 doi:10.20944/preprints202107.0430.v1

Article

In BRCA1 and BRCA2 breast cancers, chromosome breaks occur near herpes tumor virus sequences

Bernard Friedenson [1,*]

[1] Dept. of Biochemistry and Molecular Genetics, College of Medicine University of Illinois Chicago; bernief@UIC.edu

* Correspondence: bernief@UIC.edu

Abstract: Inherited mutations in BRCA1 and BRCA2 genes increase risks for breast, ovarian, and other cancers. Both genes encode proteins for accurately repairing chromosome breaks. If mutations inactivate this function, broken chromosomes may not be restored correctly, allowing breaks to persist or rearrange chromosomes. These abnormalities are potentially catastrophic events that can originate from viral infections. I used bioinformatic analyses of publicly available breast cancer patient data to show that the distribution of chromosome breaks in hereditary breast cancers differs markedly from sporadic breast cancers. Then I tested hereditary breast cancer sequence data around chromosome breaks for DNA similarity to all known viruses. Human DNA flanking breakpoints usually had decisive matches to Epstein-Barr virus (EBV / HHV4) tumor variants HKHD40 and HKNPC60. Many breakpoints were near EBV genome anchor sites, human EBV tumor-like sequences, EBV-associated epigenetic marks, and some fragile sites. On chromosomes 2 and 12, sequences near EBV genome anchor sites accounted for 90% and 88% of breakpoints ($p<0.0001$), respectively. On chromosome 4, 51/52 inter-chromosomal breakpoints were close to EBV-like sequences in 19 hereditary breast cancers. In contrast, 19 sporadic breast cancers only had 12 interchromosomal breakpoint regions on chromosome 4 near EBV-like sequences. On various other chromosomes, five EBV genome anchor sites were near hereditary breast cancer breakpoints at precisely defined, disparate gene or LINE locations. Independent evidence further implicating EBV in hereditary breast cancer breakpoints is that 25 breast cancer break positions are within 1.25% of breakpoints in model EBV cancers. In addition to BRCA1 or BRCA2 mutations, all the hereditary breast cancers had mutated genes essential for immune responses. This compromise facilitates reactivation of herpes viruses which produce nucleases capable of breaking chromosomes. EBV also causes other deleterious effects: anchored EBV episomes can interfere with normal replication and obstruct DNA break repairs; even very early infection causes massive transcription changes. The results, therefore, imply proactive treatment and prevention of herpes viral infections may prevent some chromosome breaks and benefit BRCA mutation carriers.

Keywords: Breast cancer infection, breast cancer immunity, breast cancer virus, nasopharyngeal cancer, EBV cancer, hereditary breast cancer, BRCA1, BRCA2.

 doi:10.20944/preprints202107.0430.v1

1. Introduction

Inherited mutations in *BRCA1* and *BRCA2* genes increase risks for breast, ovarian and other cancers. As one of multiple functions, the two genes encode proteins that work with many others to restore broken DNA by homologous recombination. Homologous recombination uses the sister chromatid as a template in a complex process for high-fidelity repair of double-strand breaks in chromosomal DNA. Mutations that inactivate either the *BRCA1* or *BRCA2* gene force DNA break repairs into less accurate, error-producing pathways that do not use a template [1]. Without template guidance during repairs, chromosomes can undergo visible insertions, deletions, and rearrangements because large DNA fragments do not reattach at their original positions [2-4]. BRCA1 also participates in the S-phase and G2/M checkpoints [5-7], so signals from broken chromosomes become unable to pause cell replication to allow repairs. *BRCA2* may also participate in G2/M checkpoint control [7].

Chromosomes in hereditary breast cancer have multiple deletions and insertions with varying interconnections at multiple different breakpoints. The topography of these signatures is sensitive to the status of replication, chromatin organization, and transcription, [8-12]. For example, a single error during cell replication forms a bridge between sister chromatids, causing chromosome breakage and then a cascade of chromothripsis and other mutational events [13]. According to these models a catastrophic breakage event in a single cell can destabilize the entire human genome and generate many complex rearrangements. This scenario is one of four chromosome breakage models [14]. Nik-Zainal and colleagues report hundreds of DNA rearrangements and characteristic rearrangement signatures in breast cancers [10, 12, 15-17]. About six of these breast cancer rearrangement signatures have some relationship to homologous recombination defects [17].

Infections can profoundly affect the processes instrumental in producing a cancer associated rearrangement, so I wondered whether infections underlie some abnormal rearrangement signatures. Are infections a common factor that causes a cascade of cancer associated chromosome rearrangements? The main reason for asking this question is that if infection DNA can generate fundamental chromosome changes in cancer, it is conceivable both to treat the infections and to produce vaccines against hereditary cancer.

Several candidate infections or infection-related species include retroviruses [18] or the latent EBV infections found in 90-95% of humans. About 8% of human DNA probably originated from retroviruses [19-22]. Both endogenous and exogenous retroviruses can cause deletions or insertions in human chromosomes. EBV is also ubiquitous [23] and associates with a diverse group of human malignancies, including nasopharyngeal cancers (NPC), lymphomas, Hodgkin's disease, gastric cancer, and lymphoproliferative disorders. The strongest association is with undifferentiated nasopharyngeal cancer [24-26]. Unlike retroviruses, EBV does not have an integrase enzyme, and integration sequences are short, creating uncertainty in identifying and assigning them [18, 27].

Mutations or downregulation in DNA repair genes linked to *BRCA1-BRCA2* mediated repair pathways compromise immunity [28] and are common in NPC [29, 30]. NPC in human epithelial cells includes

 doi:10.20944/preprints202107.0430.v1

inappropriate DNA repairs causing gene fusions at DNA breakpoints, such as *YAP1-MAML2, PTPLB-RSRC1*, and *SP3-PTK2* [31].

The same EBV-sensitive, BRCA-related pathway related to NPC is also essential to prevent hereditary breast cancers, suggesting a role for EBV infection in hereditary breast cancers. Evidence exists supporting this relationship. Epidemiological associations between breast cancer and EBV infection exist in different geographical locations [32, 33]. EBV infection of breast epithelial cell models facilitates malignant transformation and tumor formation [34]. Breast cancer cells from biopsies express gene products from latent EBV infection (LMP-1, -2, EBNA-, and EBER) [35] even after excluding the possibility that the virus comes from lymphocytes [36]. Evidence for the EBV lytic form in breast cancer associates with a worse outcome [37].

BRCA mutations cause problems and unreliability in repairing complex DNA damage associated with viral infections [38]. These difficulties are concerning because chromosome breaks in hereditary breast cancers occur in the context of nearly universal human infection by EBV and retroviruses. The present study aimed to determine whether such ubiquitous human viral infections are in fact associated with chromosome breaks in hereditary breast cancers.

2. Materials and Methods

Breast cancer genomic sequences

Characteristics of hereditary breast cancers compared to viral cancers. The selection of breast cancer genomes for this study required patient samples with a known, typed *BRCA1* or *BRCA2* gene mutation. Breast cancer genome sequences were from the COSMIC database curated from original publications [12, 16]. 15/25 breast cancers were stage III, four were stage II, and three had no data. Six breast cancers had a typed germ-line *BRCA1* mutation, and nineteen had a typed germ-line *BRCA2* mutation. The mean age at diagnosis was 46, with a maximum of 67 and a minimum of 34. Fourteen patients were alive, and 6 patients were deceased at the time of sequence analyses. Genome sequencing was done before treatment began. Blood samples provided normal genes for comparison [12]. All 25 cancers were female ductal breast cancers.

The 25 ductal breast cancers contained many DNA missense mutations, with an average of 1.55 mutations per gene analyzed (range 1.0-2.75). A total of 275,730 mutations had a mean value of 11,029 per cancer, ranging from 1-2.75 per gene examined. The 4316 DNA breakpoints varied from 33 to 396 per different individual cancer with a mean of 173. For all 25 breast cancers, chromosomes 1 and 2 were the most frequent sites of intra-chromosome rearrangements, but the distributions of DNA breakpoints among the various chromosomes were markedly different (Table 1 and Fig. 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

Table 1 BRCA1 and BRCA2 associated breast cancers studied

Sample	*BRCA1* / *BRCA2* muta-tion status	Analyze d genes	Muta-tions	Breaks	Muta-tions per gene	Muta-tions per break	Chrom -osome with most breaks "From"	Chrom osome with most breaks "To"	Most often intra-chromo somal breaks
PD3890	*BRCA1*	5978	13401	269	2.24	49.8	3=4,1	11, 16,21	1
PD3904	*BRCA1*	6337	13559	247	2.14	54.9	6	17,18	8
PD3945	*BRCA1*	9203	23132	114	2.51	202.9	1	X	2
PD4005	*BRCA1*	6943	15093	185	2.17	81.6	3	5	1
PD4006	*BRCA1* R1835X	7433	20453	396	2.75	51.6	1 (only 1)	22,16	2 (387 intra)
PD4115*	*BRCA1*	9069	22662	180	2.50	125.9	2	17	8
PD4116	*BRCA2*	7831	19958	359	2.55	55.6	11	17	3
PD4836	*BRCA2*	4782	5381	80	1.13	67.3	10,11 ,12	18,20	10
PD4872	*BRCA2*	5095	5906	90	1.16	65.6	2	18	13
PD4874	*BRCA2*	8201	10668	226	1.30	47.2	6	20	10
PD4875	*BRCA2*	6262	8056	140	1.29	57.5	5	19	7
PD4876	*BRCA2*	5718	8317	95	1.45	87.5	5	9	12
PD4951	*BRCA2*	3178	3283	33	1.03	99.5	6	8	4
PD4952	*BRCA2*	10806	15155	332	1.40	45.6	3	14	14
PD4953	*BRCA2*	6690	8090	143	1.21	56.6	4	18	18
PD4954	*BRCA2*	5904	6605	126	1.12	52.4	1	8	2
PD4955	*BRCA2*	6654	7920	100	1.19	79.2	12	15	1
PD4956	*BRCA2*	11036	14860	238	1.35	62.4	7	10	3
PD4957	*BRCA2*	3521	3520	74	1.00	47.6	7=10	17=11	8
PD4958	*BRCA2*	8244	10956	102	1.33	107.4	1	1	4
PD6406	*BRCA2*	4976	5485	191	1.10	28.7	2,4	6,8	2
PD6416	*BRCA2*	3364	3609	71	1.07	50.8	1	20	7
PD7217	*BRCA2*	8841	11215	124	1.27	90.4	3	18	3
PD8621	*BRCA2*	7964	9617	275	1.21	35.0	1	X	1
PD8969	*BRCA2*	7335	8829	126	1.20	70.1	1	8	9

*Uncertain significance variant

The genes analyzed in *BRCA1*-associated breast cancers (Table 1) were on average almost twice as likely to be mutated as the genes analyzed in *BRCA2* associated breast cancers ($p<0.0001$). Moreover, an unpaired t-test assuming equal variances found that the means were different (*BRCA1*=94.5, *BRCA2*=63.5, $p=0.03$. Some tests added breast cancer genome data from 29 presumptive mutation carriers when the number of identified *BRCA1*,2 gene mutation carriers became limiting. Their breast cancers occurred under age 36, but the patients had not been tested for BRCA mutations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

DNA sequence data from sporadic breast cancers. Table 2 shows characteristics of 19 sporadic breast cancers diagnosed at age >=70 and randomly selected from a list [12]. Sporadic cancers in Table 2 were used for comparisons to 19 BRCA2 associated hereditary cancers. The Table includes 18 Stage II and III ductal breast cancers and one lobular breast cancer occurring at age >=70. Numbers of breaks varied from 1 to 263, but were generally lower than the hereditary cancers. Mutations per gene were lower in sporadic cancers with a mean value of 1.09[1.04-1.14]) vs BRA2 associated breast cancers: 1.28[1.12 to 1.44]. Each break accompanied more mutation in sporadic breast cancers with a mean value = 220[-53 to 494] vs. 63[53 to 74].

