Nature. Author manuscript; available in PMC 2016 May 12.

Published in final edited form as:

Nature. 2015 February 19; 518(7539): 422–426. doi:10.1038/nature13952.

Dynamics of genomic clones in breast cancer patient xenografts at single cell resolution

Peter Eirew[1,2,*], **Adi Steif**[1,2,*], **Jaswinder Khattra**[1,2,*], **Gavin Ha**[1,2], **Damian Yap**[1,2], **Hossein Farahani**[1,2], **Karen Gelmon**[3], **Stephen Chia**[3], **Colin Mar**[3], **Adrian Wan**[1], **Emma Laks**[1,2], **Justina Biele**[1,2], **Karey Shumansky**[1], **Jamie Rosner**[1], **Andrew McPherson**[1,2], **Cydney Nielsen**[1,2], **Andrew J. L. Roth**[1,2], **Calvin Lefebvre**[1,2], **Ali Bashashati**[1,2], **Camila de Souza**[1], **Celia Siu**[1], **Radhouane Aniba**[1,2], **Jazmine Brimhall**[1], **Arusha Oloumi**[1,2], **Tomo Osako**[1,2], **Alejandra Bruna**[4,5], **Jose Sandoval**[4,5], **Teresa Algara**[1,2], **Wendy Greenwood**[4,5], **Kaston Leung**[12,13], **Hongwei Cheng**[6,7], **Hui Xue**[6,7], **Yuzhuo Wang**[6,7], **Dong Lin**[6,7], **Andrew J. Mungall**[8], **Richard Moore**[8], **Yongjun Zhao**[8], **Julie Lorette**[11], **Long Nguyen**[9,10], **David Huntsman**[2,11], **Connie J. Eaves**[9,10], **Carl Hansen**[12,13], **Marco A. Marra**[8], **Carlos Caldas**[4,5], **Sohrab P. Shah**[1,2,8], and **Samuel Aparicio**[1,2,8,11]

[1]Department of Molecular Oncology, BC Cancer Agency, 675 W10th Avenue, Vancouver, BC, V5Z 1L3, Canada

[2]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, V6T 2B5, Canada

[3]Department of Medical Oncology, BC Cancer Agency, 600 W10th Avenue, Vancouver, BC, V5Z 4E6, Canada

[4]Department of Oncology, University of Cambridge, Hills Road, Cambridge, CB2 2XZ, UK

[5]Cancer Research UK Cambridge Research Institute, University of Cambridge, Li Ka Shing Centre, Cambridge, CB2 0RE, UK

[6]Department of Experimental Therapeutics, BC Cancer Agency, Vancouver, BC, V5Z 1L3, Canada

[7]The Vancouver Prostate Centre, Vancouver General Hospital and Department of Urologic Sciences, University of British Columbia, Vancouver, BC, V5Z 1M9, Canada

[8]Michael Smith Genome Sciences Centre, Vancouver, BC, V5Z 1L3, Canada

Correspondence and requests for materials should be addressed to Samuel Aparicio (saparicio@bccrc.ca) or Sohrab P. Shah (sshah@bccrc.ca).

*denotes equal contribution.

Genome data has been deposited at the European Genome-phenome Archive (EGA,http://www.ebi.ac.uk/egaunder accession number EGAS00001000952. Processed data can be viewed at www.cbioportal.org.

[Competing Interests] The authors declare that they have no competing financial interests.

[Author contributions] SA and SS designed the study and supervised the research. SA, SS, PE and AS wrote the paper. PE, AB, JS, TA, WG, HC, HX, YW and DL performed transplants and passaging. AS, PE, GH, CN, HF, AJLR, CL, AB, CS, KS, JR, HF, RA, CD, SS and SA carried out analysis. JK, DY, EL, JB, AW, JE, KL, AM, AO, RM, YZ, CH and MM assisted with sequence generations and single cell experiments. TO, JL, DH contributed to histological analysis. CE, CH, MM, CC, SS and SA provided intellectual contributions to design or interpretation.

[9]Department of Medical Genetics, University of British Columbia, Vancouver, BC, V6T 1Z3, Canada

[10]Terry Fox Laboratory, BC Cancer Agency, Vancouver, BC, V5Z 1L3, Canada

[11]Centre for Translational and Applied Genomics, BC Cancer Agency, 600 West 10th Avenue, Vancouver, BC, V5Z 4E6, Canada

[12]Centre for High-Throughput Biology, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

[13]Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, V6T 1Z1, Canada

Abstract

Human cancers, including breast cancers, are comprised of clones differing in mutation content. Clones evolve dynamically in space and time following principles of Darwinian evolution[1,2], underpinning important emergent features such as drug resistance and metastasis[3–7]. Human breast cancer xenoengraftment is used as a means of capturing and studying tumour biology, and breast tumour xenografts are generally assumed to be reasonable models of the originating tumours[8–10]. However the consequences and reproducibility of engraftment and propagation on the genomic clonal architecture of tumours has not been systematically examined at single cell resolution. Here we show by both deep genome and single cell sequencing methods, the clonal dynamics of initial engraftment and subsequent serial propagation of primary and metastatic human breast cancers in immunodeficient mice. In all 15 cases examined, clonal selection on engraftment was observed in both primary and metastatic breast tumours, varying in degree from extreme selective engraftment of minor (<5% of starting population) clones to moderate, polyclonal engraftment. Furthermore, ongoing clonal dynamics during serial passaging is a feature of tumours experiencing modest initial selection. Through single cell sequencing, we show that major mutation clusters estimated from tumour population sequencing relate predictably to the most abundant clonal genotypes, even in clonally complex and rapidly evolving cases. Finally, we show that similar clonal expansion patterns can emerge in independent grafts of the same starting tumour population, indicating that genomic aberrations can be reproducible determinants of evolutionary trajectories. Our results show that measurement of genomically defined clonal population dynamics will be highly informative for functional studies utilizing patient-derived breast cancer xenoengraftment.

To evaluate xenograft clonal dynamics (see Table S1 for definitions of terms used) we generated 30 xenograft lines by serially transplanting (up to 16 generations over 3 years) breast cancer tissue organoid suspensions from 55 patients (Table S2, Extended Figure E1, Figure S1) into highly immunodeficient *NOD*/*SCID*/*IL2rγ*$^{-/-}$ (NSG) and *NOD*/*RAG1*$^{-/-}$*IL2rγ*$^{-/-}$ (NRG) mice[11] (details in the Supplementary Information). We carried out massively parallel whole genome shotgun sequencing (WGSS) on DNA from xenograft passages of 15 patient lines (10 primary tumour-derived and five pleural effusion-derived), along with matched patient tumour and normal DNA (47 samples total, median sequencing depth 45.1,Table S3). For these plus 56 additional xenograft passage samples, we validated 3187 somatic single nucleotide variant (SNV) positions (100–300 per tumour-xenograft

series) and 132 structural variant (SV) positions by deep targeted amplicon sequencing (Table S4, Table S5, Table S6), quantifying allele ratios to high precision. We surveyed the copy number alteration (CNA) and loss of heterozygosity (LOH) landscapes by Affymetrix SNP6.0 array (Table S7, Table S8). The mutation load of somatic SNVs (range: 4.3–27.7×10^3 genome-wide, 57-1040 in coding regions), CNA and LOH (34–67% of genome) and SVs in the 15 tumour-xenograft series (Figure S2, Table S9, Figure S3) were consistent with previous genome-wide breast cancer studies[4, 12–17], though low tumour cellularity hindered mutation discovery in SA429 and SA496 originating tumours. Tumour-xenograft pairs displayed comparable nucleotide substitution patterns (Figure S2, Figure S4), suggesting mutational processes are maintained post-engraftment.

