It is made available under a CC-BY-NC-ND 4.0 International license .

1	Evolution of synchronous bilateral breast cancers provide insights into interactions
2	between host, tumor and immunity.
3	
4	Authors
5	Anne-Sophie Hamy <sup>1,2</sup> , Judith Abecassis <sup>2,3</sup> , Lauren Darrigues <sup>4</sup> , Cecile Laurent <sup>2</sup> , François
6	Zaccarini <sup>4</sup> , Benjamin Sadacca <sup>2</sup> , Myriam Delomenie <sup>4</sup> , Enora Laas <sup>4</sup> , Odette Mariani <sup>5</sup> , Thanh
7	Lam <sup>6</sup> , Beatriz Grandal <sup>2,4</sup> , Marick Lae <sup>5,7</sup> , Ivan Bieche <sup>8,9</sup> , Sophie Vacher <sup>9</sup> , Jean-Yves Pierga <sup>1</sup> ,
8	Etienne Brain <sup>1</sup> , Celine Vallot <sup>10,11</sup> , Judicael Hotton <sup>12</sup> , Wilfrid Richer <sup>11,13</sup> , Joshua Waterfall <sup>11,14</sup> ,
9	Fabien Reyal <sup>2,4</sup>
10	Affiliations
11	1Department of Medical Oncology, Institut Curie, Université de Paris, 75005 Paris, France
12 13 14	2Residual Tumor & Response to Treatment Laboratory, RT2Lab, Translational Research Department, Paris, INSERM, U932 Immunity and Cancer, Institut Curie, Université de Paris, 75005 Paris, France
15 16	3INRIA, Université Paris-Saclay, CEA, Palaiseau, 91120, France
17 18	4Department of breast, gynecological and reconstructive surgery, Institut Curie, Université de Paris, 75005 Paris, France
19	5Departement of tumor biology, Institut Curie, Paris 75005, France
20 21	6Department of Gynecology and Obstetrics, Geneva, University Hospitals, 30 bd de la Cluse, 1205 Geneva, Switzerland
22 23	7Department of Pathology, Centre Henri Becquerel, INSERM U1245, UNIROUEN, University of Normandie, 76038 Rouen, France
24 25	8INSERM U1016, Faculty of Pharmaceutical and Biological Sciences, Université de Paris, 75005 Paris, France
26	9Pharmacogenomics Unit, Department of Biopathology, Institut Curie, 75005 Paris, France
27	10CNRS UMR3244, Institut Curie, PSL University, Paris, France
28 29	11Translational Research Department, Institut Curie Research Center, PSL Research University, Paris 75005, France
30 31 32 33	12Department of Surgical Oncology, Institut de Cancérologie de Lorraine, Vandoeuvre-lès- Nancy, 6 Avenue de Bourgogne, 54519 Vandœuvre-lès-Nancy, France Department of Surgical Oncology, Institut Godinot, 1 Rue du Général Koenig, 51100 Reims, France 75005, France

It is made available under a CC-BY-NC-ND 4.0 International license .

- <sup>13</sup>PSL Research University, Institut Curie Research Center, INSERM U932, Paris, France
- <sup>14</sup>INSERM U830, Institut Curie, PSL University, Paris, France
- 37 \*: Both authors contributed equally to this manuscript
- 38

35

### **39 Correspondence:**

- 40 <u>fabien.reyal@curie.fr</u>
- 41 joshua.waterfall@curie.fr

## 42 Abstract (219 words)

43

44 Synchronous bilateral breast cancer (sBBC) occurs after both breasts have been affected by 45 the same germline genetics, reproductive life factors and environmental exposures for 46 decades. It represents an opportunity to decipher the complex interplay between host, tumor, 47 immune system and response to neoadjuvant chemotherapy (NAC). On a cohort of 17575 48 BCs treated between 2005 and 2012, sBBCs (n=404) were associated with less aggressive 49 proliferative patterns and higher rates of luminal breast cancers (BCs) when compared with 50 unilateral BCs (n=17171). The left and right tumors were concordant for the majority of 51 clinical and pathological features. Tumor pairs of concordant BC subtype were more frequent 52 than pairs of discordant BC subtype, with notably a particularly high frequency of pairs of 53 luminal BCs. Intriguingly, both the levels of tumor infiltrating lymphocytes (TILs) and the 54 response to NAC were modified by the subtype of the contralateral tumors. Whole exome 55 sequencing and RNAseq analyses revealed that left and right tumors were independent from a 56 somatic mutation and transcriptomic point of view, while primary tumors (PT) before NAC 57 and specimens with residual disease (RD) after NAC were more closely related. The analysis 58 of the TCR repertoire identified very little overlap between patients, while common clones 59 were shared in bilateral tumors within each patient. After NAC, the TCR repertoire of RD was 60 enriched and expanded with clones edited by the contralateral PT.

- 61
- 62

It is made available under a CC-BY-NC-ND 4.0 International license .

#### 63 Introduction

64

Bilateral breast cancers (BBCs) represent 2% to 11% of breast cancers  $^{1-3}$  and their incidence 65 is increasing due to advances in breast cancer imaging <sup>4</sup>. This entity includes both 66 synchronous bilateral breast cancers (sBBCs) and metachronous bilateral breast cancer 67 68 (mBBCs), according to the length of time between the two diagnoses, with typical cut-offs 69 ranging from 3 to 12 months after diagnosis of the first tumor. sBBCs are associated with a poorer survival than unilateral cancer in several studies <sup>2,5,6</sup>. In a study of 123,757 women 70 71 with a primary BC diagnosed from 1970 to 2000, the cumulative breast cancer specific 72 mortality was 45% for women with sBBC and 33% (p<0.001) in case of unilateral breast cancer<sup>2</sup>. Beyond familial history of breast cancer, which is more frequent in women with 73 74 BBC, neither synchronous nor metachronous breast cancer is associated with strong genetic determinants and only 5% of patients with BBC carry BRCA1 or BRCA2 mutations <sup>5</sup>. Hence, 75 76 per the 2020 NCCN guidelines, bilaterality is not considered as a factor clinically indicating 77 genetic testing, but it may be considered in patients diagnosed with BBC between the age of 50 and 65 years  $^7$ . 78

79 From the genomic point of view, several studies have investigated clonal relationships between BBCs, with the majority reaching the conclusion that most of – if not all - BBCs 80 were independent events  $^{8-14}$ . These studies used several methods to address the question of 81 82 the origin of BBCs, including X chromosome inactivation analysis, comparisons of allelic imbalance patterns <sup>10,12,15</sup>, or the distribution of p53 mutations <sup>11,16</sup>, karyotyping <sup>17</sup>, and 83 comparative genomic hybridization (CGH)<sup>1418</sup>. Many of these studies were based on 84 85 technologies with limited resolution for assessing clonality. Recently, analyzing a targeted 86 sequencing panel of 254 genes including all known driver genes in BC, Begg and colleagues<sup>19</sup> 87 investigated the clonality of 30 synchronous and 19 metachronous BBC pairs, and found that 88 only 3 out of 49 pairs harbored strong molecular evidence supporting a clonal relationship 89 between BBCs. However, only one pair of sBBCs was interpreted as clonally related (2 90 shared mutations out of 3 mutations identified), leaving the question of the independence 91 between sBBCs still open.

92

93 The immune microenvironment and especially the role of tumor-infiltrating lymphocytes 94 (TILs) in BC has been studied extensively in the last decade. The drivers of BC 95 immunosurveillance derive from both (i) tumor-intrinsic characteristics; and/or (ii) extrinsic 96 factors related to the host or the environment. Among endogenous tumor characteristics,

It is made available under a CC-BY-NC-ND 4.0 International license .

97 molecular features (BC subtype, proliferative patterns), expression of human leukocyte antigen (HLA)-class I, tumor mutational burden<sup>20</sup>, and activation of particular cellular 98 signaling pathways<sup>21</sup> have been found to be associated with immune infiltration. Extrinsic 99 factors including host characteristics (gender <sup>22</sup>, age <sup>23</sup>, body mass index), environment 100 101 (tobacco, alcohol), nutritional factors, diet, commensal microbiota, physical activity, and 102 hormonal exposure have been studied less extensively. Even if our understanding of the 103 immune microenvironment has improved, it remains unclear to what extent anti-tumor 104 immunity is driven by the tumor, by the host, or the interaction between the host and the 105 tumor.