Table 2. Sporadic breast cancers.

Sporadic breast cancer	age diagnosis / follow-up	Status	Stage	Analyzed genes	Mutations	Breaks	Muta-tions per gene	Muta-tions per break
PD11744	*80/?*	*remission*	II	2548	2548	1	1.00	2548.0
PD8828	74/78	*remission*	III	2711	2776	13	1.02	213.5
PD11386	74/?	*remission*	III	4120	4187	16	1.02	261.7
PD11762	74/76	*remission*	II	2315	2139	17	0.92	125.8
PD18769	*78/80*	*remission*	III	3516	3717	21	1.06	177.0
PD13767	74/78	*progression*	III	3179	3183	25	1.00	127.3
PD7220	73/74	*progression*	III	3495	3722	40	1.06	93.1
PD13428	74/75	*remission*	III	4082	4160	49	1.02	84.9
PD8617	*76/77*	*progression*	II	2439	2428	61	1.00	39.8
PD6044 (lobular)	76/78	*progression*	II	2684	2869	63	1.07	45.5
PD11374	76/?	*remission*	III	3191	3103	70	0.97	44.3
PD23559	74/74	*progression*	III	5809	6561	88	1.13	74.6
PD4959	*73/79*	*progression*	III	5016	5579	103	1.11	54.2
PD11367	*>80/?*	*progression/ d*	III	7905	9756	107	1.23	91.2
PD13765	76/76	*remission*	III	5158	6503	108	1.26	60.2
PD13620	*80/82*	*progression*	III	5668	6492	156	1.15	41.6
PD11388	75/?	*remission*	II	4670	5515	158	1.18	34.9
PD11752	*70/72*	*remission*	III	6579	8625	250	1.31	34.5
PD7215	76/78	*remission*	III	6565	7728	263	1.18	29.4

d=deceased

Hereditary and sporadic breast cancer patient DNA sequence data. Gene breakpoints for inter-chromosomal and intra-chromosomal translocations were obtained from the COSMIC catalog of somatic mutations as curated from original publications [12] and converted to the GrCH38 human genome version. DNA flanking sequences at breakpoints were

 doi:10.20944/preprints202107.0430.v1

downloaded primarily using the UCSC genome browser but did not differ from sequences obtained using the Ensembl genome browser.

Fragile site sequence data. Positions of fragile sites were from a database [39] and original publications [38]. The presence of repetitive di- and trinucleotides was used as a test for the exact positions of fragile sites. "RepeatAround" tested sequences surrounding breakpoints for 50 or fewer direct repeats, inverted repeats, mirror repeats, and complementary repeats.

Comparisons of DNA sequences. The NCBI BLAST program (MegaBLAST) and database [40, 41] compared DNA sequences around breakpoints in *BRCA1*-and *BRCA2*- mutation-positive breast cancers to all available viral DNA sequences. E(expect) values <1e-10 were considered to represent significant homology. In many cases, expect values were 0 and always far below 1e-10. Virus DNA was from BLAST searches using "viruses (taxid:10239)" with homo sapiens and uncharacterized sample mixtures excluded. EBV DNA binding locations on human chromosomes were obtained from publications [42-44], from databases, by interpolating published figures, or by determining the location of genes within EBNA1 binding sites. EBNA1 binding data was based on lymphoblastoid and nasopharyngeal cancer cell lines. When necessary, genome coordinates were all converted to the GrCH38 version. Breaks in hereditary breast cancers were compared to EBV DNA binding sites, epigenetic marks on chromatin, genes, and copy number variations. The MIT Integrated Genome Viewer (IGV) with ENCODE data loaded and from the UCSC genome browser provided locations of H3K9Me3 chromatin epigenetic modifications. The ENCODE website also provided positions of H3K9Me3 marks (www.ENCODEproject.org).

Homology among viruses was determined by the method of Needleman and Wunsch [45].

Data analyses. Microsoft Excel, OriginPro, StatsDirect, Visual basic, and Python scripts provided data analysis. Excel worksheets were often imported into Python Jupyter notebooks for extended analysis. Chromosome annotation software was from the NCBI Genome Decoration page and the Ritchie lab using the standard algorithm for spacing [46]. Statistical analyses used StatsDirect statistical software. Linear correlation, Kendall, and Spearman tests compared distributions of the same numbers of chromosome locations that matched viral DNA. Because the comparisons require the same numbers of sites, comparisons truncated data down to a minimum value of maximum homology (human DNA vs. viral DNA) of at least 400. Excel compared the positions of breast cancers vs. midpoints of genes containing repetitive DNA fragile sites on the same chromosome.

Genes associated with the immune response damaged in breast cancer. Breast cancer somatic mutations in genes were compared to genes in the immune metagenome [47-50]. Sets of genes involved in immune responses also mediate other functions and represent a vast and growing dataset (geneontology.org). Genes involved in cancer control by immune surveillance and immunoediting are not well characterized. An extensive validation included both direct and indirect effects of gene mutations. In addition, the Online Mendelian Inheritance in Man database (www.OMIM.org) was routinely consulted to determine gene function with frequent further support obtained through PubMed, Google scholar,

 doi:10.20944/preprints202107.0430.v1

GeneCards, and UniProtKB. The "interferome" was also sometimes used [www. interferome.org].

3. Results

Inter-chromosomal breakpoint distribution in *BRCA2*-Associated breast cancers.

There are significant differences in where breakpoints involved in inter-chromosomal translocations occur among 19 *BRCA2*-associated breast cancers (Fig. 1). Many studies have shown that translocation partners are limited by spatial organization in the nucleus. Inter-chromosomal breakpoints tend to cluster in specific chromosome regions for individual breast cancers and do not appear to be random. Breast cancer PD4956 has a large breakpoint cluster on chromosome 11, and PD4952 has a group on chromosome 14. A few chromosome areas rearrange in only one cancer. Large numbers of base pairs often separate the breakpoint positions.

Sporadic breast cancer patients are generally older than hereditary breast cancer patients, so mutations have had more time to accumulate. There is a positive correlation between some base substitution signatures and age [51]. The numbers of mutations in sporadic cancers in Table 2 correlate well with the numbers of breaks ($p<0.001$). Yet, inter-chromosomal translocation breakpoints are less frequent in sporadic cancers than in hereditary breast cancers (Table 1 vs. Table 2 and Fig. 1). Telomeric regions are especially noticeable as sites for inter-chromosome translocation on 17 of 23 sporadic breast cancer chromosomes. to rule out the possibility of some hidden bias in selecting sporadic breast cancers, a second set of 19 was chosen at random. Again sporadic breast cancers had fewer breakpoints involved in inter-chromosomal translocation than hereditary cancers and results are consistent with Fig.1. Supplementary Fig S1).

Considering both intra- and inter-chromosome breakpoints in hereditary vs. sporadic breast cancers reduces differences in frequency and distribution. Although many intra-chromosomal breaks are consistent with chromothripsis [52], identifying them remains challenging.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

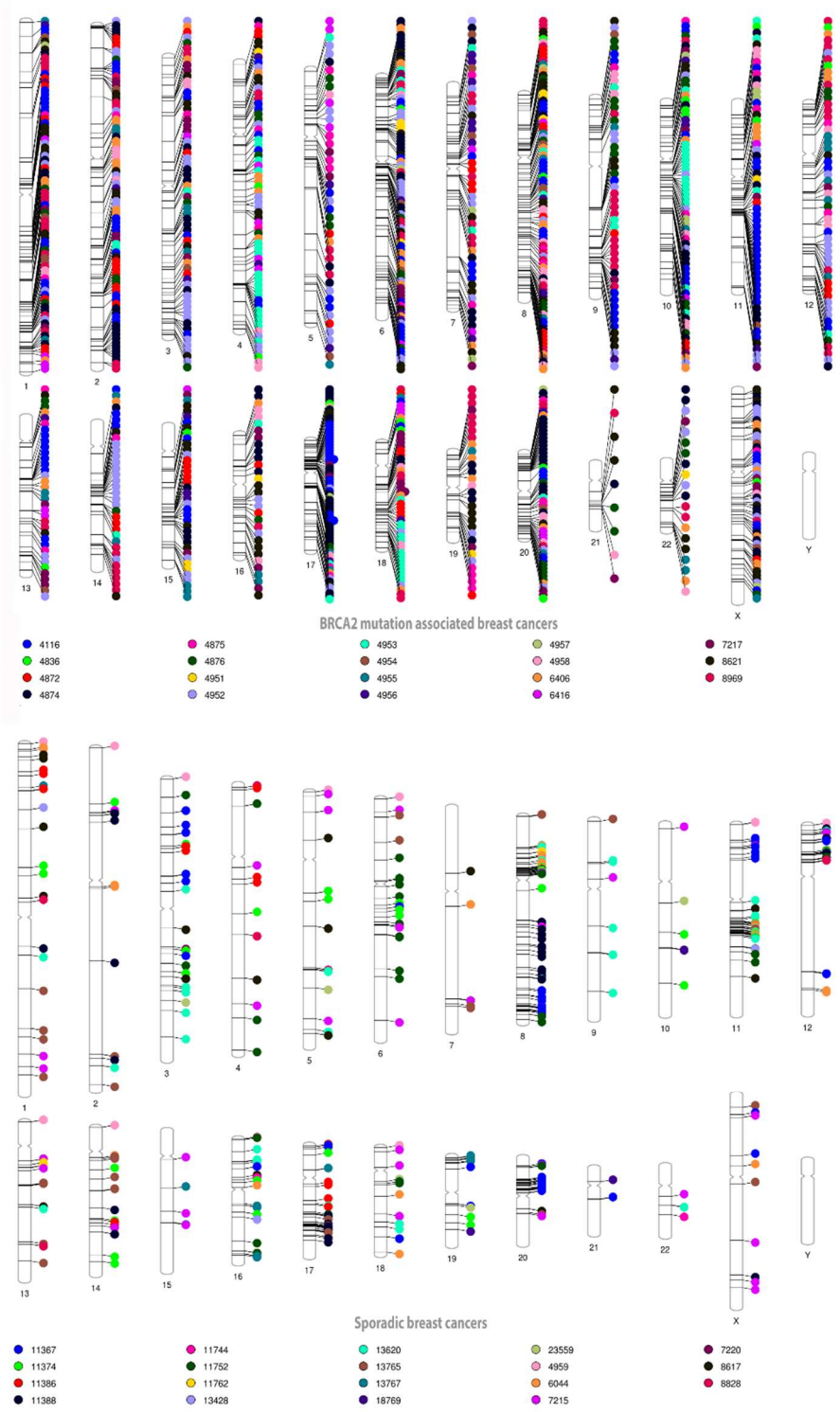


Figure 1. Distribution of breakpoints in inter-chromosome translocations in BRCA2 mutation-associated breast cancers (top) vs. sporadic breast cancers (bottom). Each breast cancer (Tables 1 and 2) has a four-digit number and a different color within each panel. All patients were female, so there are no Y chromosomes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

Inter- and intra-chromosome breakpoint comparisons in breast cancers: mutation carriers vs. women with normal BRCA genes.

