Breast Cancer Research and Treatment (2019) 175:91–104
https://doi.org/10.1007/s10549-019-05153-8

PRECLINICAL STUDY

Revealing clonality and subclonality of driver genes for clinical survival benefits in breast cancer

Yujia Lan[1] · Erjie Zhao[1] · Shangyi Luo[1] · Yun Xiao[1] · Xia Li[1] · Shujun Cheng[1,2]

Received: 11 October 2018 / Accepted: 28 January 2019 / Published online: 9 February 2019

Abstract

Purpose Genomic studies have revealed that genomic aberrations play important roles in the progression of this disease. The aim of this study was to evaluate the associations between clinical survival outcomes of the clonality and subclonality status of driver genes in breast cancer.

Methods We performed an integrated analysis to infer the clonal status of 55 driver genes in breast cancer data from TCGA. We used the chi-squared test to assess the relations between clonality of driver gene mutations and clinicopathological factors. The Kaplan–Meier method was performed for the visualization and the differences between survival curves were calculated by log-rank test. Univariate and multivariate Cox proportional hazards regression models were used to adjust for clinicopathological factors.

Results We identified a high proportion of clonal mutations in these driver genes. Among them, there were 17 genes showing significant associations between their clonality and multiple clinicopathologic factors. Performing survival analysis on BRCA patients with clonal or subclonal driver gene mutations, we found that clonal *ERBB2, FOXA1*, and *KMT2C* mutations and subclonal *GATA3* and *RB1* mutations predicted shorter overall survival compared with those with wild type. Furthermore, clonal *ERBB2* and *FOXA1* mutations and subclonal *GATA3* and *RB1* mutations independently predicted for shorter overall survival after adjusting for clinicopathological factors. By longitudinal analysis, the clonality of *ERBB2, FOXA1, GATA3*, and *RB1* significantly predicted patients' outcome within some specific BRCA tumor stages and histological subtypes.

Conclusions In summary, these clonal or subclonal mutations of driver genes have implications for diagnosis, prognosis, and treatment with BRCA patients.

Keywords Breast cancer · Clinical survival · Driver gene · Clonal mutation · Subclonal mutation

Introduction

Breast cancer is one of the most frequent cancers among women and a heterogeneous disease with multiple molecular alterations [1, 2]. These alterations affect external behaviors of cancers pertaining to invasion, metastasis, and the resistance to drug treatment. With the development of the next generation sequencing, a large number of studies have revealed that many gene mutations present in breast cancer cells [3, 4]. Moreover, driver gene mutations promote cancer progression and have major impacts on patient outcome, such as the most commonly altered genes *BRCA1* and *BRCA2* [5–7]. Therefore, identifying genomic aberrations of cancers may be critical to understand the disease progression and find new therapeutic strategies.

Accumulating evidence has shown that cancers are composed of distinct tumor cells with clonal or subclonal

Yujia Lan and Erjie Zhao are Contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10549-019-05153-8) contains supplementary material, which is available to authorized users.

✉ Yun Xiao
xiaoyun@ems.hrbmu.edu.cn

✉ Xia Li
lixia@hrbmu.edu.cn

✉ Shujun Cheng
chengshj@263.net.cn

[1] College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150086, Heilongjiang, China

[2] State Key Laboratory of Molecular Oncology, Department of Etiology and Carcinogenesis, Cancer Institute and Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing 100021, China

mutations that may influence the course of disease evolution [8, 9]. Furthermore, the subclonal driver mutations may affect the more aggressive evolution of diseases [10, 11]. In the last few years, several studies have characterized the clonal heterogeneity of tumors by multi-regional sequencing and/or multi-time point sequencing [12, 13]. However, the clonal and subclonal fraction of driver gene mutations in breast cancers remains unknown. And it also remains unclear whether there are clinical and biological differences among patients with clonal mutations, subclonal mutations, and wild type.

In this study, we investigated the mutation clonality of 55 driver genes and their clinical associations in breast cancers using the genomic data of BRCA from The Cancer Genome Atlas (TCGA). The clonal *ERBB2* and *FOXA1* mutations and subclonal *GATA3* and *RB1* mutations independently predicted for shorter overall survival (OS) after adjusting for other clinicopathological factors, such as TNM stage and American Joint Committee on Cancer (AJCC) stage. Importantly, the mutation clonality of *ERBB2, FOXA1, GATA3*, and *RB1* significantly predicted overall survival of patients with some specific tumor stages and histological subtypes. We concluded that mutation clonality of driver genes was effectively prognostic and predictive biomarkers of breast cancer patients, which could be used in clinical applications to choose effective therapies.

Method

Data source

In this study, the Affymetrix SNP6.0 array data and corresponding clinical metadata [including clinicopathological factors, overall survival, and disease-free survival (DFS)] of breast cancer patients were accessed through The Cancer Genome Atlas portal (https://portal.gdc.cancer.gov) or Broad Institute Firehose (https://gdac.broadinstitute.org). The detail mutation annotation file of every sample including variant allele frequency of each mutation was obtained from the public cBio Cancer Genomics Portal (http://www.cbioportal.org) [14, 15]. The data were filtered to exclude patients without SNP6.0 array data, mutation data, or clinical information. There were 993 patients having mutation data, 1032 having copy number data, and 1083 having available survival data. Finally, there were 979 patients presenting in all data. Among these breast cancer samples, average age of patients was 58 years old (range from 26 to 90). And there were 499 Luminal A, 197 Luminal B, 78 HER2+, and 171 Basal-like breast cancer samples. In addition, other clinicopathological characteristics were summarized in Table 1.

Table 1 Clinical and pathological characteristics of patients in TCGA breast cancer cohort

Clinical features	Category	BRCA, $n = 1083$
Gender	Female	1071
	Male	12
Age	Median (range)	58 (26–90)
Classical subtype	Basal	171
	HER2	78
	LumA	499
	LumB	197
AJCC tumor stage	I	180
	II	614
	III	246
	IV	19
	Not available	24
T stage	T1	276
	T2	628
	T3	137
	T4	39
	TX	3
N stage	N0	511
	N1	357
	N2	119
	N3	76
	NX	20
M stage	M0	900
	M1	21
	MX	162
ER status	Positive	795
	Negative	237
	Not available	51
PR status	Positive	687
	Indeterminate	342
	Negative	4
	Not available	50
Follow-up from sampling(days)	Median (range)	456.5 (1.0–7067.0)

Inferring the cancer cell fraction of point mutations

Integrating tumor purity and the local copy number of mutation sites as summarized by McGranahan et al. [16] and Landau et al. [10], we estimated the cancer cell fraction (CCF) of each mutation, and then inferred the clonal status of all gene mutations in every sample [17]. First, we calculated the tumor purity and absolute DNA copy number for each breast cancer sample using the ABSOLUTE algorithm [18]. For each mutation site, the expected variant allele frequency (VAF) was calculated using the following formula:

$$\mathrm{VAF}_{\mathrm{exp}} = \frac{\mathrm{tp} * \mathrm{CCF} * \mathrm{CN}_{\mathrm{muts}}}{\mathrm{CPN}_{\mathrm{norm}}(1 - \mathrm{tp}) + \mathrm{tp} * \mathrm{CPN}_{\mathrm{tumor}}}$$

where tp represents tumor purity, CN_{muts} denotes copy number of the allele carrying the mutation, CPN_{tumor} and CPN_{norm} denotes the local copy numbers of the mutation site in cancer cells and normal cells, respectively. Generally, let CPN_{norm} to be 2 ($CPN_{norm} = 1$ only when considering mutations in the X chromosome for males). The somatic variants present at a single copy per cancer cell. Hence, we assumed CN_{muts} to be 1 and the expected VAF can also be approximated as follows:

$$VAF_{exp} = \frac{tp * CCF}{2(1 - tp) + tp * CPN_{tumor}}$$

For a point mutation with 'c' alternate reads and 'T' sequencing coverage, we computed the probability of a given CCF using Bayesian probability theory and a binomial distribution:

$$P(CCF|(c|T)) = \frac{P((c|T)|CCF) * P(CCF)}{P(c|T)}$$

and

$$P(CCF|(c|T)) \propto Binom(c|T, VAF_{exp}(CCF))$$

Then, a distribution of CCF was obtained by calculating the P(CCF) over a uniform grid of 100 CCF values from 0.01 to 1 and normalizing by dividing their sum. Let $P(CCF = z1) > P(CCF = z2) > \ldots P(CCF = zk) > P(CCF = zk + 1) \ldots > P(CCF = z100)$. For each mutation, the estimated CCF and 95% confidence interval (CI95) were calculated as follows:

$$argmax_i(P(CCF=i))$$

$$CI95 = [\min(z_1, z_2 \ldots z_k), \max(z_1, z_2 \ldots z_k)] \Rightarrow \arg\max_k \sum_{x=z_1}^{z_k} P(CCF = x) < 0.95$$

To determine the clonal status of each mutation, we defined the Pr(clonal) to denote the probability that a mutation is clonal:

$$Pr(clonal) = P(CCF \geq 0.9) = \sum_{i=0.9}^{1} P(CCF = i)$$

Finally, we regarded mutations as clonal if the upper band of the CI95 was ≥ 1 and the $Pr(clonal) > 0.5$; otherwise, the mutations were considered subclonal. Such classification could ensure that the number of subclonal mutations was not overestimated.