106

107 Neoadjuvant chemotherapy (NAC) is currently administered to patients with locally advanced 108 breast cancers. Beyond increasing the rate of breast-conserving surgery, it serves as a test of 109 in vivo chemo-sensitivity, and the analysis of residual tumor burden may help understanding 110 resistance to treatment. Molecular BC subtypes and the density of tumor-infiltrating immune 111 cells are both considered as important predictive and prognostic factors for optimal risk 112 stratification and treatment individualization. Many studies have reported associations between high levels of TILs at diagnosis and better response to NAC<sup>24,25</sup>, and better 113 114 prognosis in both neoadjuvant and adjuvant chemotherapy, particularly for triple negative BCs (TNBC) and HER2-positive BC<sup>26</sup>. 115

116

117 sBBCs occur after both breasts have been affected by the same germline genetics, 118 reproductive life factors and environmental exposures for several decades. Two tumors 119 arising concomitantly in a host mimic a model where (i) extrinsic factors are almost fully 120 shared by the same host; (ii) *intrinsic factors* are specific to the tumor of each side; (iii) the 121 *immune tumoral microenvironment* resulting of the interaction between the same host and two 122 different tumors can be compared. In the case where these tumors are treated by NAC, sBBCs 123 represent a unique opportunity to decipher the interactions between tumor, microenvironment, 124 and response to treatment.

125

126 In this study, we describe on a large institutional cohort of sBBCs the relationships between 127 factors related to host, tumor phenotypes, levels of TILs and response to NAC. In addition, we 128 identified a rare resource of highly selected tumors, *i.e.*, sBBCs treated by NAC, and 129 performed whole exome sequencing (WES), copy number variation, and RNA sequencing

It is made available under a CC-BY-NC-ND 4.0 International license .

- 130 (RNAseq) to draw a comprehensive analysis of somatic alterations, the immune
- 131 microenvironment and the tumor evolution under treatment.

132

It is made available under a CC-BY-NC-ND 4.0 International license .

#### 133 **Results**

### 134 Patient and tumor characteristics

135 Out of 17575 BC patients in our institutional clinical database, 404 patients had synchronous 136 bilateral BC (2.3%) (FigS1). Patients with sBBC were slightly older at diagnosis (60.4 versus 137 58.2 years, p=0.001) and had a higher mean BMI (25.6 versus 24.7, p<0.001) than patients 138 with unilateral breast cancer. A lower proportion of grade 3 (20.9% versus 29.4%, p<0.001), 139 tumors of TNBC (7.7% versus 11.1%) or HER2-positive (6.1% versus 11.4%, p<0.001) 140 subtype, or with lymphovascular invasion (21.2% versus 28.4%, p<0.001) were observed in 141 bilateral tumors than in unilateral BCs (TableS1 and FigS2). ER and PR staining percentages 142 were lower in unilateral than in bilateral BCs. Out of 313 patients with invasive sBBCs, most 143 of the tumors were luminal (n=538, 87.6%), whereas TNBC (n=44, 7.2%) and HER2-positive 144 BCs (n=32, 5.2%) were rare (TableS2). Overall, paired sBBC tumors shared more common 145 characteristics than expected by chance (TableS3): the majority (84.7%) of the tumor pairs 146 were concordant regarding clinical and pathological patterns, notably regarding BC subtype 147 (Fig1A). In luminal tumors, ER and PR staining intensity of stained cells were lower when the 148 pair of tumors was discordant than when the pair of tumors was concordant (Fig1B-C).

149

# 150 Baseline immune infiltration

151 Immune infiltration levels before treatment were available for 277 pairs of tumors. TIL levels 152 were positively associated notably with a younger age, palpable, higher grade tumors, of no 153 specific histological type, with a higher number of nodes involved and the presence of 154 lymphovascular invasion (TableS4). TILs were higher in TNBC and HER2-positive BCs than 155 in luminal tumors. Interestingly, the relationship between TIL levels and BC subtype showed 156 a systemic effect, *i.e.*, it was affected by the subtype of the contralateral tumor. In luminal 157 BCs, stromal (Str) TIL levels were lower when the subtype of the contralateral tumor was 158 concordant than when it was discordant, and the same trend was observed for intratumoral 159 (IT) TILs (Fig1D-E). Conversely, in TNBCs, the IT TIL levels were lower when the subtype 160 of the contralateral tumor was concordant than when it was discordant. The interaction test 161 was highly significant (*P<sub>interaction</sub>*=0.0006), indicating that the impact of tumor subtype on IT 162 immune infiltration was significantly modified by the concordance of the BC subtype of the 163 tumor pair it belonged to. This suggests that TIL levels are not purely determined by local 164 tumor microenvironment properties.

165

It is made available under a CC-BY-NC-ND 4.0 International license .

### 166 Response to NAC

167 Out of 313 patients, 50 patients received neoadjuvant chemotherapy (NAC). Twenty two out 168 of 100 tumors reached pathological complete response (pCR). In pairs with concordant tumor 169 types, the pCR rate was higher for TNBC tumors than for luminal BC (50% versus 9%) 170 (Fig1F-G). As was seen for TIL levels, these pCR rates showed a systemic effect when the 171 contralateral tumor subtype was discordant. The pCR rate was higher in luminal BCs when 172 the contralateral pair was of different subtype (9% versus 22%); while the opposite pattern 173 was found in TNBCs (50% versus 27%), and the interaction test was significant ( $P_{interaction}$ 174 =0.07), suggesting that the response to treatment was modified by the BC subtype of the 175 contralateral tumor differentially in luminal and in TNBCs tumors.

176

# 177 Genome-scale analyses

178

179 For 6 patients with sBBCs treated with NAC, frozen material of sufficient quality was 180 available to perform tumor/normal whole exome sequencing (WES) and tumor RNA 181 sequencing (RNAseq) in both left and right pre and post NAC samples (including one patient 182 with a multicentric bilateral breast cancer) (FigS3, TableS5). Germline pathogenic mutations 183 in BC predisposition genes were identified in four patients (BRCA1, n=2, BRCA2, n=2). 184 Among the 14 primary tumors (PT), 9 were of luminal subtype and 5 were TNBCs. All 185 patients received standard sequential anthracyclines-cyclophosphamide followed by taxanes. 186 After NAC completion, 6 out of 14 tumors had residual disease (RD), while 8 tumors reached 187 pCR.

- 188
- 189

### Somatic single nucleotide and indel mutations

190 Twenty tumor samples (pre-NAC: n =14; post-NAC: n=6) were profiled by WES and 191 RNAseq and distant juxta-tumor samples were used for germline WES sequencing in each 192 patient. A median of 151.5 somatic mutations were detected per tumor (PT: 161.5 versus RD: 193 54), and a median of 3 mutations were annotated as potential drivers (splice site, nonsense, 194 frameshift or non-synonymous SNVs in COSMIC census gene) in each sample (PT: 3.5 195 versus RD: 3) (TableS6). Tumors that reached a pCR had a similar median number of 196 mutations (median 195 (driver: 5)) to those who had RD (median: 84 (driver: 2.5) and no 197 difference was seen according to BC subtype (luminal: 160 (driver: 3) versus TNBC: 182 198 (driver: 4)).

It is made available under a CC-BY-NC-ND 4.0 International license .

The three most frequently mutated COSMIC genes were *TP53* (3 patients, 8 samples), *PTEN* (3 patients, 3 samples), and *ARID1A* (2 patients, 2 samples) (Fig2A). The majority of the remaining genes were mutated only in single patients, consistent with a high genetic diversity of mutations BC. No mutation was shared between the left and right side of the PTs from any patient, consistent with the contralateral tumors developing from independent clones (Fig2B). The majority of the mutations of a tumor were shared between a PT and its RD (median percentage of shared mutations in pairs of PT/RDs: 51.3%).

206

## 207 *Copy number alterations (CNAs)*

208 Copy number analysis of the WES profiles identified recurrent gains (8 out of 12 samples) at 209 1q, 8q, 17q and recurrent losses (over 50% of the samples) at 4p, 8p, 6q, 13q, 16q and in a 210 lesser extent 1p (FigS4). Most of the alterations were not shared between the left and the right 211 side (Fig3A), while most of the CNAs were common between PT and RD (Fig3B).