Total inter-and intra-chromosome breakpoints distribute very differently in hereditary vs. sporadic breast cancers. Chromosomes 8, 4, 2, 12, 11 and 1 show these significant differences in combined inter- and intra-chromosomal breakpoint plots (Fig. 2). BRCA2 mutation carriers occasionally have non-templated insertion sequences, but none of 19 sporadic breast cancer samples had a non-templated insertion sequence. Substantial clustering of breakpoints occurs on chromosome 11, between 60 - 80 million base pairs and perhaps on some of the other chromosomes shown (Fig.2). This clustering is characteristic of breakage-fusion-bridge cycles [52].

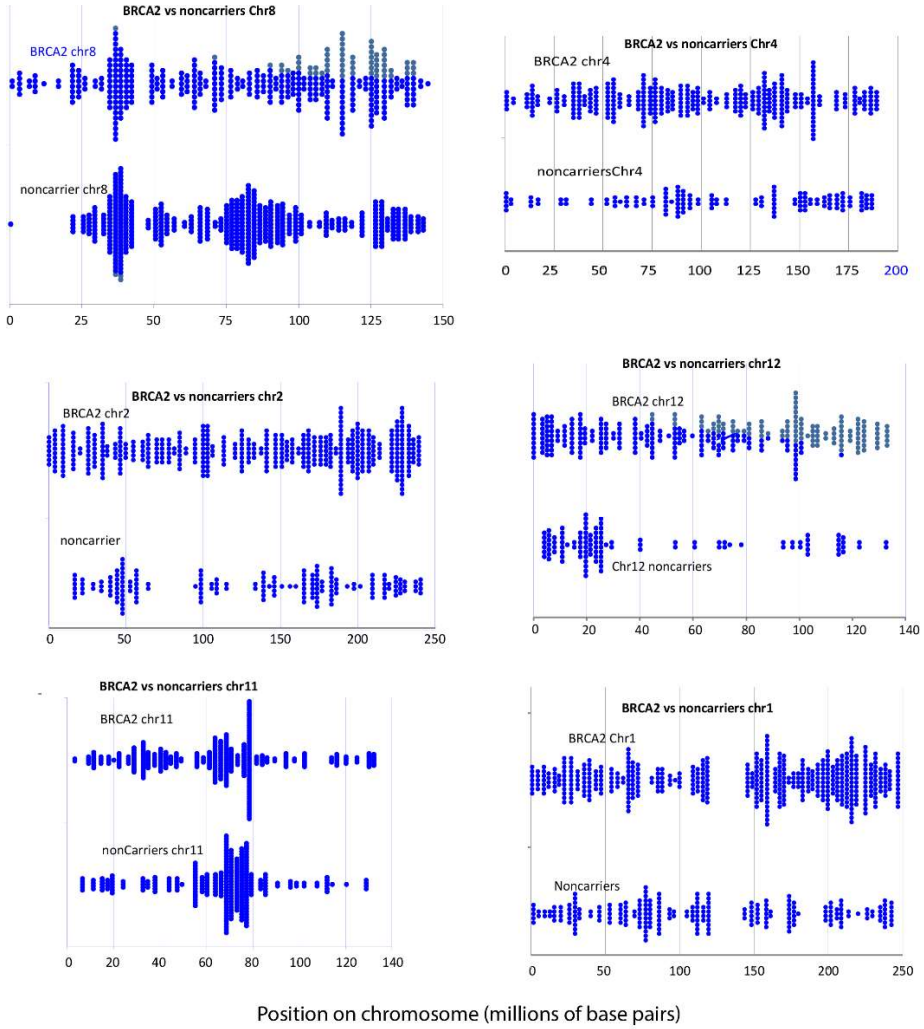


Figure 2. Distributions of inter- and intra-chromosomal breakpoints on chromosomes 8,4,2, 12,11, and 1 in nineteen hereditary BRCA2 breast cancers vs. nineteen sporadic cancers.

Virus-human homology comparisons around inter-chromosomal breakpoints produced many results like those shown in Fig. 3. Many DNA segments are virtually identical to EBV variants (Human gamma-herpesvirus 4 variants, HKNPC60 or HKHD40) at or near the different inter-chromosomal

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

breakpoints shown. A deleted fragment in panel (C) includes virus-like sequences and has viral matching sequences in its flanking regions. Maximum homology scores for human DNA vs. herpes viral DNA are over 4000 for breast cancer PD3945 and just under 4000 for PD4874. These scores represent 97% identity for up to 2462 base pairs, with E "expect" values (essentially p-values) equal to 0. The extensive homology thus represents an EBV-like breakpoint signature.

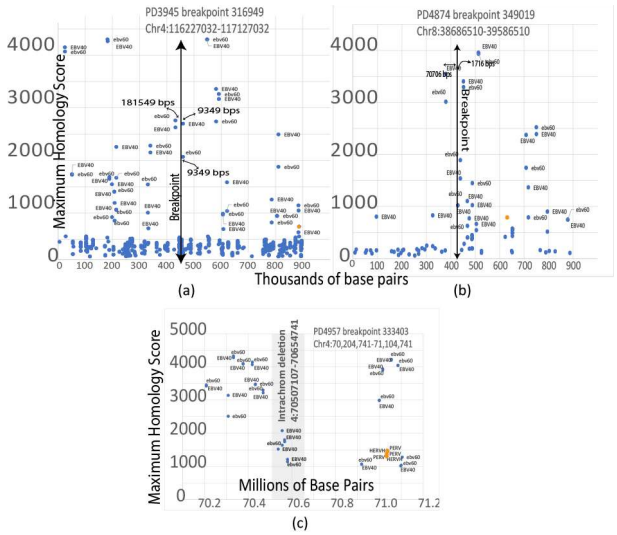


Figure 3. Examples of viral homologies at breast cancer breakpoints. The blue dots represent the start points of human DNA sequences that match EBV tumor variants, and the orange dots are retroviruses. The numbers near breakpoints are the distances in base pairs from the homology start points to the breakpoint. (a) A breakpoint on a BRCA1-associated breast cancer PD3945 is surrounded by DNA homologous to HKHD40 (EBV40) and HKNPC60 (ebv60). (b) A breakpoint in the BRCA2-associated breast cancer PD4874 is also near regions of homology to HKHD40 and HKNPC60. In the bottom panel (c), an intra-chromosome deletion on chromosome 4 (gray area) in BRCA2 associated breast cancer (PD4957) is near EBV40- and ebv60-like human sequences and also includes them.

In hereditary breast cancers, breakpoint flanking sequences often resemble EBV tumor variants.

EBV genomes can dock at many sites on chromosome 4, arbitrarily selected as a representative chromosome to estimate how often hereditary cancer breakpoints are near human sequences that resemble herpes viruses. Nearly all (51/52) inter-chromosomal breakpoints in hereditary breast cancers had statistically significant homology to herpes viruses within about 200k base pairs from the breakpoint (Fig. 4). In contrast, the set of sporadic breast cancers had only about 12 discrete areas involved in inter-chromosomal breakpoints. Breakpoints in the sporadic breast cancers had greater similarities to retroviruses, including porcine endogenous retroviruses and human endogenous retroviruses (Fig. 4). There is also similarity to the Sars-CoV-2 virus.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

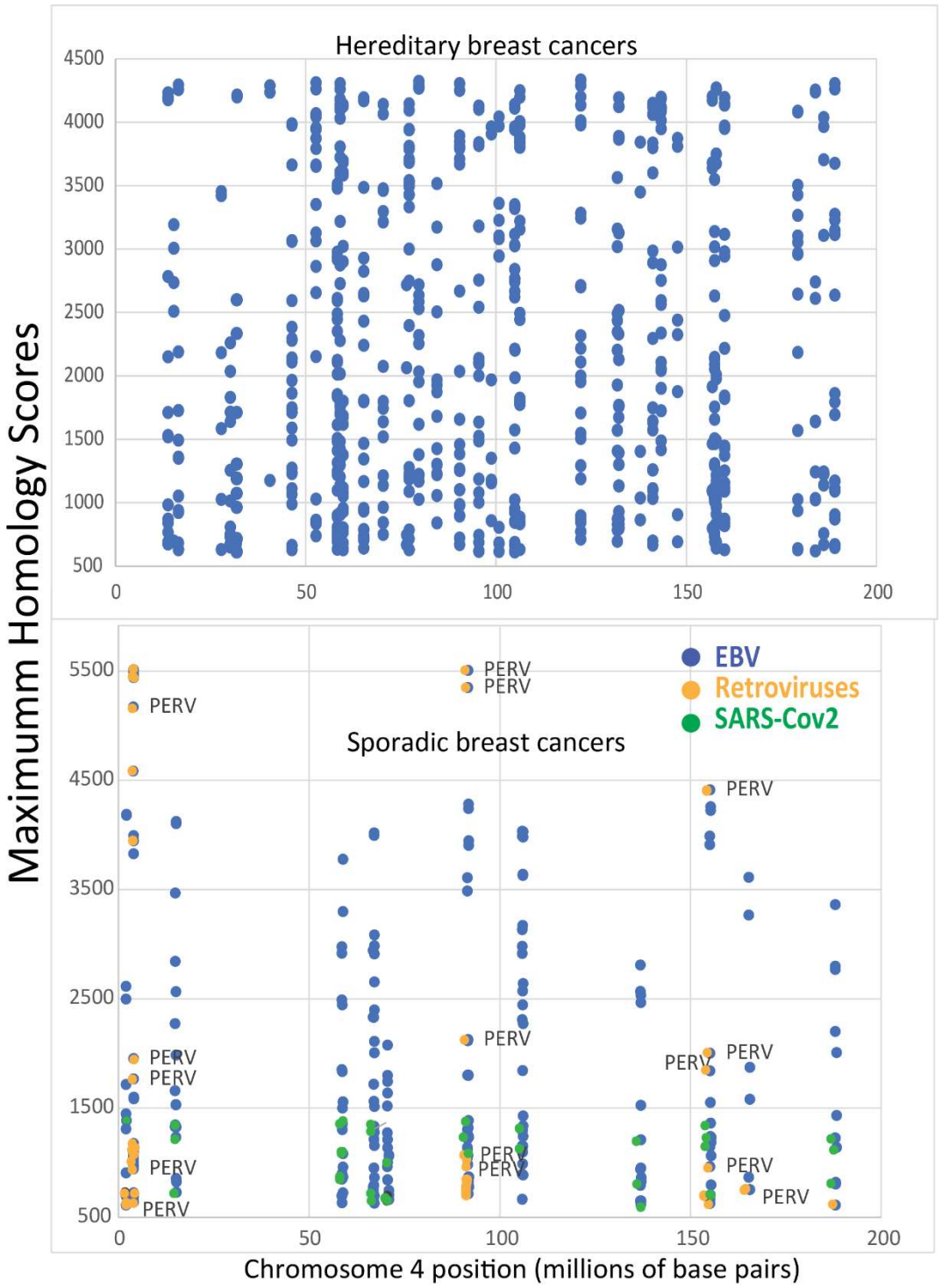


Figure 4. Statistically significant (E<<1e-10) maximum homology scores (blue dots) for human DNA from hereditary (top) and sporadic breast cancers (bottom) vs. all viruses at inter-chromosome breakpoints on the entire length of chromosome 4. For clarity and to allow labeling of the regions matching viruses, only maximum homology scores above 500 are shown.

 doi:10.20944/preprints202107.0430.v1

Hereditary breast cancer breakpoint homologies to EBV are near known EBV genome anchor sites: global comparisons on two chromosomes.