Statistical analysis

To assess the relations between the clonal status of driver gene mutations and clinicopathological factors, we used the chi-squared test. For survival analysis, overall survival and disease-free survival were used as the end points. The Kaplan–Meier method was performed for the visualization purposes and the differences between survival curves were calculated by log-rank test. Univariate and multivariate Cox proportional hazards regression models were applied to estimate the prognostic capability of clonal or subclonal driver gene mutations. The *P*-values smaller than 0.05 were considered to be statistically significant. All of the statistical analyses were performed using *R* software (http://www.r-project.org), version 3.4.4.

Results

Inference of clonality of driver genes in breast cancer samples

To systematically investigate the clonal architecture of breast cancer, we classified somatic mutations into clonal and subclonal categories by integrating single-nucleotide polymorphism arrays (SNP) and mutation data (Methods). Among 979 TCGA breast cancer samples, we successfully derived the tumor purity and local absolute copy numbers and then predicted the clonal status of each gene mutation. As a result, we identified a total of 86,248 clonal mutations (72.6% of all detected mutations, the average of 80.1 mutations per sample) and 32,479 subclonal mutations (average of 33.2 mutations per sample). Further, we used SciClone [19], an approach for clustering variants in a single sample, to assess the clonal status of mutations identified by our analysis. We found that 640 samples were detected with at least one mutation cluster. For samples with more than one cluster ($n = 449$), the mutations in each cluster of 363 samples (80.85%) were tended to be clonal or subclonal (Fisher's test, $P < 0.05$). This result showed consistency between our method and Sciclone.

Furthermore, we identified 55 breast cancer driver genes from the study by Lawrence et al. that integrated three MutSig tests (MutSigCV, MutSigCL, and MutSigFN) to identify significantly mutated driver genes (FDR ≤ 0.1) across 21 tumor types [20], and from the Wellcome Sanger Institute Cancer Gene Census that manually collected more than 700 cancer genes that were mutated and causally implicated in cancer development from literatures [21, 22]. We analyzed clonal status of these known breast cancer driver genes. Then, we found that there were seven genes (*MLL, CUL48, STAG2, RAB40A, FLNA, IRS4*, and *ZMYM3*) without clonal and subclonal mutations. We also observed there were three genes (*AKT1, CCND1*, and *PPM1D*) only containing clonal mutations. The number of clonal and subclonal mutations in the remaining 45 driver genes was shown in Fig. 1. There were 16 genes (*CBFB, MED23, MAP3K13, SALL4, ESR1, ACVR1B, BAP1, CDKN1B, GNPTAB, TBL1XR1, NOTCH1, FBLN2,*

KRAS, ETV6, SMARCD1, and *KEAP1*) with fewer ($n \leq 4$) subclonal mutations in samples. These genes harboring subclonal mutations ≤ 4 were excluded in the subsequent analysis. Among the remaining 29 driver genes, the top ten genes with the highest clonal mutation included *TP53, PIK3CA, CDH1, GATA3, MAP3L1, KMT2C, PTEN, NCOR1, RUNX1*, and *MAP2K4*. And the top ten genes with the most subclonal mutations were *KMT2C, MAP3K1, GATA3, PIK3CA, CDH1, TP53,NCOR1, PTEN, RUNX1*, and *ARID1A*. We found that mutations in these genes showed a tendency to be clonal. For instance, the *TP53* was to be clonal mutations in 321 samples, while to be subclonal mutations in 15 samples. In addition, we annotated our data using known mutations from oncoKB [23]. Among these mutations, we found 335 known mutations (including 310 clonal and 25 subclonal mutations) and 2,006 unknown mutations (including 1,575 clonal and 431 subclonal mutations). These unknown mutations tended to be subclonal mutations (21.5%, compared with 19.5% expected by chance, binomial test, $P = 0.024$), different from the known mutations (7.5%). These findings suggested that

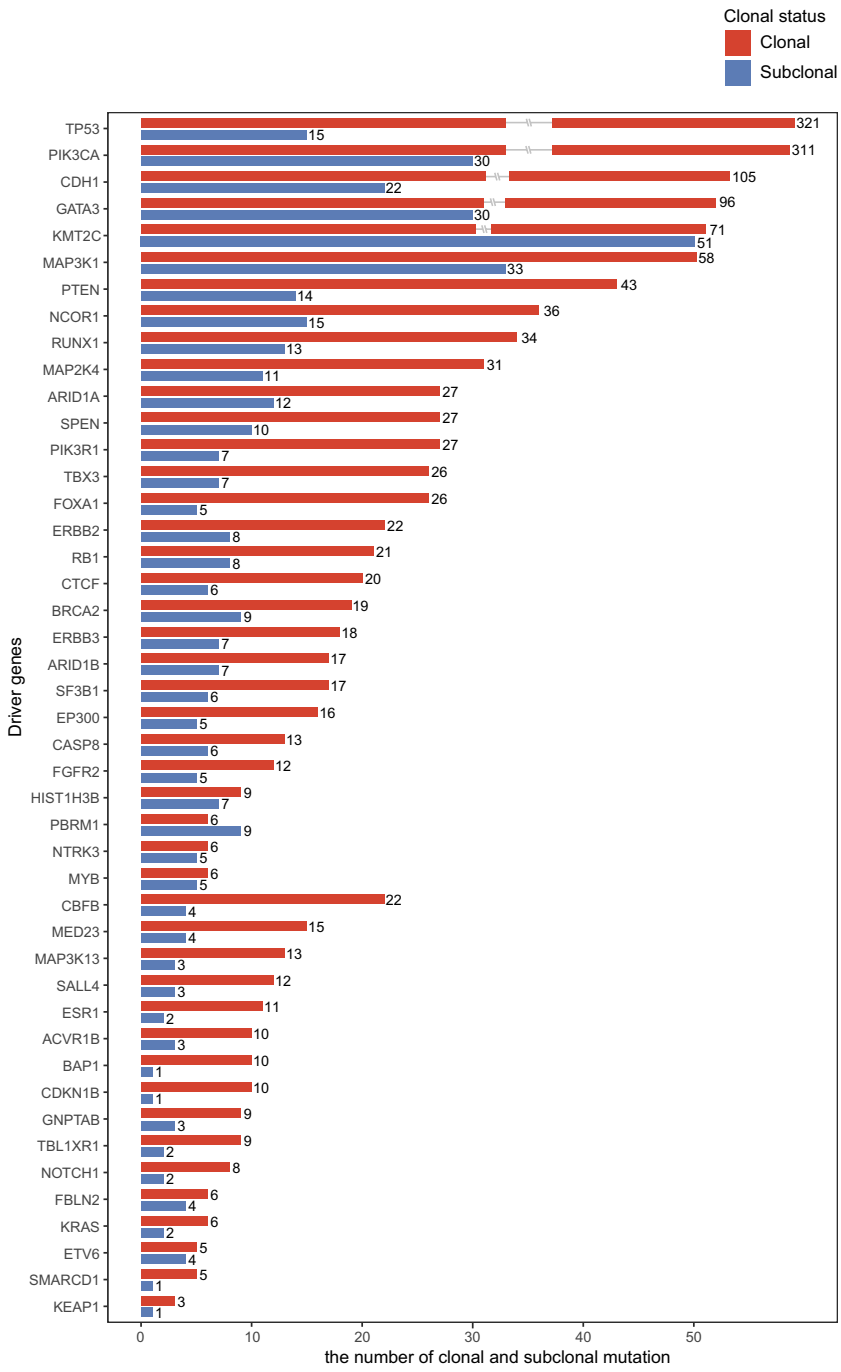

Fig. 1 The number of patients harboring clonal and subclonal driver gene mutations. Red bars correspond to clonal mutations whereas blue bars to the subclonal mutations

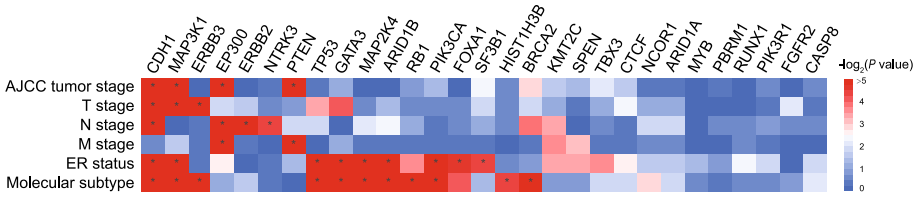

Fig. 2 The heatmap showing the significant differences between clonality of driver genes and clinicopathological characteristics. The differences were performed using Chi square test (P values < 0.05). Heatmap is color coded based on -log2(P-value)

these mutational events may be required as early events in tumorigenesis of breast cancer.