212

# 213 *Mutational signatures*

214 We analyzed mutational signatures by deconvoluting the frequency of the 96 different possible trinucleotide substitutions against known signatures of mutation patterns <sup>27</sup>. We 215 216 found that the four most abundant base substitution signatures were signature 3(34%), 217 signature 5 (17%), and signature 40 (10%) (Fig3C). Signature 3 is known to be associated 218 with BRCA1/2 mutations (Alexandrov et al., 2020) and its contribution was high (54%) in all 219 tumors from patients carrying BRCA germline mutations (Patient #1, #2, #4, #6) (Fig3C). 220 Furthermore, signature 3 was also found in the left tumor of patient #5 who carried a somatic 221 BRCA1 mutation, while it was absent from the right-side tumor which did not show any 222 BRCA1-2 mutations. Signatures 5, 40 and 1 are commonly found in diverse tumor types and 223 are associated with aging and were generally found at consistent levels between bilateral 224 tumor pairs. Similarities regarding mutational processes were lower within the left and right 225 side of the PTs than within pairs of PT-RDs (Fig3D).

- 226
- 227

## Clonality and phylogenetic evolution

We determined the phylogenetic evolution between the germline profile of the patient to the left and right primary tumors and ultimately to the residual disease if present (Fig4A-E and supplemental results). The number of clones identified in each tumor sample ranged from 2 to 6. Genomic profiling found no common clones between bilateral primary tumors of the same patient, showing that these tumors arose through unrelated tumor evolution processes. Further

It is made available under a CC-BY-NC-ND 4.0 International license .

illustrating this, left and right tumors showed distinct evolution regarding the homologous
recombination deficiency (HRD), with the emergence of a somatic mutation on the *BRCA1*gene in only one out of the two tumors of a patient (Fig1E).

For patient #3, for which we monitored evolution after NAC in both tumors in parallel, we did not observe any common clonal evolution after NAC. In all tumors, several clones disappeared under NAC (Fig4A and Fig4C, light turquoise clones; Fig4E grey clone), consistent with subclone specific responses to therapy. Finally, the majority of clones that emerged after NAC did not show identifiable additional genetic drivers, suggesting potential non genetic additional mechanisms could play a role in the tumor evolution under treatment.

Altogether, these results suggest that left and right PTs are not clonally related and that their evolution under NAC does not converge to a common profile. Hence, RD is closer to its associated primary tumor than to its concomitant contralateral tumor.

- 245
- 246

### Particular case of multicentric tumors

247 One patient of the six had a bilateral multicentric tumor (Patient #6) in the context of a 248 BRCA2 pathogenic germline mutation. While the left and the right tumors shared no common 249 mutations, the two tumors from each side shared a majority of common genetic alterations at 250 both the substitution (Fig5A) and copy number levels (Fig5B). Mutation signature analyses 251 demonstrated a dominant genomic HRD footprint (signature 3, Fig5C). On each side, 252 phylogenetic reconstruction clearly indicated that multicentric tumors were clonally related, 253 with one tumors evoluting to a neighboor tumor through the extinction / emergence of 254 particular subclones.

255

257

### 256 Transcriptomic alterations

### Tumor clustering

258 We first performed unsupervised hierarchical clustering based on transcriptomic profile of the 259 most variable genes. With the exception of tumor PT\_5L, the sample clustering first split by 260 BC subtype (luminal and TNBCs) (FigS5), while gene clustering split the 2846 genes into 261 four main clusters. The clustering first split samples into a group of luminal tumors and a 262 group of TNBCs (with the exception of the ER weakly positive (+25%) luminal tumor PT5\_L 263 clustering into the TNBC group). The group of luminal tumors was globally enriched in genes 264 from cluster 1 (early and late response to estrogens); and a subset of luminal tumors were also 265 enriched in genes from the cluster 2 (TNF signaling, myogenesis, epithelial mesenchymal 266 transition). The majority of TNBC samples were enriched in genes from cluster 3 (G2M

It is made available under a CC-BY-NC-ND 4.0 International license .

267 checkpoints, E2F targets, cellular cycle). Genes from cluster 4 showed no clear enrichment in 268 specific pathways and were expressed in complex patterns by both luminal and TNBC tumor 269 subsets. Within each subgroup, the PT samples consistently clustered with their related RD 270 rather than the tumor from the contralateral side. This suggests that beyond the well-known 271 distinct gene expression profiles according to BC subtype, PT and RD are closer from a 272 transcriptomic point of view than are left and right tumors from the same patient.

273 274

## Immune infiltration

275 We applied the CIBERSORT algorithm to deconvolute RNAseq expression profiles into 22 276 subsets of immune subpopulations. Out of the 20 samples, the top 3 most abundant immune 277 subpopulations were M2 macrophages (median: 32), CD4 memory resting T cells (median: 278 26) and M1 macrophages (median: 11), while the majority of immune subpopulations was not 279 identified at significant levels, whether in PT or RD (FigS6A). CD4 memory T cells and M2 280 macrophages were increased in RD compared with PT (FigS6B); T follicular helper cells and 281 M1 macrophages were significantly higher in TNBC than in luminal BCs, while resting mast 282 cell levels were lower (FigS6C). At the patient level, we did not identify any additional 283 patterns between the left and the paired right tumor (FigS7), nor between the PT and the 284 paired corresponding residual disease (FigS8).

We compared the predicted immune contexture patterns between samples of the cohort using a dissimilarity index (FigS9). Analyzing the entire cohort, the dissimilarity was lower among the PT samples (blue area) than among the RD samples (orange area) (mean: 0.24 *versus* 0.37, p=0.009), while the greatest difference was seen between PTs samples compared with RDs samples (yellow area, mean = 0.49)

290 Analyzing paired data, the mean dissimilarity index was significantly lower between samples 291 of the same side of patient 6 than between the remaining samples (red bordered squares, mean 292 = 0.025), while the dissimilarity indices between the left and the right sides (green bordered 293 squares, mean = 0.22), and between PT and related RDs (yellow bordered squared, mean = 294 (0.29) were not statistically different from the rest of the samples. These results suggest that 295 the composition of the immune microenvironment is strongly driven by the treatment history 296 (PT or RD), while the similarity between a tumor from one side or its contralateral side, or a 297 given tumor and its paired post-chemotherapy sample is lower.

- 298
- 299

It is made available under a CC-BY-NC-ND 4.0 International license .

300	TCD soon	oncina	analycia
300	ICK sequ	lencing	anarysis

To further investigate the T cell response to NAC and to compare infiltrating TCR repertoires in clonally distinct tumors within single patients, we extracted TCR beta CDR3 sequences from the RNAseq data, and 6708 clonotypes were identified in the whole cohort. Overall, 89.4% were unique to individual samples (n=5994) (FigS10A), 10.2% were patient specific but shared by multiple samples (n=682), and only 32 clonotypes (0.5%) were identified in different patients (FigS10B).

307

308 150 clonotypes (2.2%) were found in VDJdb, a curated database of T-cell receptor sequences 309 of known antigen specificity (FigS10C). The most common antigens for this subset were 310 CMV and Influenza, but the wide majority of clonotypes were unique to a sample. Among 311 135 clonotypes found both in PT and paired RD, the read count increased significantly after 312 NAC (p=0.005) (Fig6D).

The evolution under NAC of non-unique clones is summarized for each patient in Fig6. While the most common sharing relationships were between PT and matched RD (orange clonotypes) followed by left and right tumor, a subset of clones identified in an index PT were found in the contralateral RD (red clonotypes) or emerged *de novo* in both RD samples of the same patient (pink clonotypes). Finally, common shared TCRs were extremely rare between individuals, and most of the TCR clonotypes of a tumor were unique. However, a shared TCR repertoire was identified between

320 the left and right primary tumors. An enrichment of TCR repertoire after NAC was identified,

321 including clonotypes deriving a from the contralateral primary tumor.

It is made available under a CC-BY-NC-ND 4.0 International license .

#### 322 **Discussion**

323

In the current study, we conducted a large comprehensive overview of sBBCs, integrating clinical, pathological data and data on immune infiltration, together with genomic data generated using modern whole exome and RNA sequencing technologies. It provides important insights to understand the relationships between tumor, host, immunity and response to treatment.

329

First, we found no indication that sBBCs could be genomically related but rather we conclude
that they represent two distinct and independent diseases occurring incidentally at the same
time.

