
32. Loh JW, Khiabanian H. Leukemia's clonal evolution in development, progression, and relapse. Current Stem Cell Reports. 2019;5:73-81.
33. Emmett K, Lee A, Rabadan R. High-resolution genomic surveillance of 2014 Ebolavirus using shared subclonal variants. PLoS Curr. 2015.
34. Sobel Leonard A, Weissman DB, Greenbaum B, Ghedin E, Koelle K. Transmission bottleneck size estimation from pathogen deep-sequencing data, with an application to human influenza A virus. Journal of virology. 2017;91(14):10.1128/jvi. 00171-17.
35. Kahn JD, Miller PG, Silver AJ, Sellar RS, Bhatt S, Gibson C, et al. PPM1D-truncating mutations confer resistance to chemotherapy and sensitivity to PPM1D inhibition in hematopoietic cells. Blood, The Journal of the American Society of Hematology. 2018;132(11):1095-105.
36. Hientz K, Mohr A, Bhakta-Guha D, Efferth T. The role of p53 in cancer drug resistance and targeted chemotherapy. Oncotarget. 2016;8(5):8921.
37. Cai XB, Robert L. & Trowbridge, Jennifer J. . Clonal hematopoiesis in myeloid malignancies and solid tumors. Nat Cancer. 2025.
38. Silver AJ, Bick AG, Savona MR. Germline risk of clonal haematopoiesis. Nature Reviews Genetics. 2021;22(9):603-17.
39. Kessler MD, Damask A, O'Keeffe S, Banerjee N, Li D, Watanabe K, et al. Common and rare variant associations with clonal haematopoiesis phenotypes. Nature. 2022;612(7939):301-9.
40. Florez MA, Tran BT, Wathan TK, DeGregori J, Pietras EM, King KY. Clonal hematopoiesis: Mutation-specific adaptation to environmental change. Cell stem cell. 2022;29(6):882-904.
41. Hsu JI, Dayaram T, Tovy A, De Braekeleer E, Jeong M, Wang F, et al. PPM1D mutations drive clonal hematopoiesis in response to cytotoxic chemotherapy. Cell stem cell. 2018;23(5):700-13. e6.
42. Beauchamp EM, Leventhal M, Bernard E, Hoppe ER, Todisco G, Creignou M, et al. ZBTB33 is mutated in clonal hematopoiesis and myelodysplastic syndromes and impacts RNA splicing. Blood cancer discovery. 2021;2(5):500-17.
43. Chen C-W, Zhang L, Dutta R, Niroula A, Miller PG, Gibson CJ, et al. SRCAP mutations drive clonal hematopoiesis through epigenetic and DNA repair dysregulation. Cell Stem Cell. 2023;30(11):1503-19. e8.
44. Guryanova OA, Shank K, Spitzer B, Luciani L, Koche RP, Garrett-Bakelman FE, et al. DNMT3A mutations promote anthracycline resistance in acute myeloid leukemia via impaired nucleosome remodeling. Nature medicine. 2016;22(12):1488-95.
45. Feng Y, Yuan Q, Newsome RC, Robinson T, Bowman RL, Zuniga AN, et al. Hematopoietic-specific heterozygous loss of Dnmt3a exacerbates colitis-associated colon cancer. Journal of Experimental Medicine. 2023;220(11):e20230011.

46. Yan B, Yuan Q, Kaur P, Mckee AR, Shabashvili DE, Guryanova OA. Clonal hematopoiesis driven by Dnmt3a mutations promotes metabolic disease development. bioRxiv. 2025:2025.05. 02.651715.
47. Shibata A, Jeggo PA. ATM's role in the repair of DNA double-strand breaks. Genes. 2021;12(9):1370.
48. Goglia AG, Mustion G, Modlin LA, Levine RL, Bolton KL, Braunstein LZ. Impact of radiotherapy site, modality, and dose on subsequent clonal hematopoiesis. Blood Advances. 2025.