Correlations between clonality of driver genes and clinicopathologic characteristics

Next, we evaluated the associations between the mutation status of driver genes and clinicopathological factors, such as AJCC stage, TNM stage, and tumor molecular subtypes. Across the left 29 driver genes with clonal and subclonal mutations, there were 17 genes (*CDH1, MAP3K1, ERBB3, EP300, ERBB2, NTRK3, PTEN, TP53, GATA3, MAP2K4, ARID1B, RB1, PIK3CA, FOXA1, SF3B1, HIST1H3B*, and *BRCA2*) whose clonal status were significantly correlated with above factors (chi-squared test, P-value < 0.05, Fig. 2; Tables 2, 3, Tables S1–S15). Furthermore, we found that the clonality of 11 genes (*RB1, PIK3CA, MAP3K1, CDH1, ERBB3, TP53, GATA3, MAP2K4, ARID1B, HIST1H3B*,

Table 2 Association between clonal status of *CDH1* mutation and clinicopathological variables in breast cancer

	CDH1 status			
	Clonal	Subclonal	WT	P-value
T stage				< 0.001*
T1	19	5	228	
T2	55	13	504	
T3	29	4	76	
T4	2	0	34	
N stage				< 0.001*
N0	51	16	394	
N1	24	3	294	
N2	11	0	99	
N3	17	2	42	
M stage				0.675
M0	76	16	721	
M1	1	0	18	
AJCC stage				0.015*
I	13	2	149	
II	53	17	489	
III	36	2	171	
IV	1	16	16	
ER status				< 0.001*
Positive	99	22	597	
Negative	5	0	201	
Subtype				< 0.001*
Basal	1	0	162	
HER2	3	0	74	
LumA	91	17	384	
LumB	7	3	185	

Significant P values are labeled with *($P < 0.05$)

Table 3 Association between clonal status of *MAP3K1* mutation and clinicopathological variables in breast cancer

	MAP3K1 status			
	Clonal	Subclonal	WT	P-value
T stage				0.008*
T1	17	18	217	
T2	34	11	527	
T3	4	4	101	
T4	2	0	34	
N stage				0.853
N0	25	19	417	
N1	20	8	293	
N2	7	3	100	
N3	3	1	57	
M stage				0.341
M0	55	28	730	
M1	0	0	19	
AJCC stage				0.035*
I	10	13	141	
II	35	15	509	
III	11	5	193	
IV	0	0	17	
ER status				< 0.001*
Positive	51	32	635	
Negative	5	1	200	
Subtype				< 0.001*
Basal	3	2	158	
HER2	3	0	74	
LumA	49	23	420	
LumB	2	5	188	

Significant P values are labeled with *($P < 0.05$)

and *BRCA2*) were related to classical molecular subtypes (Luminal A, Luminal B, HER2+, Basal-like), and 9 genes (*TP53, PIK3CA, FOXA1, CDH1, MAP3K1, GATA3, MAP2K4, ARID1B*, and *SF3B1*) associated with ER status. The *ERBB2, ERBB3, NTRK3, PTEN, EP300, MAP3K1*, and *CDH1* genes were found to be highly correlated with the TNM stage. And four genes (*PTEN, EP300, MAP3K1*, and *CDH1*) were significantly associated with AJCC stage. In addition, a significant difference between the fraction of point to frameshift in clonal versus subclonal mutations of driver genes was found in our data (Wilcox test, $P=0.0005$). And we observed that more point mutations were enriched for clonal mutations (binomial test, $P=0.02$). We further found point mutations in most of driver genes and frameshift mutations in 25 driver genes were both significantly associated with molecular subtypes (Fisher's test, $P<0.05$, Fig. S1), such as *TP53* and *ERBB2*. And point mutations or frameshift mutations in a few driver genes had significant associations with other clinicopathological factors (Fisher's test, $P<0.05$, Fig. S1), such as *PBRM1, PBM1D, ESR1*, and *FGFR2*. In addition, we found that in different clinical stages or molecular subtypes, these driver genes were prone to be clonal mutations (Fig. S2). These results suggested that the clonality of known driver genes had potentially important relationships with clinicopathological factors.

Clinical characterization of driver genes mutation clonality in breast cancer

We wanted to examine whether the clonal architectures of driver genes were associated with clinical outcome of breast cancer patients. Among above 29 driver genes, the OS of patients harboring clonal *ERBB2, FOXA1* or *KMT2C* mutations were significantly shorter than those cases with unmutated genes ($P=6e{-}04$, $P=8e{-}04$ and $P=0.0346$, respectively). For these three driver genes, the patients with clonal mutations had worse OS than those with subclonal mutations. However, the results did not reach statistical significance, likely owing to the limited sample size (Fig. 3). Patients with subclonal *GATA3* or *RB1* mutations

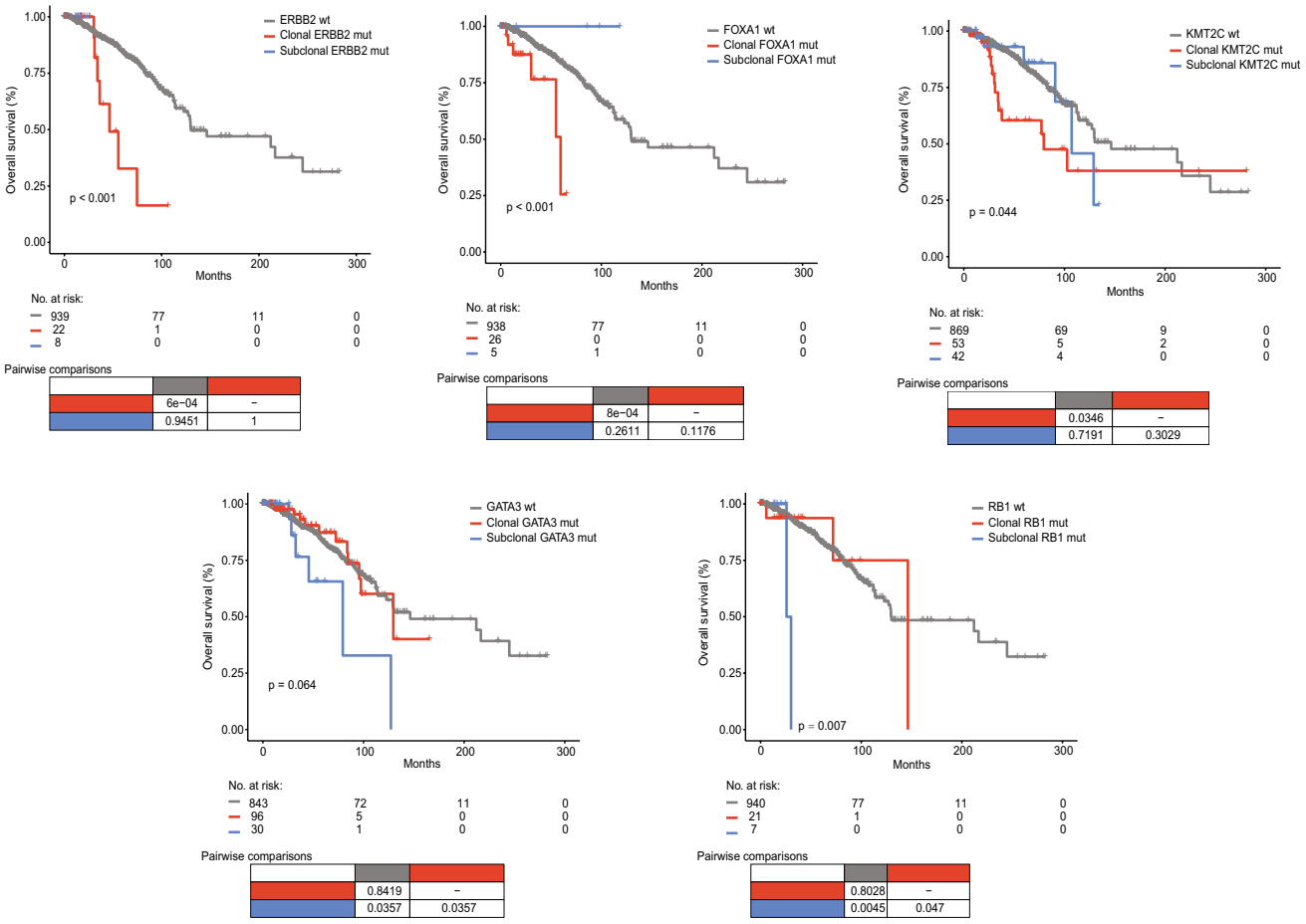

Fig. 3 The associations of driver genes with clinical outcome. Comparison of overall survival among patients carrying clonal mutations (red line), subclonal mutations (blue line), and cases harboring unmutated genes (gray line) by Kaplan–Meier analysis (with log-rank *P* values) in the cohort of breast cancer patients from TCGA. The driver genes include *ERBB2, FOXA1, KMT2C, GATA3*, and *RB1*

had significantly shorter OS compared with WT sequences ($P=0.0357$ and $P=0.0045$, respectively). Furthermore, the patients harboring subclonal *GATA3* or *RB1* mutations were significantly shorter OS than those with clonal *GATA3* or *RB1* mutations ($P=0.0357$ and $P=0.047$, respectively, Fig. 3). On account of a few silent mutations in our data, we removed silent mutations to reevaluate the prognostic values of clonality or subclonality of driver genes. We also found that patients harboring clonal *ERBB2, FOXA1*, or *KMT2C* mutations ($P=0.0048$, $P=8e{-}04$, and $P=0.025$, respectively) and subclonal *GATA3* or *RB1* mutations had significantly ($P=0.0357$ and $P=0.013$, respectively) shorter OS than those cases with unmutated genes (Fig. S3). In addition, only the patients harboring subclonal *RB1* mutations had significantly shorter DFS than those with wild type ($P=1e{-}04$, Fig. S4).