In line with previous studies <sup>28293031</sup>, we found a high concordance between the clinical and 333 334 biological patterns within pairs of sBBCs. Notably, the proportion of luminal tumors was very high <sup>32 33</sup>, while *HER2*-positive BCs were rare as previously reported <sup>30</sup>. In our study, despite 335 336 the frequent clinical and pathological concordance between the left and right tumors, we 337 demonstrate clearly that these tumors were genomically independent in terms of mutations, 338 copy number alterations, expression patterns, and clonal composition. This finding is in line with most published studies<sup>8,12,14,19</sup>, despite with older technologies. In line with previous 339 340 works, we also identified genomically related profiles in multicentric tumors, thus confirming 341 with modern technologies that multi-focal tumors often represent intramammary dissemination of a single breast cancer <sup>15</sup> <sup>14</sup>, <sup>1734</sup>. Overall, these results are consistent with the 342 fact that the occurrence of sBBCs is explained by nongenetic factors <sup>35</sup>. 343

344

345 Second, we found that the immune infiltration was not purely determined by local tumor 346 microenvironment properties, as TILs levels were different according to the BC subtype of 347 the contralateral tumor. Published studies regarding immune infiltration in sBBCs is scarce, if 348 any. Analyzing samples of sBBCs from 23 patients, one study found that nuclear grade was significantly associated with the number of FoxP3 positive TILs <sup>36</sup>. In the current work, we 349 350 analyzed immune infiltration on a large number of sBBCs and confirmed that BC subtype was 351 a master driver of immunosurveillance in BCs. We also found that immune infiltration was 352 modified by the presence of a contralateral tumor: luminal BCs associated with a luminal 353 contralateral BCs had lower str TIL levels than luminal BCs associated with a contralateral BCs of a different subtype. Similarly, a TNBC associated with a contralateral TNBC had 354 355 higher IT TIL levels than TNBCs associated with a contralateral BC of a different subtype.

It is made available under a CC-BY-NC-ND 4.0 International license .

Several hypotheses can be drawn to explain this observation. First, the immune system might be activated by an index tumor, and immune cells activated by this process might spread to the contralateral tumor; resulting in differences in TIL levels. Second, as luminal BCs associated with a contralateral tumor of another subtype were associated with a lower degree of "luminalness" (as defined by ER and PR staining levels), we cannot exclude that the highest immune infiltration is derived from such patterns rather than from the presence of the contralateral tumor.

363

364 Third, the response to NAC was also significantly modified in the case of discordant 365 contralateral tumor subtypes, as with TIL levels. Evidence regarding the influence of 366 contralateral tumor on the response to treatment has been rarely described. In a cohort of 119 patients enrolled in four German neoadjuvant trials, Reinisch and colleagues <sup>37</sup> reported that 367 368 patients with BBCs had lower pCR rates than patients with unilateral BC (12.6% versus 369 20.9%). These lower rates could partially be explained by the tumor characteristics, as the 370 percentage of hormone receptor positive tumors was significantly higher in the indicator 371 lesions of the sBBCs when compared with the unilateral cancer cohort. However, the response 372 to NAC in sBBCs was lower after statistical adjustment for baseline parameters (p = 0.077373 with a small number of pCR (13/119)). In our study, it remains unknown whether the 374 difference of the rates of pCR according to the contralateral tumor are the consequence of 375 different baseline immune infiltration levels, or if the contralateral tumor modifies *per se* the 376 chemosensitivity of an index tumor.

377

Fourth, in tumors resistant to treatment and for which we could perform NGS analysis, the primary tumor and the residual specimen shared a large majority of SNVs and copy number aberrations, but the genomic profile did reveal important changes after NAC. Notably, resistant clones emerged and new potential driver mutations appeared in an independent manner on both sides. Hence, beyond a weak influence of the contralateral tumor on either immune infiltration and/or response to NAC, clonal evolution under NAC seems specific to each tumor.

- 385
- 386

Finally, we report that the infiltration of T cell clones during NAC leads to an enrichment in the TCR repertoire of post-NAC samples, either derived from the contralateral tumor, either arising in both bilateral tumors with residual disease. At a time where bilateral tumor contexts

It is made available under a CC-BY-NC-ND 4.0 International license .

390 represent a model of growing interest to understand mechanisms underlying immune response

391 to anticancer treatment in mice  $^{31}$ ,  $^{38}$ , we provide human data regarding the temporal analysis

- 392 of the TCR repertoire in sBBCs.
- 393

394 Our study has several strengths. We analyzed genomic and transcriptomic data with modern 395 technologies. Whole exome sequencing is more informative than targeted sequencing for 396 determining clonal relationships, as it enables the comprehensive identification of mutations 397 present in both breast cancers in genomic locations that were not included in the sequencing 398 panel. Whole genome sequencing could further identify mutations in non-coding regions of 399 the genome. Second, we studied a rare and unique cohort of patients, enabling direct 400 comparison of left versus right PT together with a temporal analysis comparing paired 401 samples before *versus* after NAC. Hence, beyond all the challenges in analyzing tumor evolution from bulk sequencing data<sup>39</sup>, we were able to leverage multiple tumor samples to 402 403 reconstruct a clonal phylogeny from germline data to left and right sBBCs both before and 404 after treatment. Third, to our knowledge, our data on immune infiltration are novel 405 contributions to the literature and provide insights into the immune mechanisms underlying 406 the biology of sBBCs. This study also has limitations. First, we were only able to sequence a 407 limited number of cases. We cannot eliminate the possibility that a subset of clonally related 408 sBBCs could be identified if larger cohorts were sequenced. Second, the cohort was enriched 409 in patients with BRCA mutations (although this was not an inclusion criteria), and the latter 410 might represent tumors with particular biological patterns. Recent studies suggest however 411 that the proportion of BBC patients harboring germline pathogenic variants in cancer susceptibility genes may represent up to one third of the patients <sup>40</sup>. Finally, characterization 412 413 of the immune microenvironment by bulk sequencing approaches has inherent limitations and 414 new insights could be generated by single cell technologies.

415

To conclude, our data suggest that the similarity of molecular portraits in sBBCs derives from common environmental factors but not from genetic clonal alterations. Both tumor immune infiltration and response to treatment are influenced by the presence of a contralateral tumor. Pairs of tumors from different BC subtypes should be considered as singular entities before primary systemic treatment is considered as observed responses might deviate from wellknown profiles of response to chemotherapy.

422

It is made available under a CC-BY-NC-ND 4.0 International license .

### 423 Material and methods

## 424 *Patients and treatments*

425 We identified a cohort of 17575 patients with non-metastatic BC treated at the Institut Curie 426 (Paris and Saint Cloud, France) between 2005 and 2015 in the institutional database (CNIL 427 number 1766392 - v1). Patients were treated according to local guidelines. When indicated, 428 chemotherapy was administered in a neoadjuvant or adjuvant setting, endocrine therapy was 429 indicated in the case of positivity for hormone receptor and according to prognostic factors, 430 and patients with HER2-positive tumors received neoadjuvant and/or adjuvant trastuzumab 431 from 2007 onwards. This study was approved by the Breast Cancer Study Group and by the 432 Institutional Review Board of Institut Curie and was conducted in accordance with 433 institutional and ethical rules regarding research on tissue specimens and patients. 434 Synchronous bilateral breast cancers (sBBCs) were defined as the occurrence of primary 435 tumors occurring in both breasts with a time interval no greater than 6 months. Metachronous 436 BCs defined as a time interval greater than 6 months between the diagnoses of the first and 437 second tumours were not included in the current study. Patients with exclusive in situ 438 carcinoma (DCIS) in one of the two sides were excluded from the analyses.

439

## 440 *Tumor samples and BC subtype*

In accordance with guidelines used in France<sup>41</sup>, cases were considered estrogen receptor (ER) 441 442 or progesterone receptor (PR) positive if at least 10% of the tumor cells expressed estrogen 443 progesterone receptors (ER/PR). HER2 and/or expression was determined by 444 immunohistochemistry with scoring in accordance with American Society of Clinical Oncology (ASCO)/College of American Pathologists (CAP) guidelines<sup>42</sup>. Scores 3+ were 445 446 reported as positive, score 1+/0 as negative (-). Tumors with scores 2+ were further tested by 447 FISH. HER2 gene amplification was defined in accordance with ASCO/CAP guidelines. BC 448 subtypes were defined as follows: tumors positive for either ER or PR, and negative for HER2 449 were classified as luminal; tumors positive for HER2 were considered to be HER2-positive BC; 450 tumors negative for ER, PR, and HER2 were considered to be triple-negative breast cancers 451 (TNBC).

452

## 453 Pathological review

Bulk tumor specimens - and the corresponding pretreatment core needle biopsy specimens in
case of neoadjuvant treatment - were reviewed by an expert in breast pathology (ML). All
tumoral tissues studied were Formalin-Fixed Paraffin-Embedded (FFPE) tissue samples. In

It is made available under a CC-BY-NC-ND 4.0 International license .