To determine the extent of evolution in the SNV landscape, we first compared the genome-wide variant allele prevalences (VAP, the proportion of aligned reads at the SNV position with the variant base, see Table S1) from WGSS data in xenograft relative to tumour (SA429 and SA496 excluded due to low tumour cellularity). As expected, sizeable proportions (range 53.0–92.9%) of high-confidence SNVs are shared in tumour-xenograft pairs, with prevalences lying on a scatter plot diagonal indicating neutral dynamics (Extended Figure E2a, Figure S5a, Figure S6). Strikingly, all 15 samples also show clusters of SNVs prevalent in the xenograft while at or below the limit of detection in the tumour (range 6.5–32.1% of SNVs, see e.g. SA494, SA495, SA499), and vice versa (range 0.2–19.4%, see e.g. SA494, SA495, SA500), implying clonal selection on initial engraftment. Tumours and xenografts from SA494, SA495, SA499, SA500 and SA530 also exhibited substantial differences in SV content (Figure S7, Figure S3).

To resolve clonal dynamics and genotypes, we applied a Bayesian clustering model (PyClone[4, 18]) to SNV VAPs measured by targeted deep sequencing, accounting for the effect of copy number, LOH status and cellularity. SNVs with co-varying estimates of cellular prevalence (the proportion of tumour or xenograft cells bearing the mutation) across all time points are grouped into putative mutation clusters (Table S1). Consistent with the raw VAP measurements, several cases contained mutation clusters with high (75–100%) prevalences in the xenografts and low (0–15%) prevalences in the tumours, implying expansion of initially minor clones to dominate the xenograft (e.g. clusters 3, 4, 3, 2, 8, 2, and 2 in SA494, SA495, SA500, SA530, SA532, SA533, and SA535) (Extended Figure E2b, Figure S5b). Other series (SA493, SA499, SA501, SA531, SA534, SA536) demonstrated non-neutral clonal dynamics but involving alleles occupying much smaller proportions of total cellular populations. Interestingly, polyclonal population structure specific to the xenograft was observed after initial expansion in SA493, SA494, SA495, SA500, and SA531, suggesting initial selection on engraftment remains permissive to additional clonal evolution (Extended Figure E2b, Figure S5b). Polyclonal engraftment was evident in SA493, SA501, SA531 and SA532, suggesting multiple clones maintained their fitness post-engraftment.

Analogously, we analyzed clonal dynamics using CNAs as clonal marks, applying a probabilistic model (TITAN[19]) that infers CNA and LOH from WGSS data, accounting for mixtures of tumour and normal cells and reporting estimates of mutation cellular prevalence and mutation cluster membership (Table S10). Despite conservation of complex disruptions,

such as chromothripsis in SA429 (Figure S8) and breakage-fusion-bridge cycles in SA429 and SA494 (Figure S9,Figure S10), we identified substantial differences in copy number architecture between tumour and xenograft in all cases (Extended Figure E2c, Figure S5c). These included a xenograft-specific deletion event containing *TP53* (in SA500) that coincided with retention of a somatic SNV (Figure S11, Table S6). Notably, the predominant clonal dynamic (minor subclone expansion in SA494, SA495, SA532 and SA533; polyclonal engraftment in SA493, SA501) mirrored those seen in SNV space.

We next asked how clonal dynamics differ after initial engraftment, using PyClone predictions over serial passage generations spanning up to 3 years (Extended Figure E1). We distinguished statistically significant directional clonal dynamics by testing overlap of 90% credible intervals derived from Bayesian posterior probability distributions (Figure 1). Cases showing strongest clonal dynamics in the first engraftment passages (e.g. SA500, SA530, SA494, SA535) exhibited more stable prevalence over subsequent passages. In contrast, cases showing moderate initial clonal dynamics showed more dramatic subsequent dynamics (e.g. mutation clusters 2, 3 and 8 of SA501), in some cases leading to gradual expansion of a minor clone to dominate the xenograft over serial passages. We noted examples of all ER/HER2 subtypes and primary/metastatic cancers evolving by these two different modes. Some mutation clusters showed non-dynamic patterns over time (e.g. clusters 1, 4, 6 of SA500, clusters 1–3, 5, 7, 9–10 in SA532, as well as the highest prevalence clusters representing putative ancestral mutations that remained invariant, as expected). For two cases we noted preferential engraftment of initial transplants in MFP over SR site (SA496 4/4 MFP vs 0/4 SR; SA429 2/4 MFP vs 0/4 SR, Extended Figure E1). However, transplant site changes in *established* xenografts were not associated with unusually strong clonal dynamics (Figure 1, see SA495 X3-4, SA499 X3-4, SA429 X1-2, SA496 X1-2).

To directly validate the population-based inference of mutation clusters and clonal genotypes, we carried out single cell analyses of cases SA494 (an example of extreme initial selection) and SA501 (complex post-engraftment clonal dynamics). We performed multiplexed targeted re-sequencing of SNVs in 210 isolated tumour and xenograft nuclei, using microfluidic devices. We determined evolutionary relationships between nuclei by Bayesian phylogenetic inference[20], deriving consensus genotypes for clades representing high probability splits in the phylogenetic tree.

As predicted by PyClone, two major clades emerge in the SA494 phylogeny, comprising tumour and xenograft nuclei respectively, bearing mutually-exclusive sets of alleles in additional to a set of shared alleles (Extended Figure E3a, b, c, Figure S13). The ancestral clone SNVs (PyClone cluster 1) are common to nuclei from both clades, while SNVs in the predicted dominant tumour clone (cluster 2) and minor engrafting clone (cluster 3) are restricted to tumour and xenograft nuclei respectively (Extended Figure E3d, Genotypes A and B). This confirms the ancestral relationship between tumour and xenograft, verifies the expansion of a very minor clone (<5%) while showing unambiguously that mutation clusters inferred by PyClone represent major clonal genotypes.