To assess whether the survival prediction ability of the clonal or subclonal mutations was independent of other clinicopathologic factors in breast cancers, univariate and multivariable Cox regression analysis was performed. The covariables included age, AJCC stage, TNM stage, ER status, histologic subtype, molecular subtype, and clonal architectures of prognostic driver genes. We found that patients harboring clonal *ERBB2* mutations (HR 4.083, 95% CI 1.540 to 10.833, $P=0.005$, Table 4), clonal *FOXA1* mutations (HR 6.718, 95% CI 2.271 to 19.874, $P<0.001$, Table S16), subclonal *GATA3* mutations (HR 5.006, 95% CI 1.723 to 14.543, $P=0.003$, Table 5), and subclonal *RB1* mutations (HR 9.738, 95% CI 2.084 to 45.507, $P=0.004$, Table S17) independently predicted poor OS of the patients with BRCA. Other clinical factors, age, and AJCC stage IV, were also independently associated with shorter OS of patients (Tables 4, 5, Tables S16–S18). The results of the multivariable Cox regression analysis thus indicated that the predictive ability of clonal or subclonal driver gene mutations was independent of clinicopathological factors for OS in breast cancers.

Table 4 Multivariate analysis for *ERBB2* of overall survival in the cohort

Variables	HR	Univariate 95% CI	*P* value	HR	Multivariate 95% CI	*P* value
Age	1.035	1.017–1.054	<0.001*	1.037	1.018–1.057	<0.001*
AJCC stage						
Stage II versus I	1.485	0.739–2.982	0.267	1.315	0.423–4.090	0.636
Stage III versus I	3.379	1.634–6.988	0.001*	1.565	0.318–7.702	0.582
Stage IV versus I	10.911	4.412–26.983	<0.001*	8.065	2.883–16.014	0.008*
T stage						
T2 versus T1	1.427	0.816–2.496	0.212	0.998	0.4302–2.317	0.992
T3 versus T1	1.754	0.836–3.680	0.137	0.870	0.280–2.700	0.809
T4 versus T1	5.099	2.281–11.401	<0.001*	1.607	0.478–5.397	0.443
N stage						
N1 versus N0	1.452	0.867–2.433	0.157	1.462	0.742–2.880	0.273
N2 versus N0	3.644	1.966–6.754	<0.001*	3.227	0.968–10.762	0.057
N3 versus N0	6.804	3.280–14.115	<0.001*	3.062	0.883–10.622	0.078
M stage						
M1 versus M0	6.310	3.140–12.677	<0.001*	3.137	0.484–20.360	0.231
ER status						
Pos versus neg	0.721	0.446–1.167	0.184	0.656	0.317–1.358	0.256
Histologic subtype						
Lobular versus ductal	0.832	0.440–1.572	0.571	0.851	0.396–1.831	0.681
Molecular subtype						
Basal versus Luminal A	1.242	0.681–2.267	0.479	1.183	0.484–2.888	0.712
HER2 versus Luminal A	1.884	0.935–3.797	0.076	1.433	0.593–3.467	0.424
Luminal B versus Luminal A	1.678	0.986–2.856	0.056	1.240	0.684–2.248	0.479
ERBB2 clonality						
Clonal mut versus wt	3.915	1.576–9.724	0.003*	4.083	1.540–10.833	0.005*
Subclonal mut versus wt	0	0-Inf	0.995	0	0-Inf	0.996

Significant *P* values are labeled with *($P<0.05$)

Table 5 Multivariate analysis for *GATA3* of overall survival in the cohort

Variables	HR	Univariate 95% CI	*P* value	HR	Multivariate 95% CI	*P* value
Age	1.035	1.017–1.054	<0.001*	1.038	1.019–1.058	<0.001*
AJCC stage						
Stage II versus I	1.485	0.739–2.982	0.267	1.534	0.497–4.757	0.456
Stage III versus I	3.379	1.634–6.988	0.001*	1.924	0.385–9.630	0.426
Stage IV versus I	10.911	4.412–26.983	<0.001*	7.122	2.504–14.501	0.012*
T stage						
T2 versus T1	1.427	0.816–2.496	0.212	0.964	0.413–2.250	0.933
T3 versus T1	1.754	0.836–3.680	0.137	0.904	0.286–2.860	0.864
T4 versus T1	5.099	2.281–11.401	<0.001*	1.684	0.500–5.671	0.401
N stage						
N1 versus N0	1.452	0.867–2.433	0.157	1.231	0.632–2.397	0.541
N2 versus N0	3.644	1.966–6.754	<0.001*	2.846	0.835–9.697	0.095
N3 versus N0	6.804	3.280–14.115	<0.001*	2.525	0.719–8.871	0.149
M stage						
M1 versus M0	6.310	3.140–12.677	<0.001*	4.106	0.615–27.392	0.145
ER status						
Pos versus neg	0.721	0.446–1.167	0.184	0.584	0.281–1.213	0.149
Histologic subtype						
Lobular versus ductal	0.832	0.440–1.572	0.571	0.988	0.460–2.123	0.976
Molecular subtype						
Basal versus Luminal A	1.242	0.681–2.267	0.479	1.253	0.514–3.056	0.620
HER2 versus Luminal A	1.884	0.935–3.797	0.076	1.515	0.622–3.692	0.360
Luminal B versus Luminal A	1.678	0.986–2.856	0.056	1.168	0.646–2.112	0.608
GATA3 clonality						
Clonal mut versus wt	0.744	0.323–1.714	0.488	1.187	0.493–2.855	0.703
Subclonal mut versus wt	3.181	1.151–8.789	0.026*	5.006	1.723–14.543	0.003*

Significant *P* values are labeled with *($P<0.05$)

Identification of prognostic clonal or subclonal mutations during diverse BRCA groups

We explored whether the clonality of prognostic driver genes was effective for TNM stage patients by Kaplan–Meier survival analysis. The patients with clonal *FOXA1* or subclonal *GATA3* mutations significantly predicted OS in T2 stage BRCA patients ($P=2e{-}04$ and $P=0.0267$, respectively) than those with wild type, with significant differences between clonal and subclonal *GATA3* mutations ($P=0.0267$, Fig. S5, Table S19). Compared with wild type patients, the clonal *ERBB2* mutations ($P<0.0001$), clonal *FOXA1* mutations ($P=0.0022$) and subclonal *RB1* mutations ($P=0.0053$) were significantly associated with OS of N0 stage patients (Fig. 4a, Table S20). While, for N1 stage patients, the subclonal *GATA3* ($P=4e{-}04$) or *RB1* mutations ($P=1e{-}04$) significantly predicted shorter OS. And the subclonal and clonal *GATA3* mutations were significant differences ($P=0.0028$, Fig. 4a, Table S20). Furthermore, the clonal *ERBB2, FOXA1*, or *KMT2C* mutations ($P=9e{-}04$, $P<0.0001$, and $P<0.0282$, respectively) and subclonal *GATA3* or *RB1* mutations ($P=0.0455$ and $P=8e{-}04$, respectively) had significantly shorter OS of M0 stage patients compared with those with wide type (Fig. 4b, Table S21). In addition, for M0 stage patients, there were significant differences between clonal and subclonal mutations in *GATA3* and *RB1* ($P=0.0089$ and $P=0.0047$, respectively, Fig. 4b, Table S21).