Most breakpoints on chromosomes 2 and 12 are near EBV genome anchor sites. On a 21 Mb section of chromosome 2, (Fig. 5a), the R^2 statistic was calculated to estimate correlation between EBV genome anchoring sites and breast cancer breakpoints positions. According to R^2, EBV anchor sites account for about 91% of chromosome 2 breakpoints ($p<0001$). The normal plot of the residuals from this analysis is approximately linear, suggesting that the residuals are distributed relatively randomly about zero. Breakpoints on a 14 Mb section of chromosome 12 also correlate with viral genome anchor sites ($r^2 = 0.88$, $p<0.0001$) and share many features with chromosome 2 breakpoint region (Figs. 5b vs. 5a).

All breakpoints on both chromosome 2 and 12 sections are near regions of human DNA sequence homology to the EBV variant tumor viruses HKNPC60 and HKHD40. Breakpoints are all accessible as indicated by DNase hypersensitivity, and many breakpoints disrupt gene regulation, gene interaction, and transcription. Many breaks affect cancer-associated (COSMIC) genes. Breakpoints all go through ENCODE candidate cis-regulatory elements (cCREs). Some breakpoints disrupt the epigenetic stimulator H3K27Ac, an enhancer mark on histone packaging proteins associated with increased transcription. Most breast cancer breakpoints are near inhibitory epigenetic H3K9Me3 peaks in CD14+ primary monocytes (RO-01946). These markings occur around EBV genome anchor sites, where they contribute to viral latency and repress transcription [43]. Both regions on chromosomes 2 and 12 are rich in these sites (Figs. 5a and 5b). Both chromosome 2 and 12 sections appear to be a focus for structural variation such as CNV's, inversions, and short insertion/deletions.

Multiple breakpoints on either chromosome 2 or 12 disrupt reference genes. Human reference sequence genes in the chromosome 2 region include *KYNU, GTDC1, ACVR2A, KIF5C, STAM2, KCNJ3, ERMN, PKP4, BAZ2B, TANK,* and *DPP4* (Fig. 5a, bottom). Breast cancer breaks near at least some of these genes interrupt functions essential for immunity and preventing cancer. For example, *KYNU* mediates the response to IFN-gamma. *TANK* is necessary for *NFKB* activation in the innate immune system. *DPP4* is essential for preventing viral entry into cells. Reference gene functions in the breakpoint region of chromosome 12 (Fig. 5b, bottom) include vesicle trafficking (*RASSF9*), endocytosis (*EEA1*), blood cell formation (*KITLG*), interferon response control (*SOCS2*), and nerve cell patterning (*NR2C1*).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

Potential retrovirus-like contributions are different in Figs. 5a vs. 5b. On chromosome 2, porcine endogenous retrovirus (PERV) [53], human endogenous retrovirus (HERV) sequences [54], and a pseudogene (pHERV) also have significant homology to human DNA. The PERV-like sequence lies within a retroposed area on chromosome 2 within 28 Kbps 5' and 80 Kbps 3' of EBV sequences. EBV-like ends potentially generate homologies for retro-positioning and inserting PERV sequences. In contrast, significant retroviral sequence homologies are outside the breakpoint-rich stretch on chromosome 12 (Fig. 5b).

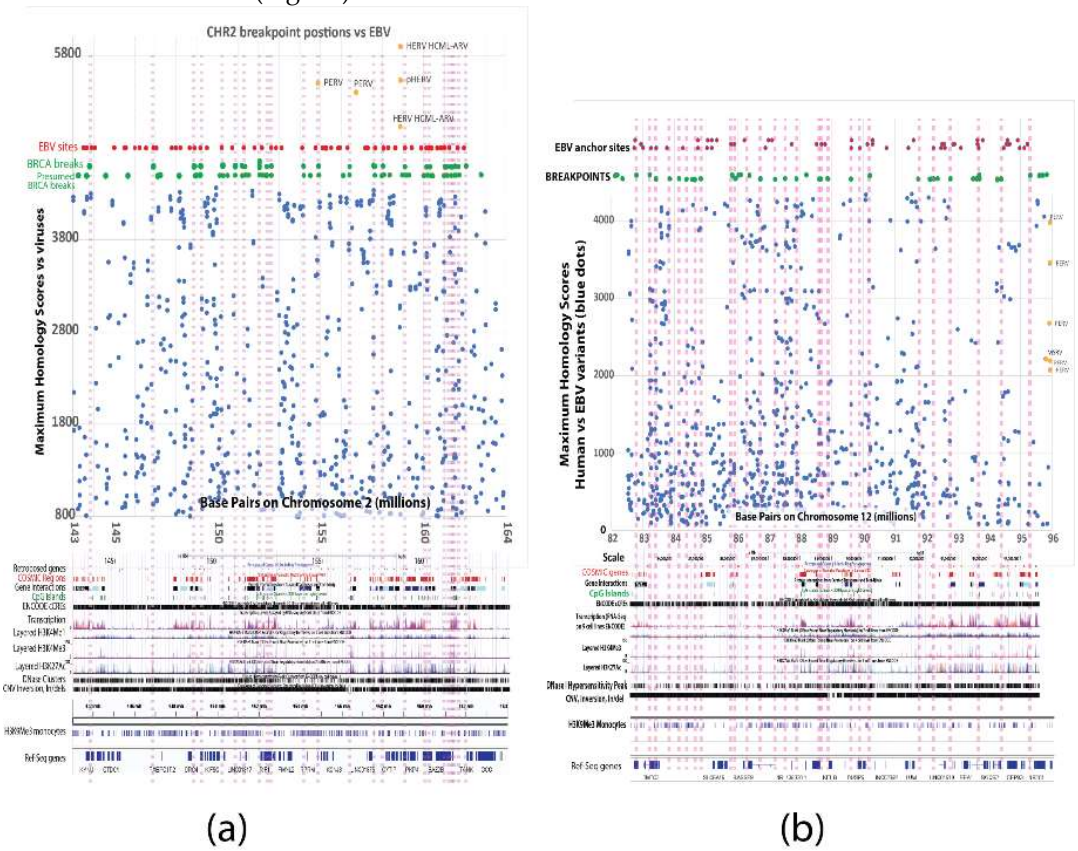


Figure 5. (a), Breakpoints on a 21 million base pair segment in human chromosome 2 and (b) on a 14 million base pair segment of chromosome 12 occur in BRCA1, BRCA2, and presumptive BRCA-associated breast cancers (green dots). Blue dots represent significant homology scores between human chromosome sequences vs. EBV variants HKNPC60 and HKHD40. Dashed lines align EBV genome anchor sites with breast cancer breakpoints. EBV anchor locations (red dots) and breast cancer breakpoints (green dots) are shown to allow comparisons. PERV and HERV variants (orange) may also contribute to breakpoints, but they distribute very differently on the two chromosome sections.

Identified EBV genome anchors near known genes match breast cancer breakpoints.

Fig. 6 further tests the relationship of breast cancer breaks to EBV genome anchor sites, precisely identified at disparate chromosome or gene locations [42]. In early-onset breast cancer PD23566, EBV episomal docking sites and EBV variant homologies fall within a deleted segment of chromosome 1. Breakpoints on chromosome 6 from three different breast

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

cancers surround EBV genome anchor sites and viral homologies. Even a chromosome 6 LINE retrotransposon has these homologies (Fig. 6a). A primary EBV genome binding site on chromosome 11 [42] matches a breakpoint in an early-onset breast cancer. A region of chromosome 5 containing an anchor site near HDAC3 has only distant EBV homologies. Instead, there are much shorter and weaker nearby similarities to retroviruses.

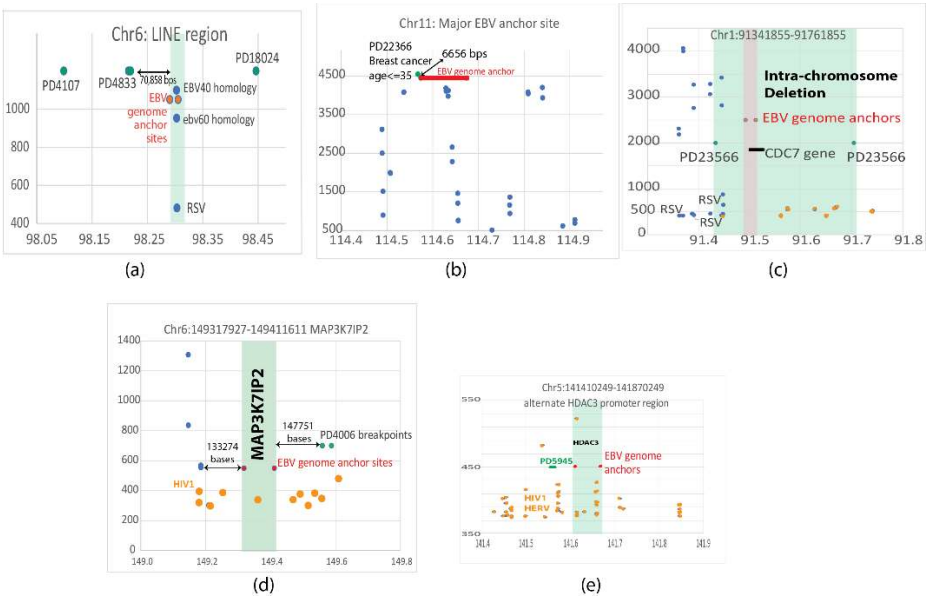


Figure 6. Maximum homology to human DNA for all viruses (y-axis) is plotted for known EBV genome anchor sites vs. LINE or gene coordinates. (a) A direct EBV binding site is near the region Chr6: 98,850,000-98,860,000, which contains a transposable LINE element (within the red dots) [26]. The area is near breakpoints in presumptive BRCA-associated breast cancers (green dots). Blue dots indicate significant HKHD40 and HKNPC60 homologies. Weaker homology to respiratory syncytial virus (RSV) also exists. In (b), a primary EBV binding site is within 6656 base pairs of a breakpoint in breast cancer PD22366. (c) EBV variant homologies near CDC7 fall within a deleted segment in an intra-chromosomal rearrangement on chromosome 1. In (d), EBV anchors are within 148 Kbps of a breast cancer breakpoint. The MAP3K7IP2 gene has EBV anchor sites at its boundaries and nearby viral homologies at 133,274 bases. In (e), breakpoints in breast cancer PD5945 are close to EBV-genome binding near the HDAC3 gene, but EBV variant homologies are more distant. Weaker homologies exist with HIV1 variants and endogenous retroviruses (orange dots).