PyClone analysis of SA501 (Figure 2 and Figure S12) revealed a dynamic and complex clonal architecture, with gradual expansion of minor mutation clusters observed over

consecutive passages, and expansion followed by decline of other clusters (Figure 2c). The major mutation clusters and their gradual change in prevalence over time predicted by PyClone were confirmed by the clonal genotypes of single cells from SA501 passages X1, X2 and X4 (Figure 2b, Figure S13). Phylogenetic inference resolved the clonal genotypes of five major clades (Figure 2a, d), with cascading acquisition of mutations from parental to descendent clone (Figure 2e). Genotypes A and B belong to sibling clades defined by the addition of cluster 5 and cluster 4 mutations, respectively, to the ancestral genotype defined by clusters 1 and 8; Genotype C was derived from Genotype B with the addition of mutations in cluster 7; Genotype D derived from Genotype C with the addition of mutations defined by cluster 2; and Genotype E from Genotype D with the addition of cluster 3 mutations and loss of cluster 8 mutations (Figure 2a, d, e). The clonal dynamics measured in the population was reflected in the relative abundance of single cell genotypes in each xenograft tumour (Figure 2f), mirroring bulk population predictions (Figure 2c). Both X1 and X2 sampled nuclei show an admixture of clones defined by genotypes A, B, C, and D (relatively rare in X1). Genotype E is confined exclusively to X4 nuclei, suggesting that by passage 4, this clone had nearly exhaustively outcompeted its ancestor and sibling clones. Its eventual dominance is mirrored by the decline of Genotype A (initially present in X1 and X2), suggesting the descendants of Genotype B outcompeted those of Genotype A over time.

Taken together, these single cell genotyping experiments combined with phylogenetic inference have recapitulated population level PyClone predictions in a simple (SA494) and a complex (SA501) clonal expansion model. Thus, single cell genotyping validates PyClone mutation clusters as genomic markers of major clonal genotypes, while providing additional insight into the ancestral lineages of cell populations.

Finally, to determine whether directional clonal dynamics might be associated with deterministic as opposed to stochastic processes (such as random genetic drift), we tested whether similar clonal dynamics occurred when the same tumour population was multiply transplanted into different mice. In 4/5 series examined, parallel clonal dynamics of the same mutation cluster(s) were observed (arrows in Figure 3a, b and Extended Figure E4a, b: SA501 2/2 replicate mice at passage X3 and 4/4 at X4; SA535 X1 3/3; SA532 X1 3/3, X2 3/7 and X3 2/2; SA429 X2 3/5). These include reproducible expansions of initially minor subclones, implying a high likelihood of shared deterministic mechanism rather than repeated rare stochastic events (for example, arising from transplants close to limiting dilution). In SA501 the same pattern (expansion of cluster 3 mutations mirrored by decline of cluster 5 mutations) was independently observed in transplants at passage 2, 3 and 4 (Figure 3a, 2B, 3B, 4A–D), suggesting shared clonal fitness but variable timing. We also observed instances of divergence, for example expansion of SA532 cluster 4 specific to branch 2A-3A-4A-5A (Extended Figure E4a). SA535 (Figure 3b) and SA532 showed examples of clonal expansion patterns replicated in related but different immunodeficient mouse strains (NSG, NRG). To control against shared clonal structure imposed through joint inference of the datasets, we also carried out independent PyClone analyses that excluded all but one transplant at each passage, and observed high correlations of inferred mutation prevalences between same-passage replicates (Extended Figure E5; median Pearson correlations 0.94, 0.93, 0.91, 0.91, 0.46 for SA501, SA535, SA532, SA429, SA496). These

data indicate that clonal genotypes defined by somatic aberrations (and/or closely co-segregating genomic factors) can be biologically meaningful determinants of fitness, leading to consistent and reproducible clonal dynamics.

We show here that patient derived xenograft clonal dynamics on initial transplant vary from polyclonal engraftment with only moderate clonal selection, in which tumour and xenograft clonal prevalence are broadly similar (a minority of cases), to highly skewed dynamics in which initially minor prevalence clones expand to dominate the xenograft (the majority of cases). Expansion of minor subclones has been suggested in previous xenotransplantation studies using malignant epithelial[10, 21–23] or hematopoietic[24, 25] cells, without formal resolution of the clonal genotypes or pattern of subsequent clonal dynamics. In contrast with preliminary studies of xenoengraftment, we find correlated dynamics of clones defined by SNVs or copy number aberrations as clonal marks. Expansion patterns are most often pronounced in the initial establishment passage, however in cases where initial clonal selection is weak, subsequent evolution over passaging is more evident. Furthermore, polyclonal sub-structure may emerge even in xenografts that have undergone a modest population bottleneck on initial engraftment. These dynamic processes are not evident from histopathological or imaging characteristics, which remain broadly stable, consistent with previous reports[8, 9, 23].

Importantly, we find that the population dynamics of genomically-defined clones are replicated when transplants are carried out in multiple mice, implying that the basis of selection is non-random and likely closely linked to the particular mutation genotype (or epigenotype) that defines the clone. The most parsimonious explanation for repeated observation of these clonal dynamics is that the clones are mostly pre-existing and variations in clonal fitness explain the dynamic behaviour, as opposed to de-novo somatic mutation. Furthermore, cases in which conversion from minor to dominant clone occurs monotonically over multiple passages demonstrate that selective fitness can be persistent rather than transient. Thus, specific somatic genotypes are likely to act as genetic markers of clonal growth and fitness advantages, yielding predictable and reproducible clonal dynamics. Determination of the precise aberrations that give rise to selective clonal fitness still faces considerable challenges. In this regard, we believe that ascertainment of clonal dynamics will prove essential for fully informed future studies of drug response and tumour biology in xenografts of human breast cancers.