Furthermore, Kaplan–Meier survival analysis after patient stratification according to breast cancer molecular subtypes or AJCC stage was performed. We demonstrated that the AJCC stage II patients harboring clonal *FOXA1* or subclonal *GATA3* mutations ($P<0.0001$ and $P=0.0197$, respectively) had significantly shorter OS than wild type cases (Fig. 5a, Table S22). The significant differences between clonal and subclonal *GATA3* mutations were found in AJCC stage II and stage III patients ($P=0.0468$ and $P=0.0245$, respectively, Fig. 5a). For BRCA classical subtypes, the subclonal *ERBB2* mutations significantly predicted shorter OS of lumina A patients comparing than those with wild type ($P=0.0024$, Fig. 5b, Table S23). We also found that the significant differences between clonal and

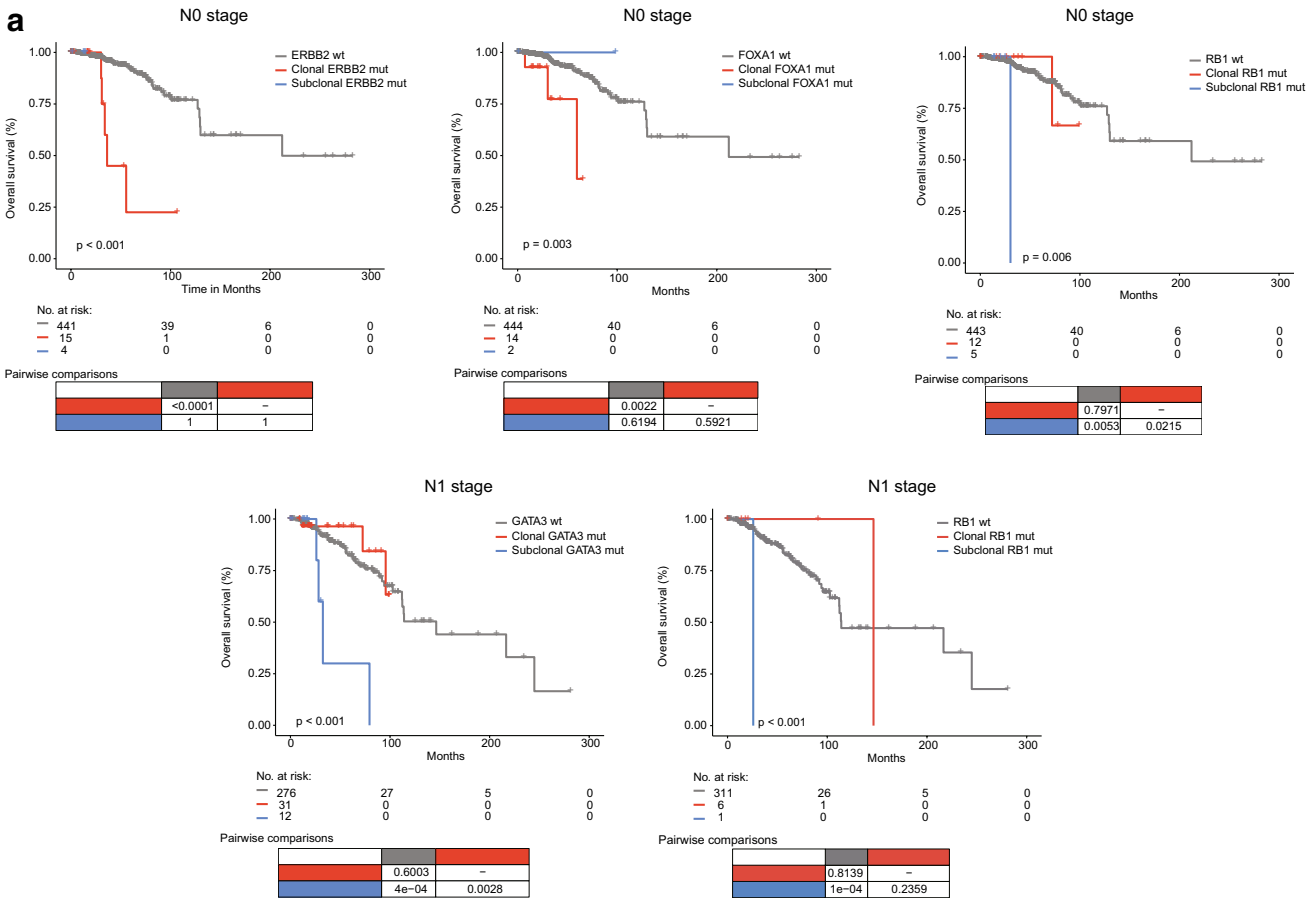

Fig. 4 Survival analysis for breast cancer patients according to clinicopathological factor TNM stage. **a** Kaplan–Meier estimates overall survival in N0 stage patients harboring clonal or subclonal *ERBB2, FOXA1* and *RB1* mutations and N1 stage patients harboring clonal or subclonal *GATA3* and *RB1* mutations. **b** Kaplan–Meier estimates overall survival in M0 stage patients harboring clonal or subclonal *ERBB2, FOXA1, KMT2C, GATA3*, and *RB1* mutations

subclonal *GATA3* mutations in these patients ($P = 0.0191$, Fig. 5b, Table S23). In addition, the patients harboring subclonal *GATA3* mutations had a trend toward shorter OS than wild type patients. These findings suggested that the clonality of driver genes had prognostic significance in early-stage breast cancer patients.

Associations between the clonality of other mutation genes with clinical outcome of patients

Indeed, there are a large number of other mutation genes not included in the known breast cancer driver genes. To investigate the prognostic values of the clonality of mutations in these genes, we performed survival analysis for the mutation genes that have subclonal mutations in at least four samples. We found 139 genes with clonal mutations and 76 genes with subclonal mutations were significantly associated with OS of patients than those cases with unmutated genes (log-rank test, $P < 0.05$). To further assess whether the survival prediction ability of the clonal or subclonal mutations was independent of other clinicopathologic factors in breast cancer, we performed univariate and multivariable Cox regression analysis of prognostic genes. Finally, we identified 40 genes with clonal mutations and 41 genes with subclonal mutations independently predicted OS of the patients with BRCA (Table S24). These findings indicate that the clonality status of other mutated genes may have a bearing on breast cancer outcomes.

Discussion

Genomic studies in cancers have revealed that these diseases have complex heterogeneity [24–26]. With the development of whole-genome and exome sequencing technologies, a large number of driver gene mutations and the associations between their clonal heterogeneity and outcome of patients were researched. Recently, some studies have confirmed that the clonal architecture of driver genes have prognostic values in diverse cancers [10, 11, 27, 28]. Therefore, we

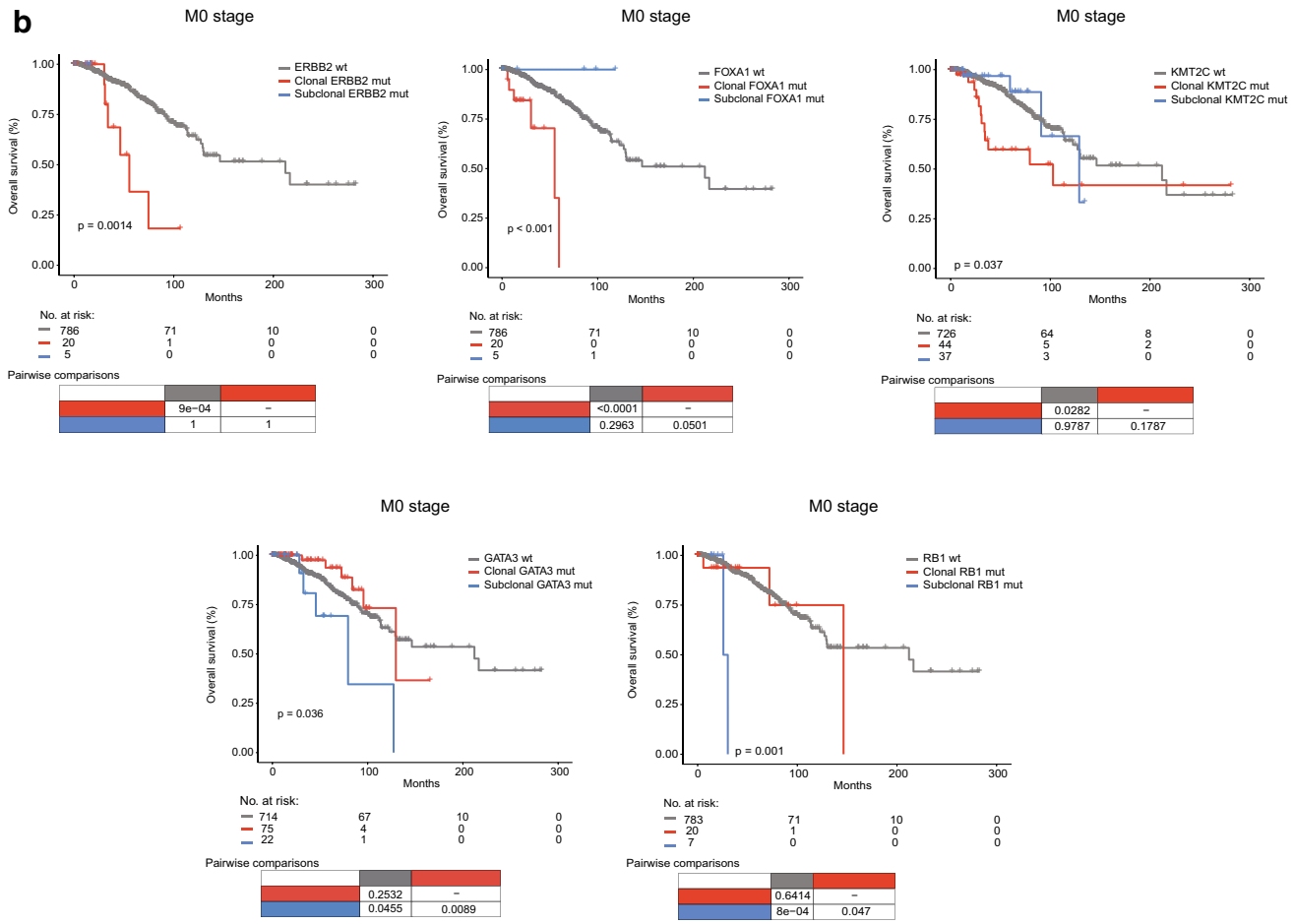

Fig. 4 (continued)