457 the cases with bilateral invasive carcinoma, tumor infiltrating lymphocytes (TILs) were 458 reviewed specifically for the purposes of this study, between September 2016 to March 2017. In accordance with the recommendations of the international TILs Working Group<sup>43</sup>, we 459 460 checked for presence of a mononuclear cell infiltrate in the stroma on hematoxylin and eosin-461 stained sections without additional staining, after excluding areas around ductal carcinomas in 462 situ (DCIS), and tumor zones with necrosis and artifacts. Infiltrates were scored on a 463 continuous scale, as the mean percentage of the stromal area occupied by mononuclear cells. 464 Intratumoral TILs (IT TILs) were defined as intraepithelial mononuclear cells within tumor 465 nests or in direct contact with tumor cells and stromal TILs (Str TILs) were defined as 466 mononuclear inflammatory cells within intratumoral stromal area and were reported as 467 percentage of stromal area. After NAC, we assessed TIL levels within the borders of the residual tumor bed, as defined by the RCB and recommended by the TILs working group  $^{44}$ . 468 We simultaneously determined the RCB index, as described by Symmans index <sup>45</sup>, with the 469 470 web-based calculator the freely available via Internet 471 (www.mdanderson.org/breastcancer\_RCB). We defined pathological complete response 472 (pCR) as the absence of invasive residual tumor from both the breast and axillary nodes 473 (ypT0/is N0).

474

# 475 Concordance between the tumors of sBBCs pairs

We evaluated the concordance of the clinical and pathological characteristics between the two
tumors within the same patient. Pairs of sBBCs composed of tumors of the same BC subtypes
were classified as concordant and where otherwise classified as discordant.

479

# 480 Sample's preparation and next generation sequencing (NGS) analyses

481 DNA and RNA were obtained from the Biological Resource Center (BRC of Institut Curie). 482 After selecting patients treated with NAC, tumors from whom sufficient frozen material from 483 both left and right tumors was available in the institutional tissue bank both before (defined as 484 primary tumor (PT)) and after treatment in case of residual disease (RD) were included. 485 Fresh-frozen samples were subjected to genomic DNA extraction and DNA qualification 486 using QuBit system.

487

It is made available under a CC-BY-NC-ND 4.0 International license .

489 lug of genomic DNA from each sample were subjected to shearing using Covaris system and 490 Illumina compatible libraries were performed according to Agilent SureSelect XT2 library 491 protocol consisting in repairing DNA ends and ligating Illumina barcoded adapters followed 492 by a PCR amplification. Libraries were pooled in equimolar condition before being 493 hybridized on dedicated biotinylated RNA probes targeting whole exome sequences (Agilent 494 Human all Exon V5 capture probes). After selection using streptavidin beads and PCR 495 amplification, enriched library pools were subjected to qPCR quantification using the KAPA 496 library quantification kit (Roche). Sequencing was carried out on the HiSeq2000 instrument 497 from Illumina based on a 2 x 100 cycles mode (paired-end reads, 100 bases) using high output 498 flow cells to produce over 50 and 170 million paired-end reads for 30X (germline) and 100X 499 (tumors) respectively. Samples were sequenced to a median depth of coverage of 153 reads, 500 with 95 % of exonic bases passing  $50 \times coverage$ . 2 samples were discarded from the analysis 501 due to low purity (RD4\_R and RD6B\_R). Reads were aligned on the human genome reference hg19/GRCh37 by Burrows-Wheeler Aligner <sup>46</sup>v0.7.5a; filtering of reads based on 502 target intersection, mapping quality and PCR duplicates removal, using Picard<sup>47</sup>, Bedtools<sup>48</sup> 503 and Samtools<sup>49</sup>, and preprocess using GATK<sup>50</sup> for local realignment around indels, and base 504 score recalibration. Preliminary variant calling was performed using Mutect2<sup>51</sup> for tumor 505 samples, and haplotype caller<sup>52</sup> for normal samples. SuperFreq v.  $1.3.2^{53}$  performed 506 507 annotation and filtering of somatic indels and SNVs, copy number and purity estimation, and 508 subclonal reconstruction. SuperFreq was run to analyze together the samples of the same side for each patient. We performed an additional filtering of alterations present in either dbSNP<sup>54</sup> 509 or ExAC<sup>55</sup> at a frequency greater than 0.2 or after manual review of the alignment on the 510 Integrative Genomics Viewer (IGV)<sup>56</sup> or a Somatic Score computed by SuperFreq greater 511 512 than 0.5. Due to insufficient purity, 2 samples were discarded from the analyses (RD4\_R and 513 RD6B R).

- 514
- 515

#### Somatic mutations interpretation

Somatic variants were annotated using VEP (version 104) <sup>57</sup>. A variant was denoted as driver if the mutation was present as either as a splice site, a nonsense, a frameshift or a nonsynonymous SNV or indel in COSMIC census gene<sup>58</sup>. The percentage of shared mutations in pairs (pair left right; pair PT-RD; pair of two multicentric tumors of the same side) was defined as the intersect between the mutations present in both tumors \*2 divided by the sum of the mutation in the two tumors of the pair \*100.

522

It is made available under a CC-BY-NC-ND 4.0 International license .

523

## Mutational signature deconvolution

524

525 The contribution of mutational signatures to individual tumor samples were explored using the signatures deconvoluted by Alexandrov et al.<sup>27</sup> and referenced in the Cosmic database<sup>59</sup>. 526 We restricted the analyses to the 13 signatures previously evidenced in BC. Signature 527 activities were estimated using the decompTumor2Sig algorithm <sup>60</sup> in the musicatk (version 528 1.0.0) R package<sup>61</sup>. The percentage of mutational signatures was calculated by summing the 529 530 relative contribution of each signature in PT samples to the whole tumor spectrum, divided by 531 the number of sampled and the result was multiplied by 100. To avoid overrepresentation of 532 the patient #6 to the cohort and because signatures profiles were similar on each side, we 533 averaged the values of the left side on the one hand, and the values of the right side on the 534 other hand.

- 535
- 536

# RNA sequencing

537 Total RNA extracts from tumor samples were subjected to quality control on a Bioanalyzer 538 instrument and only RNA with RIN (RNA Integrity Number) > 7 were used for sequencing. 539 RNA quantification was achieved using absorbance at 260nm with a Nanodrop 540 spectrophotometer. RNA sequencing libraries were prepared from 1 ug of total RNA using 541 the Illumina TruSeq Stranded mRNA Library preparation kit (following provider's 542 instructions) which allows performing a strand-specific sequencing. Briefly, a first step of 543 polyA+ RNA selection using oligodT magnetic beads is done to focus sequencing on 544 polyadenylated transcripts. After fragmentation, cDNA synthesis was performed and resulting 545 fragments were used for dA-tailing and then ligated to the TruSeq indexed adapters. PCR 546 amplification is finally achieved to create the final cDNA library. After qPCR quantification 547 using the KAPA library quantification kit (Roche), sequencing was carried out using 2 x 100 548 cycles (paired-end reads 100 bases). Sequencing was performed by multiplexing barcoded 549 libraries on with the Illumina HiSeq2000 instrument using high output flow cells to obtain 550 100 million paired-end reads per sample. Alignments were performed on human reference sequences using TopHat<sup>62</sup> v.2.0.621. Reads with mapping quality <20 and reads marked as 551 duplicates by Picard v.1.97 were excluded from further analysis. Gene- level read counts were 552 obtained using HTSeq-count<sup>63</sup> and RefSeq hg19/GRCh37. 553

- 554
- 555

Selection of the genes with the most variant expression and clustering

It is made available under a CC-BY-NC-ND 4.0 International license .

556 We selected the most variant genes, based on the inflection point of the interquartile range 557 (IQR) distribution for gene expression. The gene expression was previously rlog transformed with  $DESeq2^{64}$ . This method is more data-driven than a fixed threshold to define the 558 559 proportion of genes with the highest level of variation. For each gene, we applied the 560 following procedure: (1) we calculated the IQR for all samples; (2) we sorted the IQR values 561 of the genes in ascending order, to generate an ordered distribution; (3) we estimated the 562 major inflection point of the IQR curve as the point on the curve furthest away from a line 563 drawn between the start and end points of the distribution; (4) we retained genes with an IQR 564 higher than the inflection point. Hierarchical clustering was performed with Pearson distance 565 and Ward linkage.