Extended Data

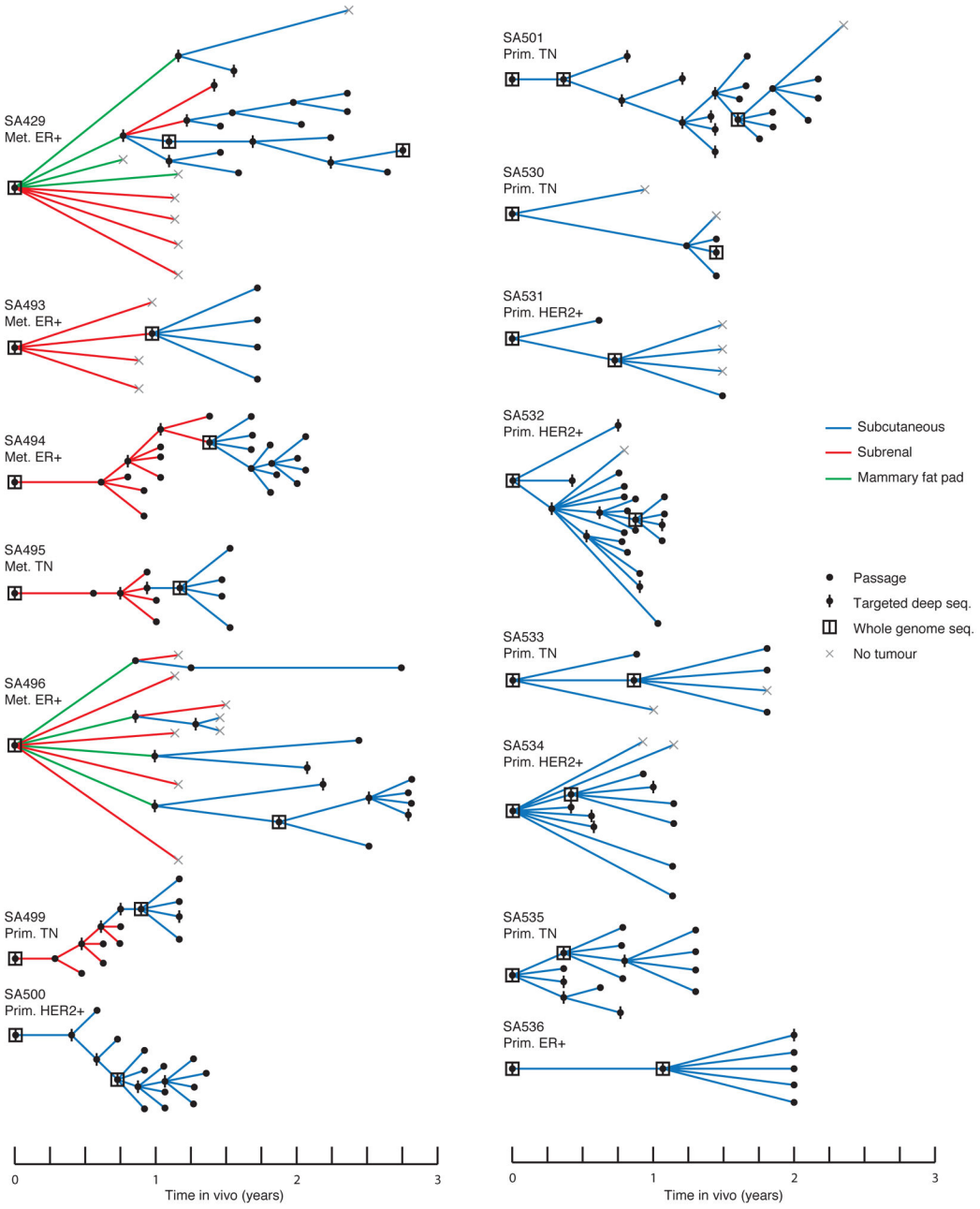

Figure E1.

Transplant History. Diagrams show the transplant history of each xenograft line. Line segment colours represent the site used for each transplant (blue=subcutaneous, red=subrenal capsule, green=mammary fat pad). Black points indicate the passage of an engrafted xenograft to the next mouse generation. Grey crosses indicate transplants that did not result in palpable tumours. Samples analyzed by whole genome and/or targeted deep sequencing are indicated (black squares and vertical lines, respectively). The cumulative time in vivo is shown on the x-axis. The originating tumour site (Prim.=primary breast,

Met.=pleural effusion) and immunohistochemical expression of biomarkers (ER=estrogen receptor, PR=progesterone receptor, TN=triple negative for ER, PR and HER2) are shown.

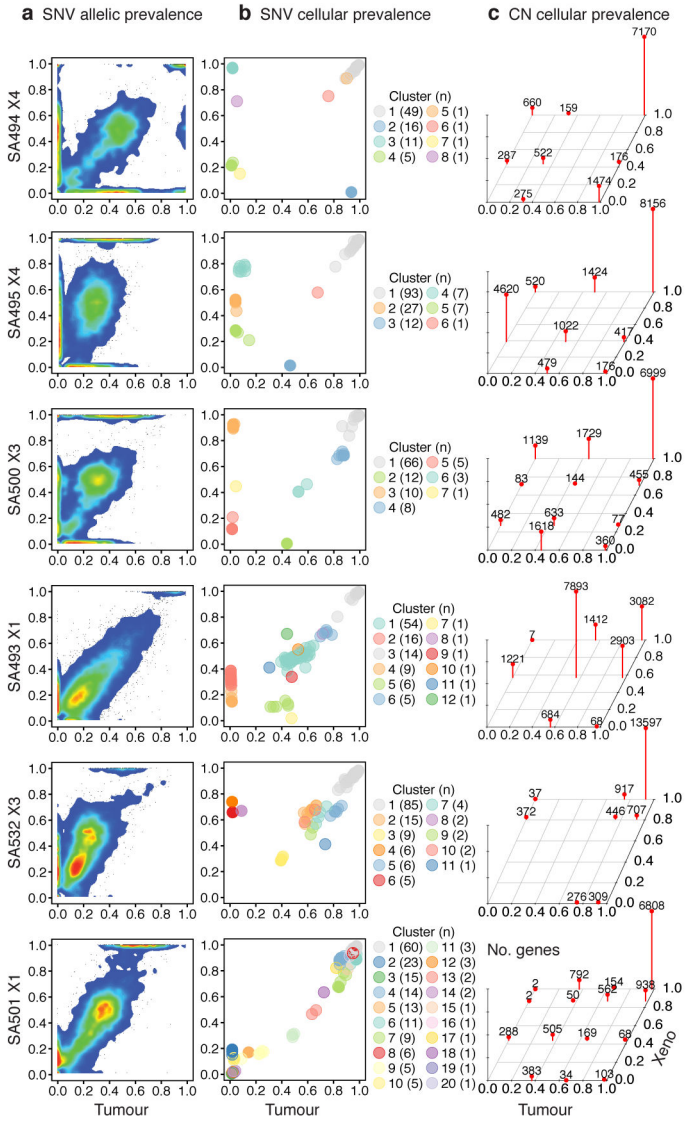

Figure E2.

Comparison of the prevalence of mutations in six originating tumors and subsequent xenografts in SNV and CNA spaces. **a**, Density scatter plots showing the WGSS variant allele prevalence of genome-wide high-confidence SNVs in tumours (x-axis) and xenografts (y-axis). SNVs in clones undergoing neutral dynamics lie along a diagonal, and SNVs in clones undergoing expansion or contraction lie on/towards the y- and x-axes respectively. **b**, Scatter plots showing the mutation cellular prevalence of selected SNVs in tumours and xenografts, inferred by PyClone from population targeted deep sequencing. Circles represent individual SNVs, colours indicate clusters of mutations for which mutation cellular prevalences vary together over all sample time points. **c**, Scatter plots show co-occurrence of CNA/LOH events inferred by TITAN in tumours and xenografts. The z-axis height of each

bar shows the number of genes belonging to a unique mutation cluster and present at the indicated mutation cellular prevalence in tumour (x-axis) and xenograft (y-axis).