comprehensively assessed the clinical relevance of clonal or subclonal driver gene mutations on the outcome of breast cancer patients using the TCGA-BRCA data cohort.

As well as we know, the clonal mutation is carried by all tumor cells, while the subclonal mutation is carried by a fraction of tumor cells. In this report, we found that the number of clonal mutations in driver genes was substantially greater than subclonal mutations in breast cancer. The finding suggests that the majority of known cancer genes are fundamental to the tumorigenesis of cancer, generally occurring in the trunk of the phylogenetic tree during the cancer evolutionary process. Indeed, many previous studies also confirmed the phenomenon in many cancer types [29–31]. Kim et al. found that 67.9% somatic single nucleotide variations were to be clonal and 29.8% to be subclonal in glioblastoma. Shi et al. also found there was a significantly lower fraction of subclonal mutations in four driver genes in the tumor evolution analysis of lung adenocarcinoma. These results highlight a complex dynamic of tumor evolution.

In this report, we found that the clonal *ERBB2, FOXA1*, and *KMT2C* mutations and subclonal *GATA3* and *RB1* mutations predicted shorter OS of breast cancer patients compared with those with wild type. Furthermore, after adjusting for other clinicopathological factors, including TNM stage, AJCC stage, molecular subtype, histologic subtype, and ER status, the clonal *ERBB2* and *FOXA1* mutations and subclonal *GATA3* and *RB1* mutations independently predicted for patients outcome. Importantly, the mutation clonality of *ERBB2, FOXA1, GATA3*, and *RB1* significantly predicted patients' outcome with some specific BRCA tumor stages and histological subtypes. Therefore, these pivotal findings may have potential implications for clinical trials and, possibly, for disease treatment of breast cancer patients.

Using the exome sequencing data and copy number data from TCGA BRCA, we identified clonal and subclonal mutations in 55 driver genes. These driver gene mutations were preferentially clonal in breast cancers, which was consist with a previous study [32]. We confirmed that subclonal *RB1* mutations were unfavorable for OS in BRCA patients with no lymph nodes metastasis and distant metastasis. While in the progression from primary tumors to metastases of breast cancer, Schrijver et al. observed the

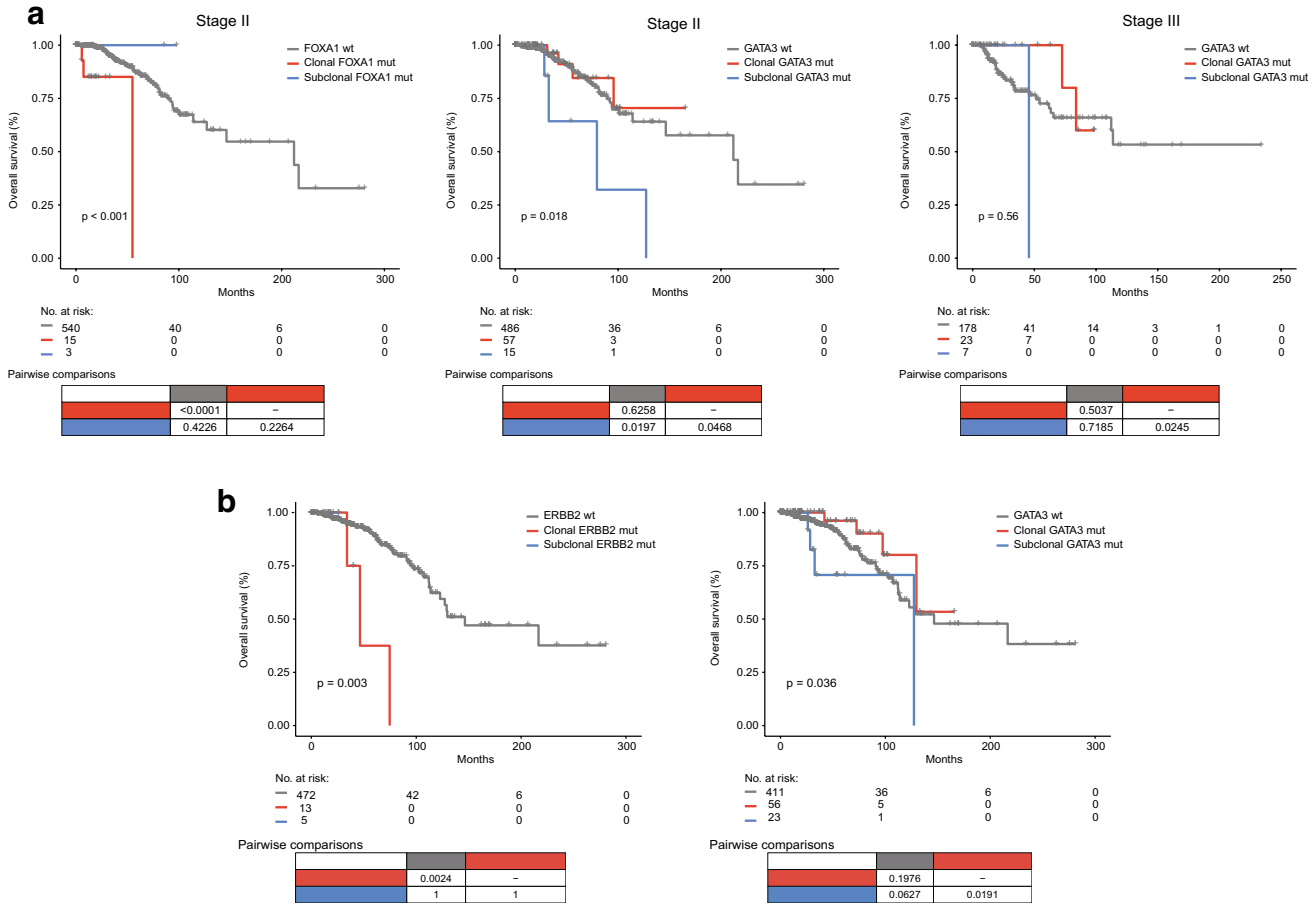

Fig. 5 Survival analysis for breast cancer patients according to AJCC stage and molecular subtypes. **a** Kaplan–Meier estimates overall survival in AJCC stage II patients harboring clonal or subclonal *FOXA1* and *GATA3* mutations and AJCC stage III patients harboring clonal or subclonal *GATA3* mutations. **b** Kaplan–Meier estimates overall survival in M0 stage patients harboring clonal or subclonal *GATA3* mutations

clonal shifts of *RB1* gene affected by somatic mutations [33]. Our findings suggested the subclonal mutations of *RB1* mainly played important functions in primary tumors. Moreover, previous exome sequencing found that the clonal *ERBB2* and *KMT2C* mutations presented in the high-grade triple-negative breast cancer [34]. We observed that the patients harboring clonal *ERBB2* and *KMT2C* mutations had shorter OS. In addition, some studies have shown that *GATA3* mutation was primarily associated with a favorable prognosis in BRCA patients [35–37]. However, in our analysis, there was no significant survival difference between patients with clonal *GATA3* mutations and WT sequences. Similarly, in TCGA cohort, we could not observe a difference in overall survival between patients harboring all *GATA3* mutations and with wild-type *GATA3* (log-rank test, $P = 0.35$), consistent with a previous study [37]. In contrast, the patients with subclonal *GATA3* mutations had significantly shorter OS compared with WT sequences in current study. This result indicated that fewer cells with *GATA3* subclonal mutations benefit from *GATA3* mutation and formed a more aggressive tumor phenotype [10, 35]. The same phenomena was also observed in some important cancer genes, such as *TP53* and *NOTCH1*[10, 11]. So their subclonal driver mutations are often associated with more aggressive tumor evolution and more adverse outcomes.