566 567

### Statistical analysis

568 The study population was described in terms of frequencies for qualitative variables, or 569 medians and associated ranges for quantitative variables. Chi-squared tests were performed to 570 search for differences between subgroups for each variable (considered significant for p-571 values  $\leq$  0.05). Continuous variables were compared between groups in Wilcoxon–Mann– 572 Whitney tests for groups of fewer than 30 patients and for variables following multimodal 573 distributions. Student's t-tests were used in all other cases. Pre and post-NAC TIL levels were 574 analyzed as continuous variables. In case of categorical variables, the kappa coefficient <sup>65</sup> 575 was computed as a measure of concordance between the left and the right side of a same 576 patient (varying from - 1 as absolute discordance, to + 1 as absolute concordance, 0 as 577 absence of concordance); otherwise in case of numeric or integer variables, the kendall test 578 was used.

579 All *P*-values no greater than 0.05 were considered significant. In the case where we tested the 580 hypothesis of potentially different effects of the concordant or discordant status of the tumor 581 pair regarding BC subtype on immune infiltration or response to treatment, we included an 582 interaction term in a linear regression model or logistic regression model respectively. A p-583 value of 0.10 was selected to determine the statistical significance of the interaction term, as it has been suggested due to a low power of the test in the interaction setting <sup>66</sup>. When appliable, 584 585 all statistical tests were two-sided. In boxplots, lower and upper bars represent the first and 586 third quartile, respectively, the

587 medium bar is the median, and whiskers extend to 1.5 times the inter-quartile range. Data 588 were processed and statistical analyses were carried out with R software version 3.1.2 589 (www.cran.r-project.org).

It is made available under a CC-BY-NC-ND 4.0 International license .

590 591 Data availability statement 592 The genomic and transcriptomic data generated during the current study are available in the 593 European genome phenome archive repository, [PERSISTENT WEB LINK TO 594 DATASETS]. 595 596 597 CIBERSORT 598 CIBERSORT is an analytical tool quantifying the levels of distinct immune cell types within 599 a complex gene expression mixture (https://cibersort-stanford-edu.proxy.insermbiblio.inist.fr). 600 We applied the original CIBERSORT gene signature LM22 defining 22 immune cell subtypes 601 to all the samples of the cohort, the number of permutations being set to 100, and the mode 602 being set to "absolute". For each immune subpopulation, (i) we calculated the difference 603 between a given sample and the rest of the cohort; (ii) we squared the result; (iii) and we 604 summed the difference between this patient and each other sample of the cohort, resulting in a 605 dissimilarity index. We displayed the overall results on a correlogram. 606

607

## TCR-sequencing analysis.

608 We applied the MixCR algorithm on RNASeq data of 20 samples to identify and quantify TCR Beta-chain CDR3 sequences. MiXCR (version 2.1.5) <sup>67</sup>was used with its default 609 610 parameters to extract and quantify TCR b-chain CDR3 sequences from RNA-seq fastq files. 611 From the MiXCR output, we obtained for each sample the total number of distinct TCRb 612 clones and the number of reads supporting each clone, and we normalized the result by the 613 total number of reads. We investigated if TCR sequences identified by MixCR were identified 614 in the VDJDB database, a curated database of T-cell receptor sequences of known antigen specificity<sup>68</sup> (Details available at the following URL : https://github.com/antigenomics/vdjdb-615 616 db)

617

It is made available under a CC-BY-NC-ND 4.0 International license .

# 618 Acknowledgments

- 619 We thank Corinne Azrijov for collecting patient's consents, Celine Meaudre for the
- 620 management of the samples, Audrey Rapinat for DNA and RNA extraction, Antoine Mader
- 621 for help in designing the figures, Lounes Djerroudi for his advice.

# 622 Author's contributions

- 623 FR and ASHP designed and conceptualized the study.
- 624 LD, MD, EL and ASHP identified the Curie cohort and the patient tumor samples.
- 625 ASHP and LD managed the clinical database, including performing retrospective chart
- 626 review.
- 627 OM, SV, and MA provided patient tumor samples.
- 628 ML provided pathology support and performed pathological review and TILs scoring with TL
- 629 and BG.
- 630 FZ, JA and JH performed literature review.
- 631 IB, JYP provided scientific advice and expert guidance on tumor biology.
- 532 JA, CL, CV, and ASHP performed the whole exome sequencing analyses.
- 633 BS and ASHP performed the RNASeq analyses.
- 634 WR provided guidance for MixCR analyses.
- 635 JW provided scientific direction.
- 636 FR and ASHP allocated funding.
- 637 All the authors participated in the preparation of the manuscript.
- 638 639
- 059
- 640

It is made available under a CC-BY-NC-ND 4.0 International license .

641	References
071	<b>INCIULUIUU</b>

- 642
- 643 1. Chen, Y., Thompson, W., Semenciw, R. & Mao, Y. Epidemiology of contralateral breast
- 644 cancer. *Cancer Epidemiol. Biomarkers Prev.* **8**, 855–861 (1999).
- 645 2. Hartman, M. et al. Incidence and prognosis of synchronous and metachronous bilateral
- 646 breast cancer. J. Clin. Oncol. 25, 4210–4216 (2007).
- 647 3. Vuoto, H. D. *et al.* Bilateral breast carcinoma: clinical characteristics and its impact on
  648 survival. *Breast J* 16, 625–632 (2010).
- 649 4. Sakai, T. et al. National trends of synchronous bilateral breast cancer incidence in the
- 650 United States. Breast Cancer Res. Treat. (2019) doi:10.1007/s10549-019-05363-0.
- 5. Carmichael, A. R., Bendall, S., Lockerbie, L., Prescott, R. & Bates, T. The long-term
- outcome of synchronous bilateral breast cancer is worse than metachronous or unilateral
  tumours. *Eur J Surg Oncol* 28, 388–391 (2002).
- 654 6. Kheirelseid, E. A. H. et al. Bilateral breast cancer: analysis of incidence, outcome,
- 655 survival and disease characteristics. *Breast Cancer Res. Treat.* **126**, 131–140 (2011).
- 656 7. Daly, M. B. et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment:
- Breast, Ovarian, and Pancreatic, Version 1.2020. *J Natl Compr Canc Netw* **18**, 380–391
- 658 (2020).
- 659 8. Brommesson, S. et al. Tiling array-CGH for the assessment of genomic similarities
- among synchronous unilateral and bilateral invasive breast cancer tumor pairs. *BMC Clin*
- 661 *Pathol* **8**, 6 (2008).
- 662 9. Chunder, N., Roy, A., Roychoudhury, S. & Panda, C. K. Molecular study of clonality in
  663 multifocal and bilateral breast tumors. *Pathol Res Pract* 200, 735–741 (2004).
- 10. Imyanitov, E. N. et al. Concordance of allelic imbalance profiles in synchronous and
- 665 metachronous bilateral breast carcinomas. *Int J Cancer* **100**, 557–564 (2002).

It is made available under a CC-BY-NC-ND 4.0 International license .

- 666 11. Janschek, E. et al. Contralateral breast cancer: molecular differentiation between
- 667 metastasis and second primary cancer. *Breast Cancer Res Treat* **67**, 1–8 (2001).
- 12. Shibata, A. *et al.* Clonal analysis of bilateral breast cancer. *Clinical Cancer Research* 2,
- 669 743–748 (1996).
- 670 13. Song, F. et al. Comparative genomic analysis reveals bilateral breast cancers are
- 671 genetically independent. *Oncotarget* **6**, 31820–31829 (2015).
- 672 14. Teixeira, M. R. et al. Assessment of clonal relationships in ipsilateral and bilateral
- 673 multiple breast carcinomas by comparative genomic hybridisation and hierarchical
- 674 clustering analysis. *British Journal of Cancer* **91**, 775–782 (2004).
- 675 15. Noguchi, S. et al. Discrimination between multicentric and multifocal carcinomas of the
- breast through clonal analysis. *Cancer* **74**, 872–877 (1994).
- 16. Stenmark-Askmalm, M., Gentile, M., Wingren, S. & Ståhl, O. Protein accumulation and
- gene mutation of p53 in bilateral breast cancer. South-East Sweden Breast Cancer Group.
- 679 *Acta Oncol* **40**, 56–62 (2001).
- 680 17. Teixeira, M. R. et al. Discrimination between multicentric and multifocal breast
- 681 carcinoma by cytogenetic investigation of macroscopically distinct ipsilateral lesions.
- 682 *Genes Chromosomes Cancer* **18**, 170–174 (1997).
- 18. Ghazani, A. A. et al. Genomic alterations in sporadic synchronous primary breast cancer
- using array and metaphase comparative genomic hybridization. *Neoplasia* **9**, 511–520
- 685 (2007).
- Begg, C. B. *et al.* Contralateral breast cancers: Independent cancers or metastases? *Int. J. Cancer* 142, 347–356 (2018).
- 688 20. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic
- properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61
- 690 (2015).