The known EBV binding site at the *HDAC3* region on chr5:141610249-141660249 [42] was close to breakpoints in presumptive hereditary cancer PD5945 (at 141,557,433 and 141,564,233), but did not contain nearby significant EBV variant homologies. Instead, there were lower maximum homology matches to retroviruses. The *HDAC3* results suggest alternate mechanisms involving retroviral participation in breast cancer breakpoints (Fig. 6).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

Viral homologies around breakpoints in *BRCA* - associated breast cancers resemble model cancers caused by EBV.

DNA sequences at recurrent breakpoints in known EBV mediated cancers (nasopharyngeal cancer and Burkitt's lymphoma) have homologies to EBV variants, just like the breast cancers (Supplementary Figs. S2 and S3).

Supplementary Fig. S2 shows distances to EBV homologous regions starting from breakpoints in one nasopharyngeal cancer (NPC-5989) calculated from data in reference 12. Like hereditary breast cancers, NPC-5989 breakpoints are close to human genome regions that resemble the EBV variants. In Supplementary Fig. S3, the *PTK2* and *MAML* genes targeted by rearrangements in nasopharyngeal cancer (NPC) include strong similarities to EBV. The *RSRC1* breakpoint in a gene fusion of *RSRC1* to *PTPLB* in NPC is close to regions of strong EBV homology. However, the flanking sequences around the *PTPLB* breakpoint have homology to retroviral sequences.

The known EBV cancers and hereditary breast cancers break at some of the same chromosome positions (Table 3). Comparable breakpoint positions in BRCA-breast vs. known EBV cancers (NPC and. BL) differ by less than 1.25%). Normality plots of the breakpoint positions in Table 3 for all three types of cancers are identical, go through zero, and follow the line y=x. One set of data can calculate the other using the equation Breast Cancer Breaks=1.00021(NPC_BL Breaks) +9012 ($p<0.0001$). An unpaired t-test did not find significant differences between the two sets of data, $p<0.001$.

Table 2 Comparisons of Cancer Chromosome Breakpoints

Chro mo-some	NPC /BL Break coordinate	Hereditary breast cancer break coordinate	% Difference
1	151,928,027	151,922,525	0.00
1	155,119,764	155,020,623	0.06
1	154,117,249	154,111,293	0.00
1	197,593,291	197,867,477	0.14
1	203,179,378	203,428,861	0.12
1	204,435,983	204,386,502	0.02
2	66,324,118	66,473,128	0.22
3	158,537,092	158,531,593	0.00
4	101,138,921	101,009,819	0.13
4	152804117	152,852,058	0.03
5	43086276	42897159	0.44
6	74142321	74016919	0.17
6	104668449	104,676,699	0.01
8	98073546	98,164,150	0.09
8	128,468,417	128,300,594	0.13
8	128,481,840	128,300,594	0.14

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

11	102,114,930	102,384,027	0.26
11	102,115,180	102,384,027	0.26
11	101,983,850	101,895,690	0.09
13	74133903	74,125,809	0.01
15	41381710	41,532,838	0.37
18	43243754	43103981	0.32
19	14770552	14,943,490	1.17
19	33307575	33,086,916	0.66
X	3,996,343	4,044,529	1.21

Damage to genes needed for the immune system in *BRCA1* and *BRCA2* associated breast cancers.

All 25 hereditary breast cancers have significant damage to genes needed for immune system functions (Table 4). A total of at least 1307 immune-related genes had mutations. Significant differences exist in distributions and numbers of mutations among the breast cancer genomes [9, 10, 12, 15]. Still, mutations commonly cripple some aspect of innate immunity, its regulation, or its connections to adaptive immunity (Table 4). Deregulation of innate immune responses may increase mutagenesis and drive multiple human cancers [55]. The damage interferes with responses to antigens, pathogens, the ability to remove abnormal cells and likely allow latent EBV infections to escape from control. The top 20 mutations also universally affect the nervous system, perhaps increasing susceptibility to herpes viral infection.

Table 4. The top 20 most commonly mutated genes in hereditary breast cancers are associated with the immune and nervous systems.

Gene	Mutations in 25 *BRCA1* or *BRCA2* breast cancers	Type of Immunity listed in database [56]	Example of function in immunity	Known or likely connection to the nervous system
AKT3	20	Innate	AKT3 amplifies innate immune responses to DNA or RNA virus infections [57]. AKT3 binds interferon response factor 3, enhancing its activation and stimulating interferon responses.	✓
FRMD4A	20	Innate	Depends on Interferon Regulatory Factor 5 [58]. Affects antigen presentation and subsequent effects on dendritic cells. Mediated by cytoskeletal actin and endocytosis	✓
ARHGAP15	19	Innate	Negative regulator of neutrophil function. Affects mast cell function	✓

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

CAMTA1	19	Adaptive	Pattern of methylation distinguishes subsets of T-cells [59]	✓
DPYD	19	Innate	Interferon response [60]. Affects infection severity in mice. Natural killer cell function	✓
CHRM3	18	Adaptive	Cholinergic receptor muscarinic 3 Deficiency associated with immune impairment [61] Targeting by miRNAs related to inflammation [62].	✓
ADAMTS12	17	Adaptive	Member of a group of related proteins with many connections to the immune system and immune disorders [63]	✓
DAB1	17	Adaptive	Associated with microbial profile [64]	✓
ADAMTS3	16	Innate	Member of a group of related proteins with many connections to the immune system and immune disorders [63] Mast cell function	✓
COL23A1	16	Adaptive	Enriched in plasmacytoid dendritic cells	✓
DLC1	16	Adaptive	Immune regulation of stem cells through interactions with *NOTCH-1* [65] Type 2 T helper cell	✓
ADAM12	15	Adaptive	Controls response to infectious disease [66]. Central memory CD8 t-cells	✓
CD36	15	Adaptive	*TLR4-TLR6-Cd36* activation is a common molecular mechanism by which atherogenic lipids and amyloid-beta stimulate sterile inflammation. Gamma delta T cell function	✓
CDH2	15	Innate	Maintains stem cell quiescence [67]. Natural killer cell function	✓
DACH1	15	Innate	Downregulated in lymphoid cell progenitor [68]	✓
CREB5	14	Innate	CREB factors limit proinflammatory responses, provide anti-apoptotic signal, regulate Th1, Th2, Th3 responses, generate and maintain regulatory T cells [69]	✓
ANK1	13	Adaptive	Type 17 T helper cell	✓
HDAC9	13	Adaptive	Known to alter B-cell responses in autoimmune diseases Immature B cell function.	✓

 doi:10.20944/preprints202107.0430.v1

COL4A1	12	Adaptive	Central memory CD4 T cell	✓
F13A1	12	Adaptive	Phagocytes at CNS borders [70]	✓

Comparisons of viral variants to other viruses

The EBV tumor viruses (HKHD40 and HKNPC60) were typical of many other herpesvirus isolates, with some haplotypes conferring a high NPC risk [26]. About 100 other gamma herpes viral variants strongly matched HKHD40 and HKNPC60 in regions with enough data to make comparisons possible. HKNPC60 was 99% identical to the EBV reference sequence at bases 1-7500 and 95% identical at bases 1,200,000-1,405,000. HKHD40 gave values of 99 and 98% identity for comparisons to the same regions.

Comparisons of breast cancer mutations to known or likely viral cancers

Sets of genes mutated in breast cancers are similar to known or likely viral cancers. A collection of breast cancers share 1143 identical genes mutated with viral cervical cancer and about 50% of genes mutated in Burkitt's lymphoma and viral liver cancers.

Some fragile site breaks are near BRCA1/2 breast cancer breakpoints

Lu et al. found 4785 EBNA1 binding sites with over 50% overlapping a repetitive sequence element [42], such as runs of consecutive A-T bases. Kim et al. reported that EBNA1 anchor sites have A-T rich flanking sequences [43].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

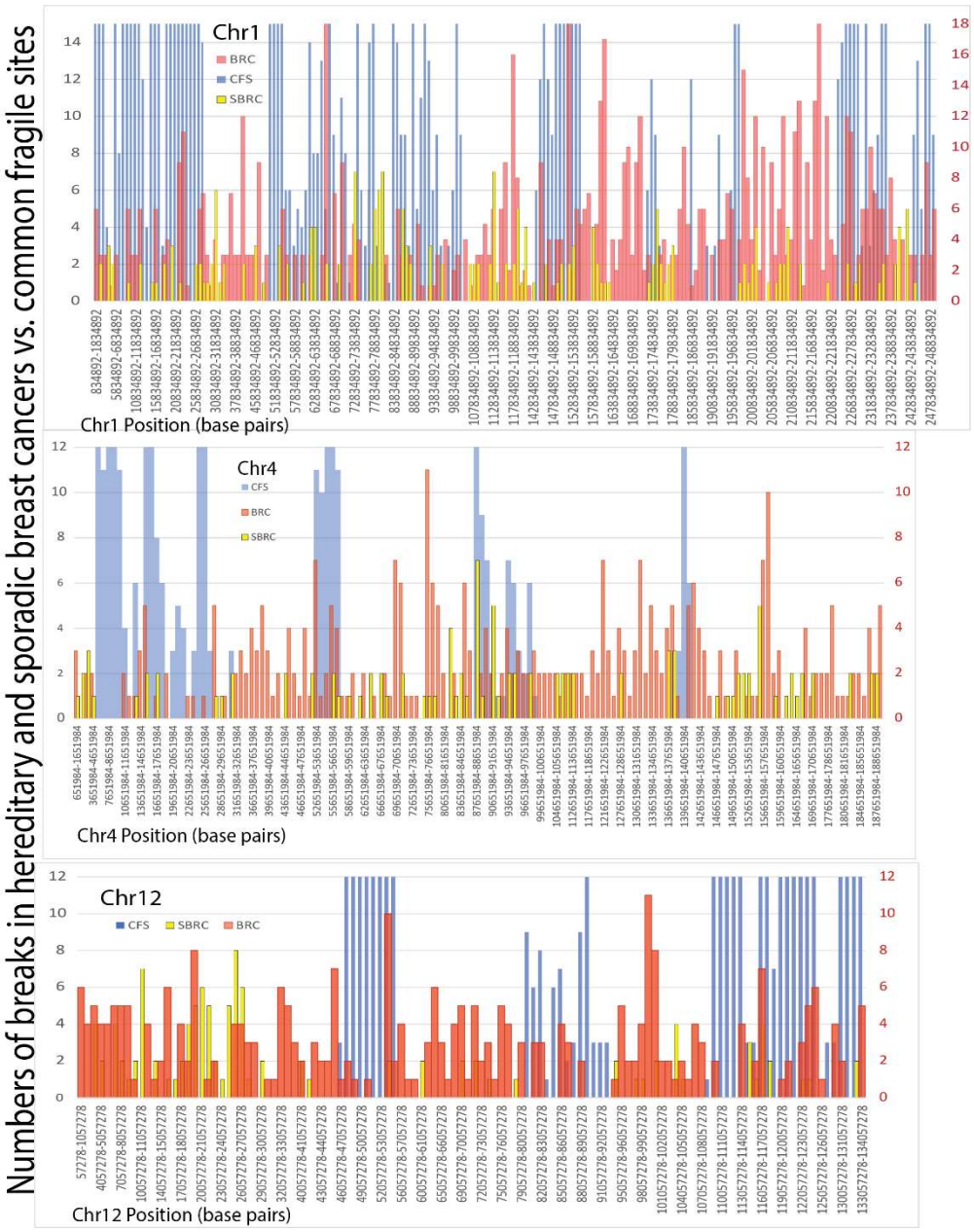


Figure 7. Histograms showing the relative positions of all breakpoints in BRCA1 or BRCA2 associated breast cancers on chromosome 1 vs the positions of common fragile site sequences (blue). Each chromosome was divided into 200 bins with values shown on the horizontal axis. The vertical axis is the number of times breakpoints in hereditary (red) or sporadic (yellow) breast cancer or CFS values fell into each bin.