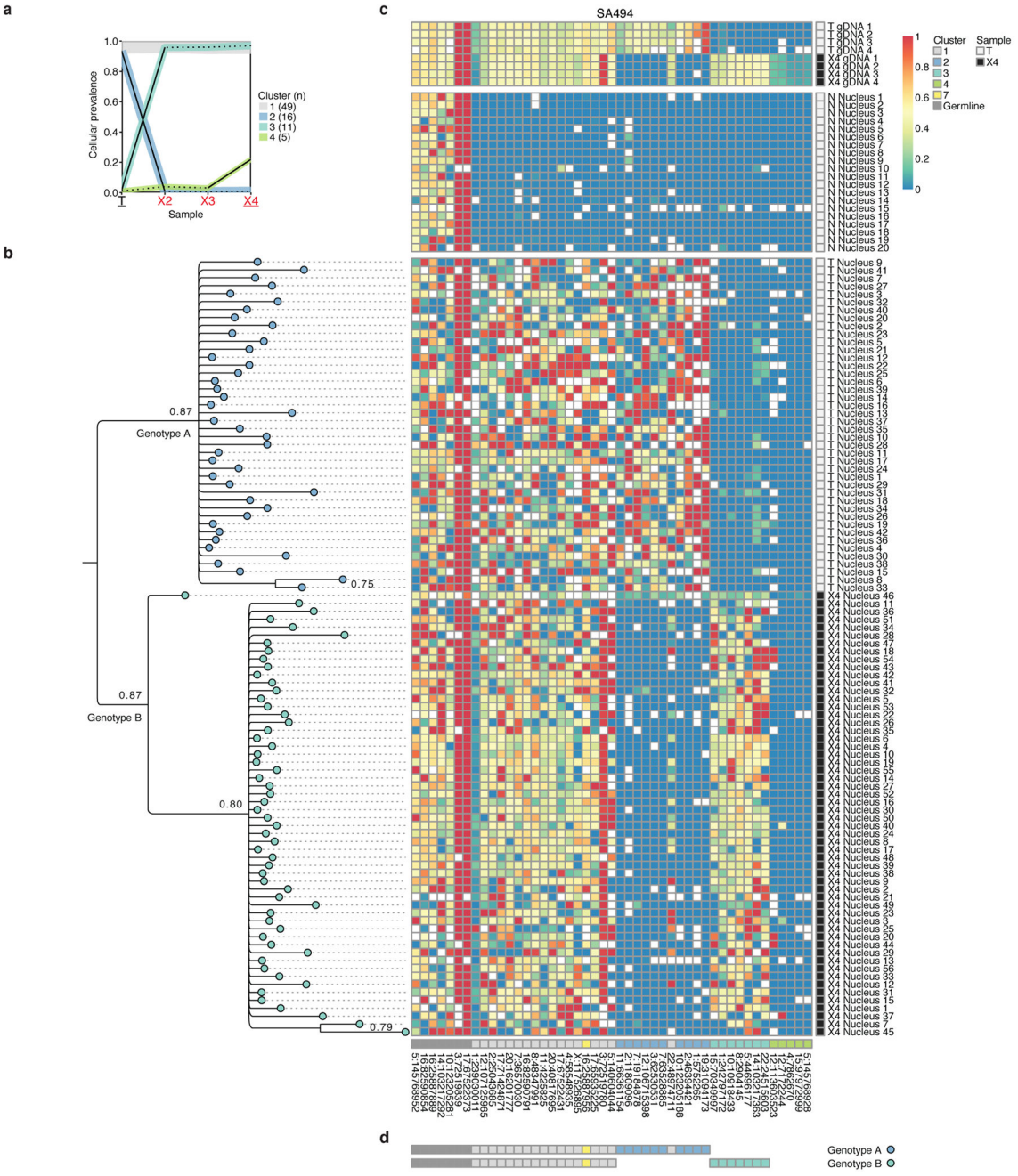

Figure E3.

Single cell determination of clonal genotypes recapitulates population-based prediction of minor clone selection. DNA prepared from 62 individual SA494 tumor and 58 passage 4 xenograft nuclei was amplified in single reactions using a panel of multiplexed PCR primer pairs targeting amplicons containing 40 SNV and 7 germline variants, and the variant allele ratios were determined by targeted deep sequencing. **a** Mutation clusters inferred by the

PyClone model from bulk population measurements. **b**, Bayesian phylogenetic tree derived from multi-locus genotypes of individual nuclei. The tumour and xenograft nuclei group in distinct clades. **c**, Heatmap depicts the multi-locus variant allele prevalences (blue/yellow/red corresponds to wild-type/heterozygous/homozygous loci) at variant positions (horizontal axis) in individual nuclei (vertical axis, ordered by phylogenetic grouping in (**b**)). Upper two blocks show genomic DNA controls and normal cell nuclei present in tumour sample. The PyClone mutation cluster corresponding to each SNV is indicated by colour in the lowermost horizontal bar. **d**, Consensus genotypes derived from high-probability splits in the phylogenetic tree confirm a set of high prevalence tumour-specific and xenograft-specific mutations, consistent with the expansion of a minor originating clone to dominance in the xenograft, as well as mutations shared in tumour and xenograft nuclei.