The clonality of mutation reflects the relative time of mutation occurring during tumorigenesis, and the time of somatic driver events was found to influence disease outcomes [38–40]. The accuracy and comprehensiveness of disease classification have a vital effect on cancer therapy. Because the solid tumor is highly heterogeneous, its accurate classification is difficult. Therefore, considering the clonal architecture of driver genes is necessary for developing a new cancer classification strategy. Moreover, drugs against a subclonal driver present in only a subset of tumor cells may allow the expansion of wild-type subclones, leading to the acquisition of drug resistance [41–44]. Patients harboring clonal *ERBB2* and *FOXA1*,

subclonal *GATA3*, and *RB1* mutations, could be potential drug targets. In summary, our results present a systematical characterization of the clonal or subclonal in driver gene mutations on survival of breast cancer patients and offer new insights, at the molecular level, for classification of BRCA. The position in the clonal hierarchy may refine the future treatment of breast cancers and further help achieve the goal of precision medicine for breast cancers.

Author contributions SJC, XL, and YX conceived and designed the project. YJL, EJZ, and SYL acquired the data. YJL and EJZ performed the statistical analysis and analyzed and interpreted all the data. YJL, EJZ, and SYL prepared the figures and tables. YX and YJL wrote the paper. EJZ and SYL reviewed and revised the manuscript. All authors approved the final manuscript.

Funding This study was funded in part by the National Program on Key Basic Research Project [973 Program, Grant No. 2014CB910504], the National Natural Science Foundation of China [Grant Nos. 61473106, 61573122], the China Postdoctoral Science Foundation (2016M600260), Wu lien-teh youth science fund project of Harbin medical university [Grant No. WLD-QN1407], Special funds for the construction of higher education in Heilongjiang Province [Grant No. UNPYSCT-2016049], the Heilongjiang Postdoctoral Foundation (LBH-Z16098).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours Nature 490(7418):61–70. https://doi.org/10.1038/nature11412
2. Polyak K (2011) Heterogeneity in breast cancer. J Clin Invest 121(10):3786–3788. https://doi.org/10.1172/jci60534
3. Kollareddy M, Dimitrova E, Vallabhaneni KC, Chan A, Le T, Chauhan KM, Carrero ZI, Ramakrishnan G, Watabe K, Haupt Y, Haupt S, Pochampally R, Boss GR, Romero DG, Radu CG, Martinez LA (2015) Regulation of nucleotide metabolism by mutant p53 contributes to its gain-of-function activities. Blood 6:7389. https://doi.org/10.1038/ncomms8389
4. Stemke-Hale K, Gonzalez-Angulo AM, Lluch A, Neve RM, Kuo WL, Davies M, Carey M, Hu Z, Guan Y, Sahin A, Symmans WF, Pusztai L, Nolden LK, Horlings H, Berns K, Hung MC, van de Vijver MJ, Valero V, Gray JW, Bernards R, Mills GB, Hennessy BT (2008) An integrative genomic and proteomic analysis of PIK3CA, PTEN, and AKT mutations in breast cancer. Cancer Res 68(15):6084–6091. https://doi.org/10.1158/0008-5472.can-07-6854
5. Volinia S, Galasso M, Sana ME, Wise TF, Palatini J, Huebner K, Croce CM (2012) Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. Proc Natl Acad Sci USA 109(8):3024–3029. https://doi.org/10.1073/pnas.1200010109
6. Pharoah PD, Easton DF, Stockton DL, Gayther S, Ponder BA (1999) Survival in familial, BRCA1-associated, and BRCA2-associated epithelial ovarian cancer. United Kingdom Coordinating Committee for Cancer Research (UKCCCR) Familial Ovarian Cancer Study Group. Cancer Res 59(4):868–871
7. Verhoog LC, Brekelmans CT, Seynaeve C, van den Bosch LM, Dahmen G, van Geel AN, Tilanus-Linthorst MM, Bartels CC, Wagner A, van den Ouweland A, Devilee P, Meijers-Heijboer EJ, Klijn JG (1998) Survival and tumour characteristics of breast-cancer patients with germline mutations of BRCA1. Lancet 351(9099):316–321
8. Burrell RA, McGranahan N, Bartek J, Swanton C (2013) The causes and consequences of genetic heterogeneity in cancer evolution. Nature 501(7467):338–345. https://doi.org/10.1038/nature12625
9. Yates LR, Campbell PJ (2012) Evolution of the cancer genome. Nat Rev Genet 13(11):795–806. https://doi.org/10.1038/nrg3317
10. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, Sougnez C, Stewart C, Sivachenko A, Wang L, Wan Y, Zhang W, Shukla SA, Vartanov A, Fernandes SM, Saksena G, Cibulskis K, Tesar B, Gabriel S, Hacohen N, Meyerson M, Lander ES, Neuberg D, Brown JR, Getz G, Wu CJ (2013) Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell 152(4):714–726. https://doi.org/10.1016/j.cell.2013.01.019
11. Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, Jares P, Navarro A, Martin-Garcia D, Bea S, Salaverria I, Oldreive C, Aymerich M, Suarez-Cisneros H, Rozman M, Villamor N, Colomer D, Lopez-Guillermo A, Gonzalez M, Alcoceba M, Terol MJ, Colado E, Puente XS, Lopez-Otin C, Enjuanes A, Campo E (2016) Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. Blood 127(17):2122–2130. https://doi.org/10.1182/blood-2015-07-659144
12. Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, McLean CY, Fouse SD, Yamamoto S, Ueda H, Tatsuno K, Asthana S, Jalbert LE, Nelson SJ, Bollen AW, Gustafson WC, Charron E, Weiss WA, Smirnov IV, Song JS, Olshen AB, Cha S, Zhao Y, Moore RA, Mungall AJ, Jones SJM, Hirst M, Marra MA, Saito N, Aburatani H, Mukasa A, Berger MS, Chang SM, Taylor BS, Costello JF (2014) Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science 343(6167):189–193. https://doi.org/10.1126/science.1239947
13. Suzuki H, Aoki K, Chiba K, Sato Y, Shiozawa Y, Shiraishi Y, Shimamura T, Niida A, Motomura K, Ohka F, Yamamoto T, Tanahashi K, Ranjit M, Wakabayashi T, Yoshizato T, Kataoka K, Yoshida K, Nagata Y, Sato-Otsubo A, Tanaka H, Sanada M, Kondo Y, Nakamura H, Mizoguchi M, Abe T, Muragaki Y, Watanabe R, Ito I, Miyano S, Natsume A, Ogawa S (2015) Mutational landscape and clonal architecture in grade II and III gliomas. Nat Genet 47(5):458–468. https://doi.org/10.1038/ng.3273
14. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2(5):401–404. https://doi.org/10.1158/2159-8290.cd-12-0095
15. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 6(269):pl1. https://doi.org/10.1126/scisignal.2004088
16. McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C (2015) Clonal status of actionable driver events and