Breakpoints in hereditary breast cancers were more likely to occur near common fragile sites (CFS) than breakpoints in sporadic breast cancers (SBRC) (Fig. 7). The alignment of hereditary cancer breakpoints with common fragile sites varied depending on the chromosome. On chromosome 1, common fragile sites fell into the same bins with 353

 doi:10.20944/preprints202107.0430.v1

hereditary breast cancer breakpoints vs. 97 sporadic breast cancer breakpoints. Based on the fragile site database, chromosome 1 contains 658 fragile site genes, the most of any chromosome [39]. Flanking sequences were rich in A-T bases.

On chromosomes 4 and 12, breast cancer breaks near fragile sites are more likely in hereditary than in sporadic breast cancers (329 vs 30 on chromosome 4 and 75 vs 16 on chromosome 12). On chromosome 4, interchromosomal breakpoints were more frequent and more likely to be near fragile sites in hereditary than in sporadic breast cancers (10 vs 1). Almost half the chromosome 4 fragile sites were in histogram bins near BRCA2-associated cancer breakpoints (56 of 126 CFS).

While there is considerable overlap between fragile sites and BRCA2 associated breakpoints, fragile sites are not sufficient to explain BRCA- and sporadic breast cancer associated breaks. On all the chromosomes tested, there were many more breakpoints than fragile sites. Many breast cancer breaks do not occur near common fragile sites (Fig. 7). Some hereditary breast cancer breakpoints were tested for repeats likely to generate fragile sites because the repeats are difficult to replicate. This test did not find such sequences (supplementary Table S1). In contrast, essentially all chromosome 4 interchromosomal breaks are close to human EBV-like sequences. According to one model, sites of replication errors in even one cell can be a sudden catastrophe that cascades into further breaks, destabilizing the entire genome [13]. This cascade is more likely in hereditary breast cancers with their deficits in homologous recombination repair.

Discussion

In carriers of *BRCA1* and *BRCA2* mutations, the association of EBV variant sequences with chromosome aberrations does not require viral integration or the continuing presence of active viruses anywhere within the resulting tumor. Because virtually everyone carries EBV infection, viral participation in breast cancer has been very difficult to distinguish from its role in seemingly normal cells.

This work produces a working model for how EBV variants associate with chromosome breaks in hereditary breast cancers. Multiple lines of evidence support the model given below (Fig. 8). For example, positions of viral anchors (EBNA1) are near chromosome breakpoints; human herpes virus-like sequences are frequently found in sequences flanking breast cancer breakpoints; patterns of viral homology are challenging to distinguish from cancers known to be caused by EBV. Although some breast cancer inter-chromosomal breakpoints are near fragile sites, they are more often near EBV-like human sequences.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

Fig. 8 summarizes the proposed model. Pathogenic mutations in either BRCA gene increase persistence of chromosome breaks caused by replication errors, mutagens, DNA damaging agents, and inflammation. Inactive BRCA genes also compromise homologous recombination pathways needed to restore fragmented chromosomes to their original condition. In this context, almost every human harbors latent EBV infection as a circular viral DNA parasite genome. Even early EBV infection causes massive changes in host cell gene transcription programs. Viral episomes can establish up to four different premalignant latent gene expression programs, with most viral genes shut off by epigenetic marks. The immune response typically controls EBV, allowing its four latent forms to express only the few genes that do not trigger immunity. One of these viral genes expresses EBNA1, which docks EBV genomes to host chromosomes at hundreds of sites, where the virus may still alter control of host cellular genes. Viral genomes remain attached as parasites even during cell division in mutation carriers and interfere with it. This attachment maintains stable numbers of episome copies in multiplying host cells [71, 72] and increases replication stress. When chromosomes break in BRCA-associated breast cancers, attached EBNA1 and viral DNA obstruct access to machinery required for repair.

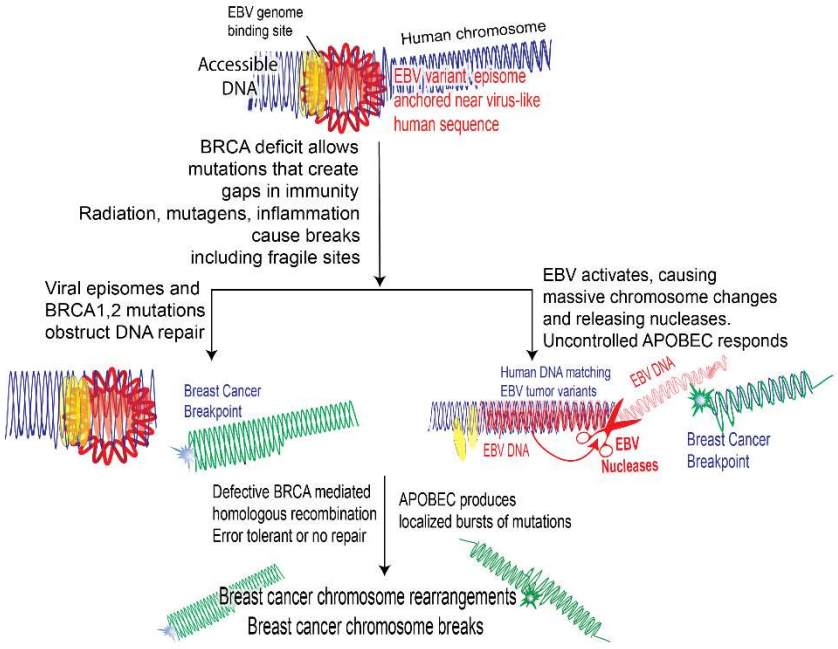


Figure 8. Model for the role for EBV variants in chromosome breaks in breast cancers associated with BRCA1 and BRCA2 gene mutations

This interference becomes so significant in mouse models that B-cell tumors develop [73]. Flanking the viral docking sites on human chromosomes are regions that match the viral sequence, are accessible to DNases. The flanks sometimes contain a common fragile site, and are rich in repetitive A-T DNA sequences.

 doi:10.20944/preprints202107.0430.v1

Translocations can generate a burst of localized mutations through the actions of APOBEC ("kataegis") [74]. APOBEC is typically a response to inactivate viral infections and APOBEC3 probably mediates EBV-induced carcinogenesis [55, 75]. Gene regulation disruptions in breast cancers around translocation breakpoints could easily deregulate APOBEC3 (Figs. 5). On chromosome 2, breakpoints go through genes essential for immune responses. Whether or not immune compromise happens before or after chromosome breaks, mutations in *BRCA1* and *BRCA2* associated breast cancers compromise host immune responses, favoring the conversion of inactive circular EBV to active linear forms. On activation, the virus instructs host cells to remove inhibitory epigenetic marks from the viral DNA [76] and the dissociation of EBNA1 oligomers occurs [77]. The activated viral forms cause massive host chromosome changes as they force host cells into viral production. The infection spreads, and more host cells become infected with latent forms. Viral production that does not go to completion allows viral enzymes to persist as cancer drivers in abnormal cells [78, 79]. Latent circular viral episomes obstruct the host cell's ability to access a broken human chromosome for repair. Activated EBV in tumors produces nucleases BGLF5 and BALF3, which cause additional DNA breaks and inhibit host protein synthesis needed for immunity [80, 81]. The viral nucleases cause further host chromosome aberrations, micronuclei [80], and they block the production of host DNA repair proteins. Relatively unopposed EBV replication then drives EBV-mediated malignancies, [73] especially when cells survive because viral production does not go to completion to cause host cell lysis [79]. Evidence for lytic EBV forms in cancer is abundant [18, 73]. Whatever repairs are made produce abnormal chromosome structures.

There are countless agents and mechanisms known to break DNA and chromosomes that do not involve viruses. There is selectivity among breast cancers in partners for improper repair (Fig. 1). Partners for inter-chromosomal rearrangements and translocations are typically close to each other. An individual chromosome resides in its own spatial domain in the nucleus relative to other chromosomes [52]. Interference from viral DNA adds to this spatial limitation to change translocation partners. Large differences exist in the distributions of inter-chromosome translocations in hereditary vs. sporadic breast cancers. EBV cause massive changes in the spatial distribution of chromosomes so that broken chromosome fragments have new nearby translocation partners.

EBV contributes to these differences because nearly all the inter-chromosome breakpoints on the entire length of chromosome 4 occur within the approximate number of base pairs in lytic EBV variant sequences (roughly 200,000 in Fig. 4). Many breast cancer breakpoints are separated from viral homologous sequences by numbers of base pairs consistent with circular viral episome lengths (roughly 65,000 base pairs, Fig. 7). Circular viral episomes could also become a dominant carcinogen by blocking access to proteins and nucleic acids needed for repair.

Comparisons of chromosome break sites in hereditary cancers to actual locations of EBNA1 anchor binding further supports their relationship. Comparisons on chromosome 2 (21 million base pairs) and chromosome 12 (14 million base pairs) give similar results (Fig. 5). DNA comparisons for a few EBV binding sites at precisely known LINE transposable element or gene

 doi:10.20944/preprints202107.0430.v1

sequences (Fig. 6) also show that viral genome binding, breast cancer breaks, and viral human homologies are often close together.

EBV episomes bind to human chromosomes using EBNA1-dimers as anchors for subsequent viral attachment [77, 82]. The dimerized anchor itself can alter human chromosome structure and mediate distant interactions related to transcription during infection. EBNA1 anchor sites favor repetitive elements, especially LINE 1 retrotransposons (Fig. 6), and correlate weakly with histone modifications [42].

Retroviruses and retrotransposons [83] may also participate in breast cancer breaks. Participation from porcine endogenous retroviruses is actionable by thoroughly cooking pork products. However, despite assertions that xenotransplantation with pig cells is safe, it is concerning that up to 6500 bps in human chromosome 11 are virtually identical to pig DNA (Fig. 5a).