- 691 21. Loi, S. et al. RAS/MAPK Activation Is Associated with Reduced Tumor-Infiltrating
- 692 Lymphocytes in Triple-Negative Breast Cancer: Therapeutic Cooperation Between MEK
- and PD-1/PD-L1 Immune Checkpoint Inhibitors. *Clin. Cancer Res.* 22, 1499–1509
- 694 (2016).
- 695 22. Fish, E. N. The X-files in immunity: sex-based differences predispose immune responses.
- 696 Nat. Rev. Immunol. 8, 737–744 (2008).
- 697 23. Lazuardi, L. et al. Age-related loss of naïve T cells and dysregulation of T-cell/B-cell

698 interactions in human lymph nodes. *Immunology* **114**, 37–43 (2005).

- 699 24. Denkert, C. et al. Tumor-associated lymphocytes as an independent predictor of response
- to neoadjuvant chemotherapy in breast cancer. J. Clin. Oncol. 28, 105–113 (2010).
- 701 25. Denkert, C. et al. Tumor-infiltrating lymphocytes and response to neoadjuvant
- chemotherapy with or without carboplatin in human epidermal growth factor receptor 2-
- positive and triple-negative primary breast cancers. J. Clin. Oncol. 33, 983–991 (2015).
- 26. Denkert, C. et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of
- breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. Lancet
- 706 *Oncol.* **19**, 40–50 (2018).
- 707 27. Alexandrov, L. B. et al. The repertoire of mutational signatures in human cancer. Nature
- 708 **578**, 94–101 (2020).
- 28. Russnes, H. G. et al. Paired distribution of molecular subtypes in bilateral breast

710 carcinomas. *Cancer Genet* **204**, 96–102 (2011).

- 711 29. de la Rochefordiere, A. et al. Simultaneous bilateral breast carcinomas: a retrospective
- review of 149 cases. International Journal of Radiation Oncology, Biology, Physics 30,

713 35–41 (1994).

It is made available under a CC-BY-NC-ND 4.0 International license .

- 714 30. Kim, H. *et al.* Distribution of tumor subtypes in bilateral breast cancer: Comparison
- between synchronous and metachronous cancer. Asia Pac J Clin Oncol (2020)
- 716 doi:10.1111/ajco.13444.
- 717 31. Chen, I. X. et al. A bilateral tumor model identifies transcriptional programs associated
- 718 with patient response to immune checkpoint blockade. *Proc Natl Acad Sci U S A* **117**,
- 719 23684–23694 (2020).
- 720 32. Pak, L. M. et al. Tumor phenotype and concordance in synchronous bilateral breast
- cancer in young women. *Breast Cancer Res Treat* **186**, 815–821 (2021).
- 722 33. Beinart, G. et al. Clinical course of 771 patients with bilateral breast cancer:
- characteristics associated with overall and recurrence-free survival. *Clin. Breast Cancer* 7,
- 724 867–874 (2007).
- 34. Mosbah, R. *et al.* Pathological Characteristics of Both Tumors in Bifocal and Bicentric
  Breast Cancer. *Anticancer Res* 35, 5111–5116 (2015).
- 727 35. Hartman, M. et al. Genetic implications of bilateral breast cancer: a population based
- 728 cohort study. *Lancet Oncol.* **6**, 377–382 (2005).
- 729 36. Goto, R. et al. The number of FoxP3-positive tumor-infiltrating lymphocytes in patients
- with synchronous bilateral breast cancer. *Breast Cancer* 27, 586–593 (2020).
- 731 37. Reinisch, M. et al. pCR rates in patients with bilateral breast cancer after neoadjuvant
- anthracycline-taxane based-chemotherapy A retrospective pooled analysis of individual
- patients data of four German neoadjuvant trials. Breast (Edinburgh, Scotland) 32, 73–78
- 734 (2017).
- 735 38. Tsunoda, M. et al. T cell receptor (TCR) repertoire analysis reveals a highly conserved
- TCR repertoire in a bilateral tumor mouse model. *bioRxiv* 2021.05.13.443732 (2021)
- 737 doi:10.1101/2021.05.13.443732.

- 738 39. Tarabichi, M. et al. A practical guide to cancer subclonal reconstruction from DNA
- 739 sequencing. *Nat Methods* **18**, 144–155 (2021).
- 40. Fanale, D. et al. Detection of Germline Mutations in a Cohort of 139 Patients with
- 741 Bilateral Breast Cancer by Multi-Gene Panel Testing: Impact of Pathogenic Variants in
- 742 Other Genes beyond BRCA1/2. *Cancers (Basel)* **12**, E2415 (2020).
- 41. GEFPICS-FNCLCC. [Recommendations for the immunohistochemistry of the hormonal
- receptors on paraffin sections in breast cancer. Update 1999. Group for Evaluation of
- 745 Prognostic Factors using Immunohistochemistry in Breast Cancer (GEFPICS-FNCLCC)].
- 746 *Ann Pathol* **19**, 336–343 (1999).
- 42. Wolff, A. C. et al. American Society of Clinical Oncology/College of American
- Pathologists guideline recommendations for human epidermal growth factor receptor 2
- 749 testing in breast cancer. J. Clin. Oncol. 25, 118–145 (2007).
- 43. Salgado, R. *et al.* The evaluation of tumor-infiltrating lymphocytes (TILs) in breast
- cancer: recommendations by an International TILs Working Group 2014. Ann. Oncol. 26,
- 752 259–271 (2015).
- 44. Dieci, M. V. et al. Update on tumor-infiltrating lymphocytes (TILs) in breast cancer,
- 754 including recommendations to assess TILs in residual disease after neoadjuvant therapy
- and in carcinoma in situ: A report of the International Immuno-Oncology Biomarker
- 756 Working Group on Breast Cancer. Semin. Cancer Biol. 52, 16–25 (2018).
- 45. Symmans, W. F. et al. Measurement of residual breast cancer burden to predict survival
- after neoadjuvant chemotherapy. J. Clin. Oncol. 25, 4414–4422 (2007).
- 46. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler
- 760 transform. *Bioinformatics* **25**, 1754–1760 (2009).
- 761 47. Picard Tools By Broad Institute. https://broadinstitute.github.io/picard/.

- 48. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing
- 763 genomic features. *Bioinformatics* **26**, 841–842 (2010).
- 49. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25,
- 765 2078–2079 (2009).
- 50. McKenna, A. et al. The Genome Analysis Toolkit: A MapReduce framework for
- analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297–1303 (2010).
- 51. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and
- heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
- 52. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of
- 771 samples. *bioRxiv* 201178 (2018) doi:10.1101/201178.
- 53. Flensburg, C., Sargeant, T., Oshlack, A. & Majewski, I. J. SuperFreq: Integrated mutation
  detection and clonal tracking in cancer. *PLoS Comput. Biol.* 16, e1007603 (2020).
- 54. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**,
- 775 308–311 (2001).
- 55. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536,
  285–291 (2016).
- 56. Robinson, J. T. et al. Integrative genomics viewer. Nat Biotechnol 29, 24–26 (2011).
- 57. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122
- 780 (2016).
- 58. Sondka, Z. *et al.* The COSMIC Cancer Gene Census: describing genetic dysfunction
- across all human cancers. *Nat Rev Cancer* **18**, 696–705 (2018).
- 59. Forbes, S. A. *et al.* COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* 45, D777–D783 (2017).
- 60. Krüger, S. & Piro, R. M. decompTumor2Sig: identification of mutational signatures
- active in individual tumors. *BMC Bioinformatics* **20**, 152 (2019).