the timing of mutational processes in cancer evolution. Sci Transl Med 7(283):283ra254. https://doi.org/10.1126/scitranslmed.aaa1408
17. Zhang H, Liao J, Zhang X, Zhao E, Liang X, Luo S, Shi J, Yu F, Xu J, Shen W, Li Y, Xiao Y, Li X (2018) Sex difference of mutation clonality in diffuse glioma evolution. Neuro Oncol. https://doi.org/10.1093/neuonc/noy154
18. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, Laird PW, Onofrio RC, Winckler W, Weir BA, Beroukhim R, Pellman D, Levine DA, Lander ES, Meyerson M, Getz G (2012) Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 30(5):413–421. https://doi.org/10.1038/nbt.2203
19. Miller CA, White BS, Dees ND, Griffith M, Welch JS, Griffith OL, Vij R, Tomasson MH, Graubert TA, Walter MJ, Ellis MJ, Schierding W, DiPersio JF, Ley TJ, Mardis ER, Wilson RK, Ding L (2014) SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. PLoS Comput Biol 10(8):e1003665. https://doi.org/10.1371/journal.pcbi.1003665
20. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505(7484):495–501. https://doi.org/10.1038/nature12912
21. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR (2004) A census of human cancer genes. Nat Rev Cancer 4(3):177–183. https://doi.org/10.1038/nrc1299
22. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, Stefancsik R, Harsha B, Kok CY, Jia M, Jubb H, Sondka Z, Thompson S, De T, Campbell PJ (2017) COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 45(D1):D777–D783. https://doi.org/10.1093/nar/gkw1121
23. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, Rudolph JE, Yaeger R, Soumerai T, Nissan MH, Chang MT, Chandarlapaty S, Traina TA, Paik PK, Ho AL, Hantash FM, Grupe A, Baxi SS, Callahan MK, Snyder A, Chi P, Danila D, Gounder M, Harding JJ, Hellmann MD, Iyer G, Janjigian Y, Kaley T, Levine DA, Lowery M, Omuro A, Postow MA, Rathkopf D, Shoushtari AN, Shukla N, Voss M, Paraiso E, Zehir A, Berger MF, Taylor BS, Saltz LB, Riely GJ, Ladanyi M, Hyman DM, Baselga J, Sabbatini P, Solit DB, Schultz N (2017) OncoKB: a precision oncology knowledge base. JCO Precis Oncol. https://doi.org/10.1200/PO.17.00011
24. Burke KA, Piscuoglio S, Berman SH, Reis-Filho JS, Weigelt B, van Diest PJ, Moelans CB, Dagogo-Jack I, Shaw AT (2018) Tumour heterogeneity and resistance to cancer therapies. Cancer Res 15(2):81–94. https://doi.org/10.1038/nrclinonc.2017.166
25. McGranahan N, Swanton C (2015) Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer Cell 27(1):15–26. https://doi.org/10.1016/j.ccell.2014.12.001
26. Landau DA, Carter SL, Getz G, Wu CJ (2014) Clonal evolution in hematological malignancies and therapeutic implications. Leukemia 28(1):34–43. https://doi.org/10.1038/leu.2013.248
27. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, Yoon CJ, Ellis P, Wedge DC, Pellagatti A, Shlien A, Groves MJ, Forbes SA, Raine K, Hinton J, Mudie LJ, McLaren S, Hardy C, Latimer C, Della Porta MG, O'Meara S, Ambaglio I, Galli A, Butler AP, Walldin G, Teague JW, Quek L, Sternberg A, Gambacorti-Passerini C, Cross NC, Green AR, Boultwood J, Vyas P, Hellstrom-Lindberg E, Bowen D, Cazzola M, Stratton MR, Campbell PJ (2013) Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 122(22):3616–3627. https://doi.org/10.1182/blood-2013-08-518886 (**quiz 3699**)
28. Molenaar RJ, Thota S, Nagata Y, Patel B, Clemente M, Przychodzen B, Hirsh C, Viny AD, Hosano N, Bleeker FE, Meggendorfer M, Alpermann T, Shiraishi Y, Chiba K, Tanaka H, van Noorden CJ, Radivoyevitch T, Carraway HE, Makishima H, Miyano S, Sekeres MA, Ogawa S, Haferlach T, Maciejewski JP (2015) Clinical and biological implications of ancestral and non-ancestral IDH1 and IDH2 mutations in myeloid neoplasms. Leukemia 29(11):2134–2142. https://doi.org/10.1038/leu.2015.91
29. Shi J, Hua X, Zhu B, Ravichandran S, Wang M, Nguyen C, Brodie SA, Palleschi A, Alloisio M, Pariscenti G, Jones K, Zhou W, Bouk AJ, Boland J, Hicks B, Risch A, Bennett H, Luke BT, Song L, Duan J, Liu P, Kohno T, Chen Q, Meerzaman D, Marconett C, Laird-Offringa I, Mills I, Caporaso NE, Gail MH, Pesatori AC, Consonni D, Bertazzi PA, Chanock SJ, Landi MT (2016) Somatic genomics and clinical features of lung adenocarcinoma: a retrospective study. PLoS Med 13(12):e1002162. https://doi.org/10.1371/journal.pmed.1002162
30. Kim H, Zheng S, Amini SS, Virk SM, Mikkelsen T, Brat DJ, Grimsby J, Sougnez C, Muller F, Hu J, Sloan AE, Cohen ML, Van Meir EG, Scarpace L, Laird PW, Weinstein JN, Lander ES, Gabriel S, Getz G, Meyerson M, Chin L, Barnholtz-Sloan JS, Verhaak RG (2015) Whole-genome and multisector exome sequencing of primary and post-treatment glioblastoma reveals patterns of tumor evolution. Genome Res 25(3):316–327. https://doi.org/10.1101/gr.180612.114
31. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, Dawson KJ, Iorio F, Nik-Zainal S, Bignell GR, Hinton JW, Li Y, Tubio JM, McLaren S, Butler SOM, Teague AP, Mudie JW, Anderson L, Rashid E, Tai N, Shammas YT, Sperling MA, Fulciniti AS, Richardson M, Parmigiani PG, Magrangeas G, Minvielle F, Moreau S, Attal P, Facon M, Futreal T, Anderson PA, Campbell KC, Munshi PJ NC (2014) Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun 5:2997. https://doi.org/10.1038/ncomms3997
32. Ng CKY, Bidard FC, Piscuoglio S, Geyer FC, Lim RS, de Bruijn I, Shen R, Pareja F, Berman SH, Wang L, Pierga JY, Vincent-Salomon A, Viale A, Norton L, Sigal B, Weigelt B, Cottu P, Reis-Filho JS (2017) Genetic heterogeneity in therapy-naive synchronous primary breast cancers and their metastases. Clin Cancer Res 23(15):4402–4415. https://doi.org/10.1158/1078-0432.ccr-16-3115
33. Schrijver W, Selenica P, Lee JY, Ng CKY, Burke KA, Piscuoglio S, Berman SH, Reis-Filho JS, Weigelt B, van Diest PJ, Moelans CB (2018) Mutation profiling of key cancer genes in primary breast cancers and their distant metastases. Cancer Res 78(12):3112–3121. https://doi.org/10.1158/0008-5472.Can-17-2310
34. Fusco N, Geyer FC, De Filippo MR, Martelotto LG, Ng CK, Piscuoglio S, Guerini-Rocco E, Schultheis AM, Fuhrmann L, Wang L, Jungbluth AA, Burke KA, Lim RS, Vincent-Salomon A, Bamba M, Moritani S, Badve SS, Ichihara S, Ellis IO, Reis-Filho JS, Weigelt B (2016) Genetic events in the progression of adenoid cystic carcinoma of the breast to high-grade triple-negative breast cancer. Mod Pathol 29(11):1292–1305. https://doi.org/10.1038/modpathol.2016.134
35. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, Pugh M, Jones L, Russell R, Sammut SJ, Tsui DW, Liu B, Dawson SJ, Abraham J, Northen H, Peden JF, Mukherjee A, Turashvili G, Green AR, McKinney S, Oloumi A, Shah S, Rosenfeld N, Murphy L, Bentley DR, Ellis IO, Purushotham A, Pinder SE, Borresen-Dale AL, Earl HM, Pharoah PD, Ross MT, Aparicio S, Caldas C (2016) The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun 7:11479. https://doi.org/10.1038/ncomms11479
36. Ng CK, Schultheis AM, Bidard FC, Weigelt B, Reis-Filho JS (2015) Breast cancer genomics from microarrays to massively parallel sequencing: paradigms and new insights. J Natl Cancer Inst 107 (5). https://doi.org/10.1093/jnci/djv015

37. Jiang YZ, Yu KD, Zuo WJ, Peng WT, Shao ZM (2014) GATA3 mutations define a unique subtype of luminal-like breast cancer with improved survival. Cancer 120(9):1329–1337. https://doi.org/10.1002/cncr.28566
38. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, Fisher R, McGranahan N, Matthews N, Santos CR, Martinez P, Phillimore B, Begum S, Rabinowitz A, Spencer-Dene B, Gulati S, Bates PA, Stamp G, Pickering L, Gore M, Nicol DL, Hazell S, Futreal PA, Stewart A, Swanton C (2014) Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nat Genet 46(3):225–233. https://doi.org/10.1038/ng.2891
39. Fisher R, Horswell S, Rowan A, Salm MP, de Bruin EC, Gulati S, McGranahan N, Stares M, Gerlinger M, Varela I, Crockford A, Favero F, Quidville V, Andre F, Navas C, Gronroos E, Nicol D, Hazell S, Hrouda D, O'Brien T, Matthews N, Phillimore B, Begum S, Rabinowitz A, Biggs J, Bates PA, McDonald NQ, Stamp G, Spencer-Dene B, Hsieh JJ, Xu J, Pickering L, Gore M, Larkin J, Swanton C (2014) Development of synchronous VHL syndrome tumors reveals contingencies and constraints to tumor evolution. Genome Biol 15(8):433. https://doi.org/10.1186/s13059-014-0433-z
40. Swanton C (2015) Cancer evolution constrained by mutation order. N Engl J Med 372(7):661–663. https://doi.org/10.1056/NEJMe1414288
41. Delhommeau F (2015) Effect of mutation order on myeloproliferative neoplasms. N Engl J Med 372(19):1865. https://doi.org/10.1056/NEJMc1503143
42. McGranahan N, Swanton C (2017) Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 168(4):613–628. https://doi.org/10.1016/j.cell.2017.01.018
43. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, Dunning MJ, Gale D, Forshew T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N (2013) Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 368(13):1199–1209. https://doi.org/10.1056/NEJMoa1213261
44. Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, Piskorz AM, Parkinson C, Chin SF, Kingsbury Z, Wong AS, Marass F, Humphray S, Hadfield J, Bentley D, Chin TM, Brenton JD, Caldas C, Rosenfeld N (2013) Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 497(7447):108–112. https://doi.org/10.1038/nature12065