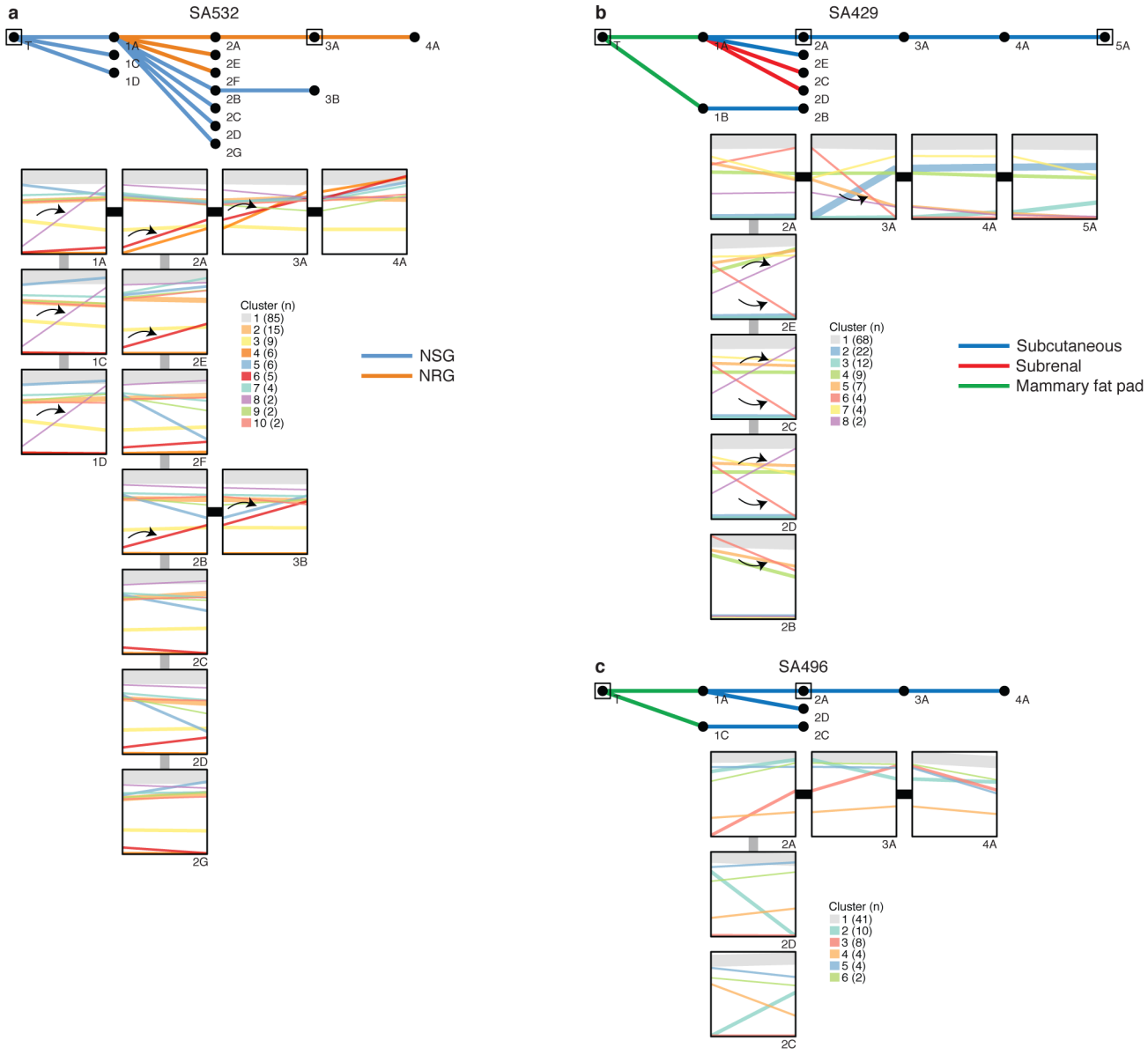

Figure E4.

Clonal dynamics are reproduced in replicate transplants (2). **a**, **b**, **c**, Upper panels: passaging history of SA532, SA429, SA496, showing transplants that resulted in successful xenografts. The transplants sites (blue=SC, red=SR, green=MFP; all SC for SA532) and host mouse strains (blue=NSG, orange=NRG; all NSG for SA429 and SA496). Lower panels: change in cellular prevalence of mutation clusters over individual transplants. Plots correspond to passages in upper panels. The clusters are inferred by PyClone using grouped data from all passages, and correspond to those displayed in Figure 1. Arrows in SA429 and SA532 show examples of parallel clonal dynamics of the same mutation cluster in multiple replicate transplants. SA496 exhibits less replicated evolution compared with other cases.

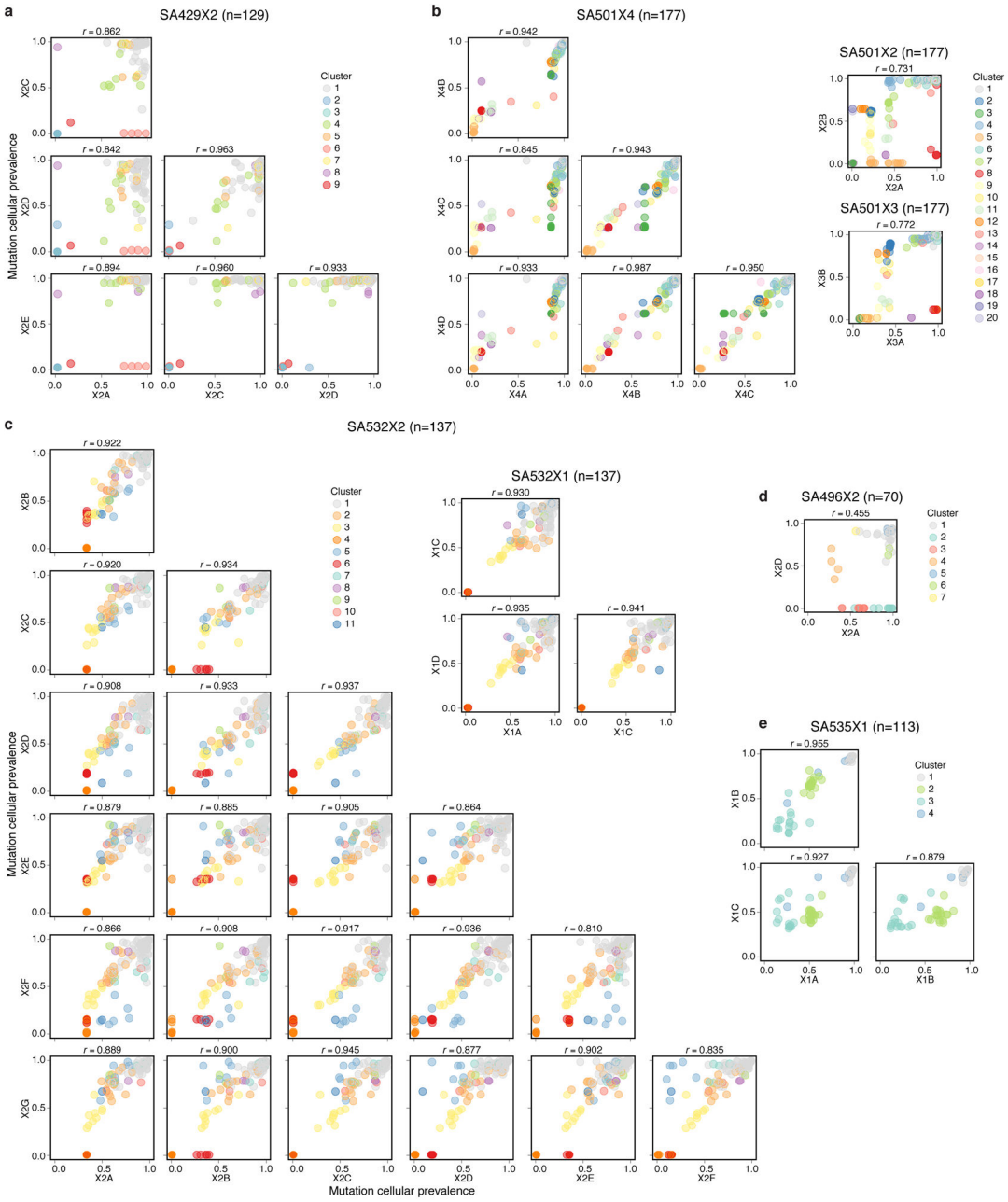

Figure E5.

Correlation of clonal dynamics in replicate transplants of SA429, SA501, SA532, SA496 and SA535. **a**, **b**, **c**, **d**, **e**, Scatter plots display the inferred mutation cellular prevalence of all SNVs in pairs of same-passage replicates, for cases SA429, SA501, SA532, SA496 and SA535 respectively. For each replicate, prevalences are inferred by a separate PyClone analysis that excludes data from other same-passage transplants. Colours indicate mutation clusters inferred in each individual PyClone analyses; the SNVs clustered and colours assigned may differ in each plot. The Pearson correlation coefficients are shown, indicating closely related evolution in most pairs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the staff of the CTAG Molecular Pathology facility, members of the Library Technical Development, Library Construction, Sequencing and Bioinformatics teams at the Michael Smith Genome Sciences Centre for technical assistance with data generation, and Steve Kalloger for assisstance with sample collection. SA and SS are supported by Canada Research Chairs. PE is supported by a Michael Smith Foundation for Health Research (MSFHR) Fellowship. AS is supported by an NSERC CREATE scholarship through the graduate program in Genome Science and Technology at UBC. SS is a MSFHR scholar. We acknowledge generous long term funding support provided by the BC Cancer Foundation. The SA, SS and CH groups receive operating funds from the Canadian Breast Cancer Foundation, Canadian Cancer Society Research Institute, Terry Fox Research Institute, Genome Canada and Canadian Institutes for Health Research (CIHR). We thank Dr Sarah Mullaly for critical reading of the manuscript.