- 787 61. Chevalier, A. et al. The Mutational Signature Comprehensive Analysis Toolkit (musicatk)
- for the discovery, prediction, and exploration of mutational signatures.
- 789 2020.11.17.385864 https://www.biorxiv.org/content/10.1101/2020.11.17.385864v2
- 790 (2021) doi:10.1101/2020.11.17.385864.
- 791 62. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient
- alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25
- 793 (2009).
- 63. Anders, S., Pyl, P. T. & Huber, W. HTSeq A Python framework to work with high-
- throughput sequencing data. *Bioinformatics* **31**, 166–169 (2014).
- 64. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion
- for RNA-Seq data with DESeq2. *bioRxiv* 1–21 (2014) doi:10.1101/002832.
- 798 65. Fleiss J., H. Statistical Methods forRates and Proportions. WILEY SERIES IN
- 799 *PROBABILITY AND STATISTICS* **16**, 539–539 (2003).
- 800 66. Selvin. Statistical analysis of epidemiologic data. (Oxford University Press, 1991).
- 801 67. Bolotin, D. A. *et al.* MiXCR: software for comprehensive adaptive immunity profiling.
- 802 *Nat Methods* **12**, 380–381 (2015).
- 803 68. Bagaev, D. V. et al. VDJdb in 2019: database extension, new analysis infrastructure and a
- T-cell receptor motif compendium. *Nucleic Acids Research* **48**, D1057–D1062 (2020).
- 805 69. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles.
- 806 Nat. Methods 12, 453–457 (2015).
- 807

It is made available under a CC-BY-NC-ND 4.0 International license .

808

## 809 Figures and tables legends

- **Fig1: Tumor's characteristics of the pairs of sBBCs included in the cohort.** Tumor's
- 811 characteristics are based on the 302 pairs with concordance subtype available out of 313 pairs.
- 812 The concordance refers to the status of both tumors within a pair of sBBCs regarding BC
- subtype, either of the same BC subtype (tumor in concordant pair), either of different BC
- 814 subtypes (tumor in discordant pair). A. Repartition of the association of the BC subtypes
- within a pair of sBBCs according to the concordance or the discordance status of the pair; B.
  ER staining intensity of the index tumor by the concordance status of the pair it belongs to
- (luminal tumors only, n=510); C. PR staining intensity of the index tumor by the
- concordance status of the pair it belongs to (luminal tumors only, n=446); D. Str TIL levels of
- the index tumor by BC subtype and by the concordance status of the pair it belongs to; E. IT
- TIL levels of the index tumor by BC subtype and by the concordance status of the pair it
- belongs to; F. pCR rates of the index tumor by BC subtype and by the concordance status of
- the pair it belongs to; G. Evolution of the 100 tumors after NAC according to the BC subtype
- 823 of the index tumor and by the concordance status of the pair of tumor.
- 824 Abbreviations: BC, breast cancer; ER, estrogen receptor; IT TIL, intra tumoral tumor
- 825 infiltrating lymphocytes. NAC, neoadjuvant chemotherapy; pCR, pathologic complete
- 826 response; PR, progesterone receptor; sBBC, synchronous bilateral breast cancers; Str TIL,
- 827 stromal tumor infiltrating lymphocytes; TNBC, triple negative breast cancer
- 828

# 829 Fig2: Tumor mutation profiles

- 830 The analyses are performed in the in left, right, pre and post-NAC samples of a cohort of 6
- patients (20 samples) with next generation sequencing data available. A. Heatmap of somatic
- 832 driver mutations (including missense, nonsense, and splicing) detected in 6 study patients and
- 833 20 samples. B. Venn diagrams showing the number of mutations shared between the left
- 834 (pink) and the right (purple) primary tumors of a same patient; and shared (intersect
- 835 turquoise) between the PT (light green) and the corresponding specimen after NAC (RD,
- residual disease, blue). The post NAC samples RD4\_R and RD6B\_R were discarded due to
- poor sample purity. Mutations from the 2 multicentric tumors of each side of patient #6 havebeen merged.
- 839 Abbreviations: L, left; NAC, neoadjuvant chemotherapy; pCR, pathologic complete response;
- 840 PT, primary tumor; R, right; RD, residual disease; TNBC, triple negative breast cancer
- 841

# Fig3: Copy number alterations (CNA) and mutational signatures profiles.

- The analyses are performed in the in left, right, pre and post-NAC samples of a cohort of 6 patients (20 samples) with next generation sequencing data available. Mutational signatures
- published by Alexandrov<sup>27</sup> are calculated using the deconstructSigs package. A. CNA are
- solution published by Filehandrov and calculated using the deconstructions patentic. Filehandrov and calculated using the deconstructions patentic structure and structure
- 847 primary tumor and its corresponding sample with RD after NAC.
- 848 C. Mutational signatures are compared among the two sides of the primary tumor of the same
- patient (Mutations from the 2 multicentric tumors of each side of patient #6 have been
- 850 merged.) and; D. Among the primary tumor and its corresponding sample with RD after851 NAC.
- 852 Abbreviations: BC, breast cancer; CAN, copy number alterations; CH-LOH, copy-neutral loss
- 853 of heterozygosity; chr, chromosome; L, left; NAC, neoadjuvant chemotherapy; R, right; RD,
- 854 residual disease
- 855

It is made available under a CC-BY-NC-ND 4.0 International license .

# 856 Fig4: Fishplot retracing phylogeny between left, right primary tumors and

# 857 corresponding residual disease.

858 Each subfigure represents the evolution of the tumors of a given patient under NAC. The

upper fish plot represents the evolution of the left tumor(s), the lower fishplot represents the

860 evolution of the right tumor(s). Each fishplot displays the prevalence of subclones throughout

treatment. Subclonal architecture was reconstructed with SuperFreq. Subclonal profiles show

- annotated common driver cancer genes. A. Both PTs were mutated for P53, but the genomic
- alteration was different on the left and the right side (right side: indel frameshift deletion
- position 7578213; left side: substitution C>T p.R175C missense substitution identified as
- pathogenic in Clinvar, and present in the RD); B. Both tumors were mutated for TP53 with 2 different mutations Both tumors were mutated for TP53 with 2 different mutations (right side
- different mutations Both tumors were mutated for TP53 with 2 different mutations (right side:
   frameshift loss of a nucleotide position 7577558; E. TP53 was mutated on both sides (left
- side: frameshift deletion position 7578213, right side: frameshift deletion position 7579522);
- Abbreviations: NAC, neoadjuvant chemotherapy; PT, primary tumor; RD, residual disease.
- 870

# Fig5: Genomic alterations in the patient with multicentric sBBCs.

- The analyses are performed in the in two left tumors ( $PT6A_L PT6B_L$ ), and in the 2 right tumors ( $PT6A_R - PT6B_R$ ). The sample with residual disease ( $RD6B_R$ ) was discarded due
- to poor sample purity. A. Tumor mutation profiles: Venn diagrams showing the number of
- 875 mutations shared between the 2 primary tumors of each side (yellow intersect); B. Copy
- number alterations (CNA) profiles, compared among the two multicentric tumors of each
- 877 side; C. Mutational signatures as from Alexandrov<sup>27</sup> compared among the two multicentric
- tumors of each side; D. Fishplot retracing the phylogeny between the two multicentric tumors
- of each side. Each fish plot can be interpreted from the left to the right, or from the right to the
- left, corresponding to the emergence or the extinction of a clone respectively.
- 881 Abbreviations: L, left; PT, primary tumor; R, right; RD, residual disease; BC, breast cancer;
- CAN, copy number alterations; CH-LOH, copy-neutral loss of heterozygosity; chr,chromosome.
- 884

# Fig6: Repartition of the shared clonal TCR repertoire assessed by TCR Beta-chain CDR3 sequences among samples at the individual patient level.

- The clonotypes are identified with the MixCR algorithm<sup>67</sup>. Only the TCR sequences shared
- within different samples of a same patient are displayed on the figure. Unique clonotypes (a, 5004) and also at the same displayed of the same transformed and the same displayed on the figure.
- (n=5994) and clonotypes shared between patients of the cohort (n=32) are not displayed on the former The energy in the 14 PT of the
- the figure. The upper panel represents the repartition of the clonotypes in the 14 PT of the
- patients #1 to #6, and the lower panel represents the repartition of clonotypes of the 6 sampleswith RD.
- Abbreviations: L, left; PT, primary tumor; R, right; RD, residual disease TCR: T cell receptor.
- 894 895







Ε IT TIL levels

G



luminal (n=218)

Str TIL levels

D



Pinteraction BC subtype and concordance status on str TILS=0.57

F Response to NAC



Pinteraction between BC subtype and concordance status=0.07

Pinteraction BC subtype and concordance status on IT TILS=0.0006

Evolution of tumors after NAC by the BC subtype 100 sBBCs (50 patients) treated with NAC



Α



























SBS2

SBS5



D

SBS41

SBS40



SBS9

SBS17a

SBS18