Nasopharyngeal cancer (NPC) and Burkitt's lymphoma (BL) have known links to EBV. At least 25 breakpoints in these known EBV-mediated cancers are probably within experimental error [84] of breakpoint positions in breast cancers (Table 3). Running these comparisons for such accepted EBV-related cancers gave results that were difficult to distinguish from breast cancers (Figs. S2, S3, and Table 3). Five of six genes involved in gene fusions are near EBV variant sequences and include EBV-like variant sequences within the gene boundaries (Fig. S3).

Breast cancer mutations cripple the immune system's ability to control cancer-causing infections, remove cells damaged by disease, and regulate inflammation. Many genes mutated in human breast cancers link to some variable function within the immune system or structural barrier defenses [28, 85, 86]. Antigen and viral recognition, signaling essential to transmit the immune responses, and immune regulation can all be impaired. Many mutations independently link to various infections, including all known cancer-causing microorganisms. The top 20 mutated genes in BRCA1 and BRCA2 breast cancers are associated with some function of the immune response (Table 4). The relationships of the top 20 mutated genes to the nervous system may also cause herpes viral infections to favor nervous tissues.

Multiple components of immune defenses mutate in the 25 hereditary breast cancer genomes. The mutations affect processes such as cytokine production, autophagy, etc. These functions depend on many genes dispersed throughout the genome, so any cancer needs only to damage one gene to cripple an immune function. Each breast cancer genome has a different set of these mutations, with the same gene only occasionally damaged. Damage affecting the nervous system was also universal. Some herpes viruses establish occult infection within the central nervous system even after other sites become virus-free [87].

The model based on the present work is potentially actionable. The current evidence adds support for developing EBV treatment and a childhood herpes vaccine. EBV causes about 200,000 cancers per year of multiple different types. The prospects for producing an EBV vaccine are promising, but the most appropriate targets are still not settled. Some immunotherapy strategies rely on augmenting the immune response, but this

approach may need modification because mutations create additional holes in the immune response.

A limitation of the results is a relatively small sample size. The breast cancer data comes from 560 breast cancer genome sequences [12], with hereditary breast cancers comprising only a small percentage. This limitation is mainly due to the rarity of hereditary breast cancers. Larger studies are clearly needed but representing the breadth of all somatic and hereditary breast cancers is a significant problem. The consistency of inter-chromosomal breakpoints in two random samples of sporadic breast cancers is nonetheless reassuring (Figs 1 and S1).

References

1. Sun Y, McCorvie TJ, Yates LA, Zhang X: **Structural basis of homologous recombination**. *Cellular and molecular life sciences : CMLS* 2020, **77**(1):3-18.
2. Murashko MM, Stasevich EM, Schwartz AM, Kuprash DV, Uvarova AN, Demin DE: **The Role of RNA in DNA Breaks, Repair and Chromosomal Rearrangements**. *Biomolecules* 2021, **11**(4).
3. Garcia-de-Teresa B, Rodriguez A, Frias S: **Chromosome Instability in Fanconi Anemia: From Breaks to Phenotypic Consequences**. *Genes (Basel)* 2020, **11**(12).
4. Venkitaraman AR: **Tumour suppressor mechanisms in the control of chromosome stability: insights from BRCA2**. *Molecules and cells* 2014, **37**(2):95-99.
5. Arlt MF, Xu B, Durkin SG, Casper AM, Kastan MB, Glover TW: **BRCA1 is required for common-fragile-site stability via its G2/M checkpoint function**. *Molecular and cellular biology* 2004, **24**(15):6701-6709.
6. Voutsinos V, Munk SHN, Oestergaard VH: **Common Chromosomal Fragile Sites-Conserved Failure Stories**. *Genes (Basel)* 2018, **9**(12).
7. Xu B, Kim S, Kastan MB: **Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionizing irradiation**. *Molecular and cellular biology* 2001, **21**(10):3445-3450.
8. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL *et al*: **Signatures of mutational processes in human cancer**. *Nature* 2013, **500**(7463):415-421.
9. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR: **Deciphering signatures of mutational processes operative in human cancer**. *Cell reports* 2013, **3**(1):246-259.
10. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA *et al*: **Mutational Processes Molding the Genomes of 21 Breast Cancers**. *Cell* 2012, **149**(5):979-993.
11. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, Raine K, Jones D, Marshall J, Ramakrishna M *et al*: **The life history of 21 breast cancers**. *Cell* 2012, **149**(5):994-1007.
12. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC *et al*: **Landscape of somatic mutations in 560 breast cancer whole-genome sequences**. *Nature* 2016, **534**(7605):47-54.
13. Umbreit NT, Zhang CZ, Lynch LD, Blaine LJ, Cheng AM, Tourdot R, Sun L, Almubarak HF, Judge K, Mitchell TJ *et al*: **Mechanisms generating cancer genome complexity from a single cell division error**. *Science* 2020, **368**(6488).
14. Cajuso T, Sulo P, Tanskanen T, Katainen R, Taira A, Hanninen UA, Kondelin J, Forsstrom L, Valimaki N, Aavikko M *et al*: **Retrotransposon insertions can initiate colorectal cancer and are associated with poor survival**. *Nat Commun* 2019, **10**(1):4022.

15. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL *et al*: **Signatures of mutational processes in human cancer**. *Nature* 2013, **500**(7463):415-421.
16. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC *et al*: **Author Correction: Landscape of somatic mutations in 560 breast cancer whole-genome sequences**. *Nature* 2019, **566**(7742):E1.
17. Morganella S, Alexandrov LB, Glodzik D, Zou X, Davies H, Staaf J, Sieuwerts AM, Brinkman AB, Martin S, Ramakrishna M *et al*: **The topography of mutational processes in breast cancer genomes**. *Nat Commun* 2016, **7**:11383.
18. Zapatka M, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M, Desai N, Sultmann H, Moch H, Pathogens P *et al*: **The landscape of viral associations in human cancers**. *Nature genetics* 2020, **52**(3):320-330.
19. Griffiths DJ: **Endogenous retroviruses in the human genome sequence**. *Genome Biol* 2001, **2**(6):REVIEWS1017.
20. Ueda MT, Kryukov K, Mitsuhashi S, Mitsuhashi H, Imanishi T, Nakagawa S: **Comprehensive genomic analysis reveals dynamic evolution of endogenous retroviruses that code for retroviral-like protein domains**. *Mob DNA* 2020, **11**:29.
21. Xiang Y, Liang H: **The Regulation and Functions of Endogenous Retrovirus in Embryo Development and Stem Cell Differentiation**. *Stem Cells Int* 2021, **2021**:6660936.
22. Goke J, Ng HH: **CTRL+INSERT: retrotransposons and their contribution to regulation and innovation of the transcriptome**. *EMBO Rep* 2016, **17**(8):1131-1144.
23. Moustafa A, Xie C, Kirkness E, Biggs W, Wong E, Turpaz Y, Bloom K, Delwart E, Nelson KE, Venter JC *et al*: **The blood DNA virome in 8,000 humans**. *PLoS pathogens* 2017, **13**(3):e1006292.
24. Germini D, Sall FB, Shmakova A, Wiels J, Dokudovskaya S, Drouet E, Vassetzky Y: **Oncogenic Properties of the EBV ZEBRA Protein**. *Cancers (Basel)* 2020, **12**(6).
25. Hau PM, Lung HL, Wu M, Tsang CM, Wong KL, Mak NK, Lo KW: **Targeting Epstein-Barr Virus in Nasopharyngeal Carcinoma**. *Front Oncol* 2020, **10**:600.
26. Xu M, Yao Y, Chen H, Zhang S, Cao SM, Zhang Z, Luo B, Liu Z, Li Z, Xiang T *et al*: **Genome sequencing analysis identifies Epstein-Barr virus subtypes associated with high risk of nasopharyngeal carcinoma**. *Nature genetics* 2019, **51**(7):1131-1136.
27. Xu M, Zhang WL, Zhu Q, Zhang S, Yao YY, Xiang T, Feng QS, Zhang Z, Peng RJ, Jia WH *et al*: **Genome-wide profiling of Epstein-Barr virus integration by targeted sequencing in Epstein-Barr virus associated malignancies**. *Theranostics* 2019, **9**(4):1115-1124.
28. Friedenson B: **Mutations in components of antiviral or microbial defense as a basis for breast cancer**. *Functional & integrative genomics* 2013, **13**(4):411-424.
29. Lung RW, Tong JH, Ip LM, Lam KH, Chan AW, Chak WP, Chung LY, Yeung WW, Hau PM, Chau SL *et al*: **EBV-encoded miRNAs can sensitize nasopharyngeal carcinoma to chemotherapeutic drugs by targeting BRCA1**. *Journal of cellular and molecular medicine* 2020, **24**(22):13523-13535.
30. Dai W, Chung DL, Chow LK, Yu VZ, Lei LC, Leong MM, Chan CK, Ko JM, Lung ML: **Clinical Outcome-Related Mutational Signatures Identified by Integrative Genomic Analysis in Nasopharyngeal Carcinoma**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2020, **26**(24):6494-6504.
31. Valouev A, Weng Z, Sweeney RT, Varma S, Le QT, Kong C, Sidow A, West RB: **Discovery of recurrent structural variants in nasopharyngeal carcinoma**. *Genome research* 2014, **24**(2):300-309.