References

1. Aparicio S, Caldas C. The implications of clonal genome evolution for cancer medicine. N Engl J Med. 2013; 368:842–51. [PubMed: 23445095]
2. Nowell PC. The clonal evolution of tumor cell populations. Science (New York, NY). 1976; 194:23–28.
3. Diaz LA Jr, et al. Nature. 2012; 486:537–40. [PubMed: 22722843]
4. Shah SP, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012; 486:395–9. [PubMed: 22495314]
5. Gerlinger M, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine. 2012; 366:883–892. [PubMed: 22397650]
6. Campbell PJ, et al. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature. 2010; 467:1109–1113. [PubMed: 20981101]
7. Bashashati A, et al. Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. The Journal of pathology. 2013; 231:21–34. [PubMed: 23780408]
8. DeRose YS, et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nat Med. 2011; 17:1514–20. [PubMed: 22019887]
9. Zhang X, et al. A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. Cancer Res. 2013; 73:4885–97. [PubMed: 23737486]
10. Ding L, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010; 464:999–1005. [PubMed: 20393555]
11. Pearson T, et al. Non-obese diabetic-recombination activating gene-1 (NOD-Rag1 null) interleukin (IL)-2 receptor common gamma chain (IL2r gamma null) null mice: a radioresistant model for human lymphohaematopoietic engraftment. Clin Exp Immunol. 2008; 154:270–84. [PubMed: 18785974]
12. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490:61–70. [PubMed: 23000897]
13. Ha G, et al. Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Res. 2012; 22:1995–2007. [PubMed: 22637570]
14. Nik-Zainal S, et al. The life history of 21 breast cancers. Cell. 2012; 149:994–1007. [PubMed: 22608083]
15. Ellis MJ, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature. 2012; 486:353–360. [PubMed: 22722193]

16. Banerji S, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature. 2012; 486:405–409. [PubMed: 22722202]
17. Curtis C, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012; 486:346–52. [PubMed: 22522925]
18. Roth A, et al. PyClone: statistical inference of clonal population structure in cancer. Nat Methods. 2014
19. Ha G, et al. Titan: Inference of copy number architectures in clonal cell populations from tumor whole genome sequence data. Genome Res. 2014
20. Ronquist F, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012; 61:539–42. [PubMed: 22357727]
21. Kreso A, et al. Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. Science. 2013; 339:543–8. [PubMed: 23239622]
22. Nolan-Stevaux O, et al. Measurement of cancer cell growth heterogeneity through lentiviral barcoding identifies clonal dominance as a characteristic of in vivo tumor engraftment. PLoS One. 2013; 8:e67316. [PubMed: 23840661]
23. Li S, et al. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. Cell Rep. 2013; 4:1116–1130. [PubMed: 24055055]
24. Notta F, et al. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. Nature. 2011; 469:362–7. [PubMed: 21248843]
25. Clappier E, et al. Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse. J Exp Med. 2011; 208:653–61. [PubMed: 21464223]

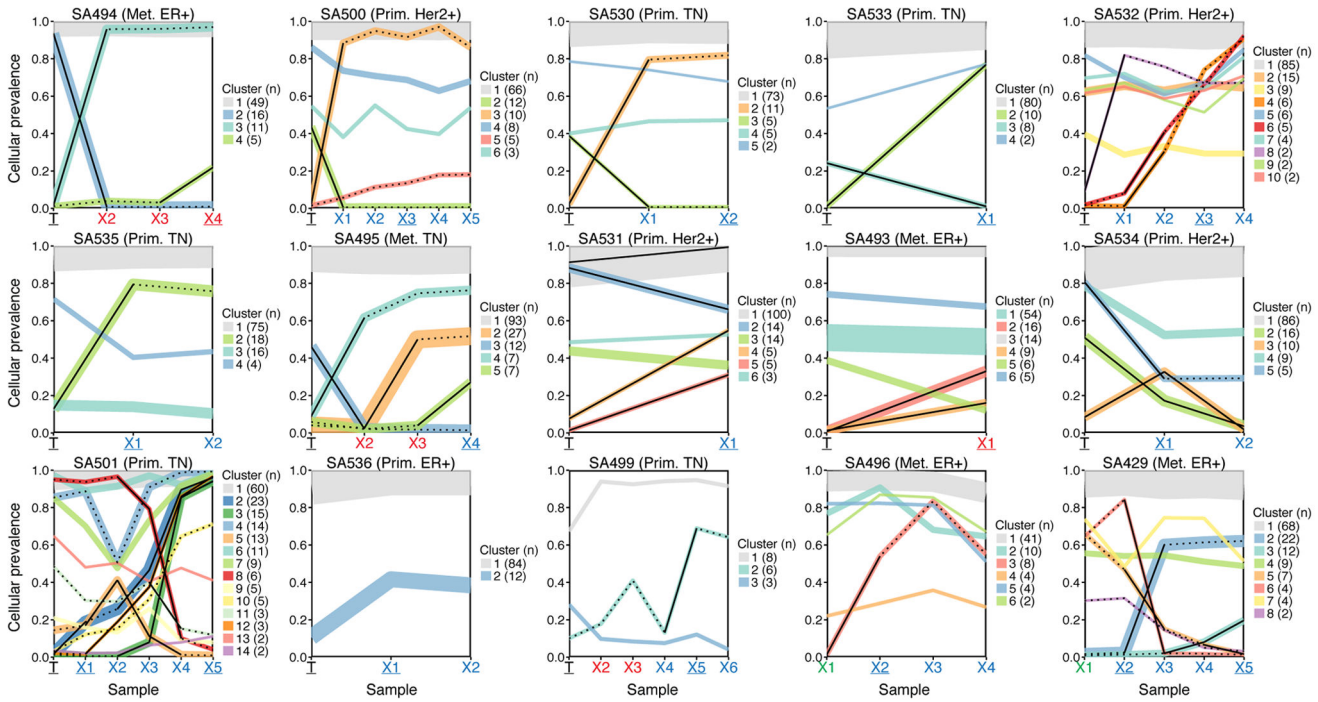

Figure 1.

Clonal dynamics over multiple passages in time. Plots display the mean cellular prevalence estimates of mutation clusters in originating tumours (T) and subsequent xenograft passages (X1, X2, etc.). The clusters and prevalences were inferred by PyClone from population targeted deep sequencing. Line widths indicate the number of SNVs comprising each mutation cluster (numbers in brackets adjacent to each plot). Black lines indicate non-neutral dynamics, assessed by non-overlap of credible intervals derived from Bayesian posterior distributions (solid=non-neutral over indicated passage, dotted=over cumulative passages since initial transplant). All passages that underwent deep sequencing are shown. Transplant sites are represented by colour (blue=subcutaneous, red=subrenal, green=mammary fat pad), tumour and passages analyzed by WGSS are underlined. The panels are ordered by the degree of initial change in mutation cellular prevalence. Singleton clusters were not displayed for clarity.