49. Corbeau A, Kuipers SC, de Boer SM, Horeweg N, Hoogeman MS, Godart J, et al. Correlations between bone marrow radiation dose and hematologic toxicity in locally advanced cervical cancer patients receiving chemoradiation with cisplatin: a systematic review. Radiotherapy and Oncology. 2021;164:128-37.
50. Buttigieg MM, Vlasschaert C, Bick AG, Vanner RJ, Rauh MJ. Inflammatory reprogramming of the solid tumor microenvironment by infiltrating clonal hematopoiesis is associated with adverse outcomes. Cell Reports Medicine. 2025;6(3).
51. Rondeau V, Bansal S, Buttigieg MM, Zeng AG, Chan DY, Chan-Seng-Yue M, et al. Response to Immune Checkpoint Blockade is Enhanced in the Presence of Hematopoietic TET2 Inactivation. Cancer research. 2026:OF845-OF57.
52. Chan IC, Zhang P, Pan X, Castro C, Fox N, Lewis AM, et al. CDK4/6 inhibition mitigates chemotherapy-induced expansion of TP53-mutant clonal hematopoiesis. Nature Genetics. 2026:1-11.
53. Tucker MA, Meadows AT, Boice Jr JD, Stovall M, Oberlin O, Stone BJ, et al. Leukemia after therapy with alkylating agents for childhood cancer. Journal of the National Cancer Institute. 1987;78(3):459-64.
54. Whelan TJ, Kim D-H, Sussman J, editors. Clinical experience using hypofractionated radiation schedules in breast cancer. Seminars in radiation oncology; 2008: Elsevier.
55. Miller CA, Walker JR, Jensen TL, Hooper WF, Fulton RS, Painter JS, et al. Failure to detect mutations in U2AF1 due to changes in the GRCh38 reference sequence. The Journal of Molecular Diagnostics. 2022;24(3):219-23.
56. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.39972013.
57. Fennel T, Homer N, Genomics F. Fgbio: tools for working with genomic and high throughput sequencing data. URL https://githubcom/fulcrumgenomics/fgbio (accessed July 2025).
58. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010;20(9):1297-303.
59. Rabadan R, Bhanot G, Marsilio S, Chiorazzi N, Pasqualucci L, Khiabanian H. On statistical modeling of sequencing noise in high depth data to assess tumor evolution. Journal of statistical physics. 2018;172(1):143-55.
60. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 2011;27(21):2987-93.
61. Wilm A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, Wong CH, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic acids research. 2012;40(22):11189-201.
62. Morgulis A, Gertz EM, Schäffer AA, Agarwala R. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. Journal of computational biology. 2006;13(5):1028-40.
63. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-43.
64. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, et al. The NCBI dbGaP database of genotypes and phenotypes. Nature genetics. 2007;39(10):1181-6.
65. Liu X, Tang Y-H, Blachly J, Edge S, Jakubek YA, McCarter M, et al. qcCHIP: an R package to identify clonal hematopoiesis variants using cohort-specific data characteristics. Bioinformatics. 2025;41(9):btaf522.
66. Vlasschaert C, Mack T, Heimlich JB, Niroula A, Uddin MM, Weinstock J, et al. A practical approach to curate clonal hematopoiesis of indeterminate potential in human genetic data sets. Blood, The Journal of the American Society of Hematology. 2023;141(18):2214-23.
67. Marzban S, Stiehl T, Xie Z, Andersen M, Snyder J, Gudmand-Høyer J, et al. Modeling the evolutionary dynamics of clonal hematopoiesis. Nature genetics. 2026:1-11.
68. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome research. 2018;28(11):1747-56.

69. Gatenbee CD, Schenck RO, Bravo RR, Anderson AR. EvoFreq: visualization of the Evolutionary Frequencies of sequence and model data. BMC bioinformatics. 2019;20(1):710.
70. Fenstermacher DA, Wenham RM, Rollison DE, Dalton WS. Implementing personalized medicine in a cancer center. The Cancer Journal. 2011;17(6):528-36.