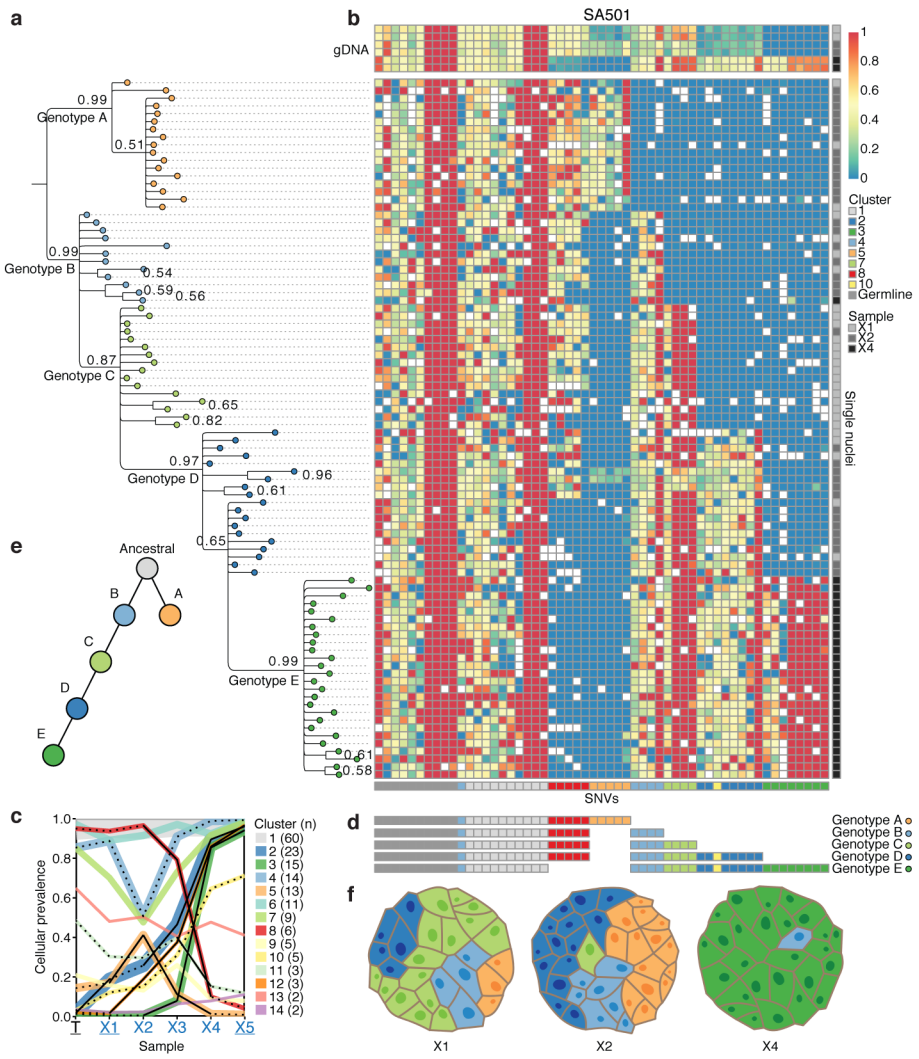

Figure 2.

Single cell determination of clonal genotypes recapitulates population-based prediction of cascading subclonal evolution. DNA was prepared from 90 individual SA501 xenograft nuclei from passages X1, X2 and X4, and the variant allele ratios were determined by targeted ultra-deep sequencing at 45 somatic SNV and 10 germline SNV positions. **a**, Bayesian phylogenetic tree derived from multi-locus genotypes of individual nuclei, depicting cascading evolution. **b**, Heatmap depicting multilocus variant allele ratios (blue/yellow/red corresponds to wild-type/heterozygous/homozygous loci). Nuclei (y-axis) are ordered according to the phylogenetic tree in (**a**). Positions (x-axis) are grouped according to the consensus genotypes derived from high-probability branch splits in a manner naive to the PyClone clustering. The cluster groupings (horizontal bar below horizontal axis) recapitulate the PyClone groupings inferred from bulk population measurements (**c**). **d**, Five consensus genotypes derived from high-probability splits in the phylogenetic tree. **e**, Schematic of the phylogeny derived from single cell genotyping depicts the sequential expansion of genomic subclones. Genotypes are coloured according to the last Py-Clone mutation cluster acquired at a given point in the phylogeny. **f**, Schematic representations of xenograft tumours X1, X2,

and X4 based on single cell genotypes. Cells are coloured according to their genotype in (**e**), and the number of cells within each schematic corresponds to the number of sequenced nuclei with the given genotype in (**b**). The relative proportions of cells with each genotype reflect predictions based on bulk measurements in (**c**).

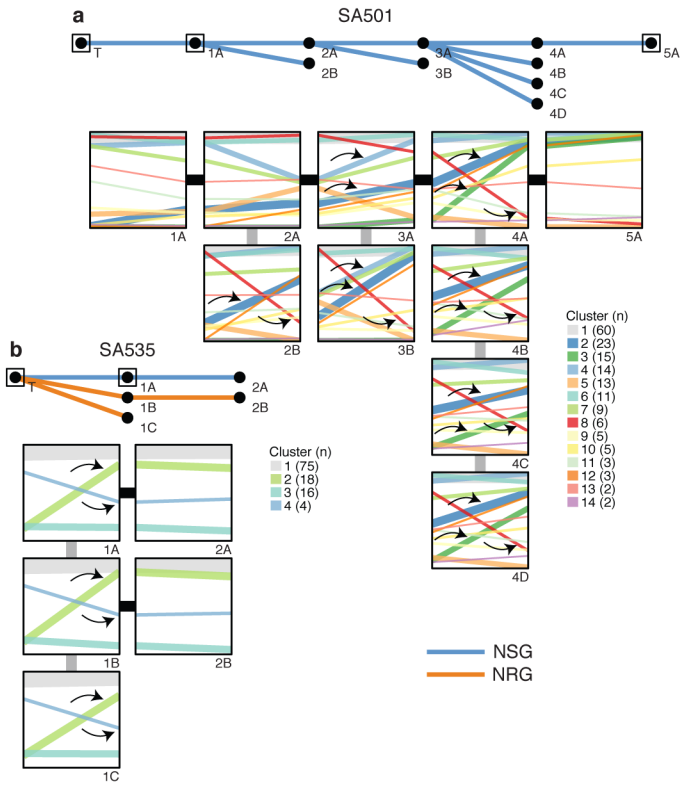

Figure 3.

Clonal dynamics are reproduced in replicate transplants (1). **a**, **b**, Upper panels: Passaging history of SA501, SA535 showing transplants that resulted in successful xenografts. The host mouse strains (blue=NSG, orange=NRG) are indicated. All transplants were in subcutaneous site. Lower panels: change in cellular prevalence of mutation clusters over individual transplants. Plots correspond to passages in upper panels. The clusters are inferred by PyClone using grouped data from all passages, and correspond to those displayed in Figure 1. Arrows show examples of parallel clonal dynamics of the same mutation cluster in multiple replicate transplants.