32. Fina F, Romain S, Ouafik L, Palmari J, Ben Ayed F, Benharkat S, Bonnier P, Spyratos F, Foekens JA, Rose C *et al*: **Frequency and genome load of Epstein-Barr virus in 509 breast cancers from different geographical areas**. *British journal of cancer* 2001, **84**(6):783-790.
33. Peng J, Wang T, Zhu H, Guo J, Li K, Yao Q, Lv Y, Zhang J, He C, Chen J *et al*: **Multiplex PCR/mass spectrometry screening of biological carcinogenic agents in human mammary tumors**. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 2014, **61**(2):255-259.
34. Hu H, Luo ML, Desmedt C, Nabavi S, Yadegarynia S, Hong A, Konstantinopoulos PA, Gabrielson E, Hines-Boykin R, Pihan G *et al*: **Epstein-Barr Virus Infection of Mammary Epithelial Cells Promotes Malignant Transformation**. *EBioMedicine* 2016, **9**:148-160.
35. Ayee R, Ofori MEO, Wright E, Quaye O: **Epstein Barr Virus Associated Lymphomas and Epithelia Cancers in Humans**. *Journal of Cancer* 2020, **11**(7):1737-1750.
36. Lorenzetti MA, De Matteo E, Gass H, Martinez Vazquez P, Lara J, Gonzalez P, Preciado MV, Chabay PA: **Characterization of Epstein Barr virus latency pattern in Argentine breast carcinoma**. *PloS one* 2010, **5**(10):e13603.
37. Marrao G, Habib M, Paiva A, Bicout D, Fallecker C, Franco S, Fafi-Kremer S, Simoes da Silva T, Morand P, Freire de Oliveira C *et al*: **Epstein-Barr virus infection and clinical outcome in breast cancer patients correlate with immune cell TNF-alpha/IFN-gamma response**. *BMC cancer* 2014, **14**:665.
38. Maccaroni K, Balzano E, Mirimao F, Giunta S, Pelliccia F: **Impaired Replication Timing Promotes Tissue-Specific Expression of Common Fragile Sites**. *Genes (Basel)* 2020, **11**(3).
39. Kumar R, Nagpal G, Kumar V, Usmani SS, Agrawal P, Raghava GPS: **HumCFS: a database of fragile sites in human chromosomes**. *BMC genomics* 2019, **19**(Suppl 9):985.
40. Mount DW: **Using the Basic Local Alignment Search Tool (BLAST)**. *CSH Protoc* 2007, **2007**:pdb top17.
41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *Journal of molecular biology* 1990, **215**(3):403-410.
42. Lu F, Wikramasinghe P, Norseen J, Tsai K, Wang P, Showe L, Davuluri RV, Lieberman PM: **Genome-wide analysis of host-chromosome binding sites for Epstein-Barr Virus Nuclear Antigen 1 (EBNA1)**. *Virology journal* 2010, **7**:262.
43. Kim KD, Tanizawa H, De Leo A, Vladimirova O, Kossenkov A, Lu F, Showe LC, Noma KI, Lieberman PM: **Epigenetic specifications of host chromosome docking sites for latent Epstein-Barr virus**. *Nat Commun* 2020, **11**(1):877.
44. Xiao K, Yu Z, Li X, Li X, Tang K, Tu C, Qi P, Liao Q, Chen P, Zeng Z *et al*: **Genome-wide Analysis of Epstein-Barr Virus (EBV) Integration and Strain in C666-1 and Raji Cells**. *Journal of Cancer* 2016, **7**(2):214-224.
45. Needleman SBaW, C.D.: **A general method applicable to search for similarities in the amino acid sequence of two proteins**. *Journal of molecular biology* 1970, **48**:453-453.
46. Wolfe D, Dudek S, Ritchie MD, Pendergrass SA: **Visualizing genomic information across chromosomes with PhenoGram**. *BioData Min* 2013, **6**(1):18.
47. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z: **Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade**. *Cell reports* 2017, **18**(1):248-262.
48. Lynn DJ, Winsor GL, Chan C, Richard N, Laird MR, Barsky A, Gardy JL, Roche FM, Chan TH, Shah N *et al*: **InnateDB: facilitating systems-level analyses of the mammalian innate immune response**. *Molecular systems biology* 2008, **4**:218.
49. Ortutay C, Siermala M, Vihinen M: **ImmTree: database of evolutionary relationships of genes and proteins in the human immune system**. *Immunome Res* 2007, **3**:4.

50. Ortutay C, Siermala M, Vihinen M: **Molecular characterization of the immune system: emergence of proteins, processes, and domains**. *Immunogenetics* 2007, **59**(5):333-348.
51. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN *et al*: **The repertoire of mutational signatures in human cancer**. *Nature* 2020, **578**(7793):94-101.
52. Leibowitz ML, Zhang CZ, Pellman D: **Chromothripsis: A New Mechanism for Rapid Karyotype Evolution**. *Annu Rev Genet* 2015, **49**:183-211.
53. Denner J: **The porcine virome and xenotransplantation**. *Virology journal* 2017, **14**(1):171.
54. Prusty BK, zur Hausen H, Schmidt R, Kimmel R, de Villiers EM: **Transcription of HERV-E and HERV-E-related sequences in malignant and non-malignant human haematopoietic cells**. *Virology* 2008, **382**(1):37-45.
55. Law EK, Levin-Klein R, Jarvis MC, Kim H, Argyris PP, Carpenter MA, Starrett GJ, Temiz NA, Larson LK, Durfee C *et al*: **APOBEC3A catalyzes mutation and drives carcinogenesis in vivo**. *The Journal of experimental medicine* 2020, **217**(12).
56. Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock RE, Brinkman FS, Lynn DJ: **InnateDB: systems biology of innate immunity and beyond--recent updates and continuing curation**. *Nucleic acids research* 2013, **41**(Database issue):D1228-1233.
57. Xiao J, Li W, Zheng X, Qi L, Wang H, Zhang C, Wan X, Zheng Y, Zhong R, Zhou X *et al*: **Targeting 7-Dehydrocholesterol Reductase Integrates Cholesterol Metabolism and IRF3 Activation to Eliminate Infection**. *Immunity* 2020, **52**(1):109-122 e106.
58. Chow KT, Driscoll C, Loo YM, Knoll M, Gale M, Jr.: **IRF5 regulates unique subset of genes in dendritic cells during West Nile virus infection**. *J Leukoc Biol* 2019, **105**(2):411-425.
59. Minskaia E, Saraiva BC, Soares MMV, Azevedo RI, Ribeiro RM, Kumar SD, Vieira AIS, Lacerda JF: **Molecular Markers Distinguishing T Cell Subtypes With TSDR Strand-Bias Methylation**. *Front Immunol* 2018, **9**:2540.
60. Zhu WP, Liu ZY, Zhao YM, He XG, Pan Q, Zhang N, Zhou JM, Wang LR, Wang M, Zhan DH *et al*: **Dihydropyrimidine dehydrogenase predicts survival and response to interferon-alpha in hepatocellular carcinoma**. *Cell Death Dis* 2018, **9**(2):69.
61. Grondin JA, Kwon YH, Far PM, Haq S, Khan WI: **Mucins in Intestinal Mucosal Defense and Inflammation: Learning From Clinical and Experimental Studies**. *Frontiers in immunology* 2020, **11**:2054.
62. Ibanez-Cabellos JS, Seco-Cervera M, Osca-Verdegal R, Pallardo FV, Garcia-Gimenez JL: **Epigenetic Regulation in the Pathogenesis of Sjogren Syndrome and Rheumatoid Arthritis**. *Front Genet* 2019, **10**:1104.
63. Mead TJ, Apte SS: **ADAMTS proteins in human disorders**. *Matrix Biol* 2018, **71-72**:225-239.
64. Reverter A, Ballester M, Alexandre PA, Marmol-Sanchez E, Dalmau A, Quintanilla R, Ramayo-Caldas Y: **A gene co-association network regulating gut microbial communities in a Duroc pig population**. *Microbiome* 2021, **9**(1):52.
65. Na T, Zhang K, Yuan BZ: **The DLC-1 tumor suppressor is involved in regulating immunomodulation of human mesenchymal stromal /stem cells through interacting with the Notch1 protein**. *BMC cancer* 2020, **20**(1):1064.
66. Aljohmani A, Yildiz D: **A Disintegrin and Metalloproteinase-Control Elements in Infectious Diseases**. *Front Cardiovasc Med* 2020, **7**:608281.
67. Florez MA, Matatall KA, Jeong Y, Ortinau L, Shafer PW, Lynch AM, Jaksik R, Kimmel M, Park D, King KY: **Interferon Gamma Mediates Hematopoietic Stem Cell Activation and Niche Relocalization through BST2**. *Cell reports* 2020, **33**(12):108530.
68. Amann-Zalcenstein D, Tian L, Schreuder J, Tomei S, Lin DS, Fairfax KA, Bolden JE, McKenzie MD, Jarratt A, Hilton A *et al*: **A new lymphoid-primed progenitor marked by Dach1**

downregulation identified with single cell multi-omics. *Nature immunology* 2020, **21**(12):1574-1584.

69. Wen AY, Sakamoto KM, Miller LS: **The role of the transcription factor CREB in immune function**. *Journal of immunology* 2010, **185**(11):6413-6419.
70. Ivan DC, Walthert S, Berve K, Steudler J, Locatelli G: **Dwellers and Trespassers: Mononuclear Phagocytes at the Borders of the Central Nervous System**. *Frontiers in immunology* 2020, **11**:609921.
71. Dheekollu J, Wiedmer A, Ayyanathan K, Deakyne JS, Messick TE, Lieberman PM: **Cell-cycle-dependent EBNA1-DNA crosslinking promotes replication termination at oriP and viral episome maintenance**. *Cell* 2021, **184**(3):643-654 e613.
72. De Leo A, Calderon A, Lieberman PM: **Control of Viral Latency by Episome Maintenance Proteins**. *Trends in microbiology* 2020, **28**(2):150-162.
73. Munz C: **Latency and lytic replication in Epstein-Barr virus-associated oncogenesis**. *Nature reviews Microbiology* 2019, **17**(11):691-700.
74. Nik-Zainal S, Morganella S: **Mutational Signatures in Breast Cancer: The Problem at the DNA Level**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2017, **23**(11):2617-2629.
75. Bobrovnitchaia I, Valieris R, Drummond RD, Lima JP, Freitas HC, Bartelli TF, de Amorim MG, Nunes DN, Dias-Neto E, da Silva IT: **APOBEC-mediated DNA alterations: A possible new mechanism of carcinogenesis in EBV-positive gastric cancer**. *International journal of cancer Journal international du cancer* 2020, **146**(1):181-191.
76. Buschle A, Mrozek-Gorska P, Cernilogar FM, Ettinger A, Pich D, Krebs S, Mocanu B, Blum H, Schotta G, Straub T *et al*: **Epstein-Barr virus inactivates the transcriptome and disrupts the chromatin architecture of its host cell in the first phase of lytic reactivation**. *Nucleic acids research* 2021, **49**(6):3217-3241.
77. Jiang L, Lung HL, Huang T, Lan R, Zha S, Chan LS, Thor W, Tsoi TH, Chau HF, Borestrom C *et al*: **Reactivation of Epstein-Barr virus by a dual-responsive fluorescent EBNA1-targeting agent with Zn(2+)-chelating function**. *Proceedings of the National Academy of Sciences of the United States of America* 2019.
78. Kirchner EA, Bornkamm GW, Polack A: **Transcriptional activity across the Epstein-Barr virus genome in Raji cells during latency and after induction of an abortive lytic cycle**. *The Journal of general virology* 1991, **72 (Pt 10)**:2391-2398.
79. Morales-Sanchez A, Fuentes-Panana EM: **The Immunomodulatory Capacity of an Epstein-Barr Virus Abortive Lytic Cycle: Potential Contribution to Viral Tumorigenesis**. *Cancers (Basel)* 2018, **10**(4).
80. Wu CC, Liu MT, Chang YT, Fang CY, Chou SP, Liao HW, Kuo KL, Hsu SL, Chen YR, Wang PW *et al*: **Epstein-Barr virus DNase (BGLF5) induces genomic instability in human epithelial cells**. *Nucleic acids research* 2010, **38**(6):1932-1949.
81. Chiu SH, Wu CC, Fang CY, Yu SL, Hsu HY, Chow YH, Chen JY: **Epstein-Barr virus BALF3 mediates genomic instability and progressive malignancy in nasopharyngeal carcinoma**. *Oncotarget* 2014, **5**(18):8583-8601.
82. **Correction for Jiang et al., Reactivation of Epstein-Barr virus by a dual-responsive fluorescent EBNA1-targeting agent with Zn(2+)-chelating function**. *Proceedings of the National Academy of Sciences of the United States of America* 2020, **117**(10):5542.
83. Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M: **Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing**. *Genome research* 2014, **24**(7):1053-1063.

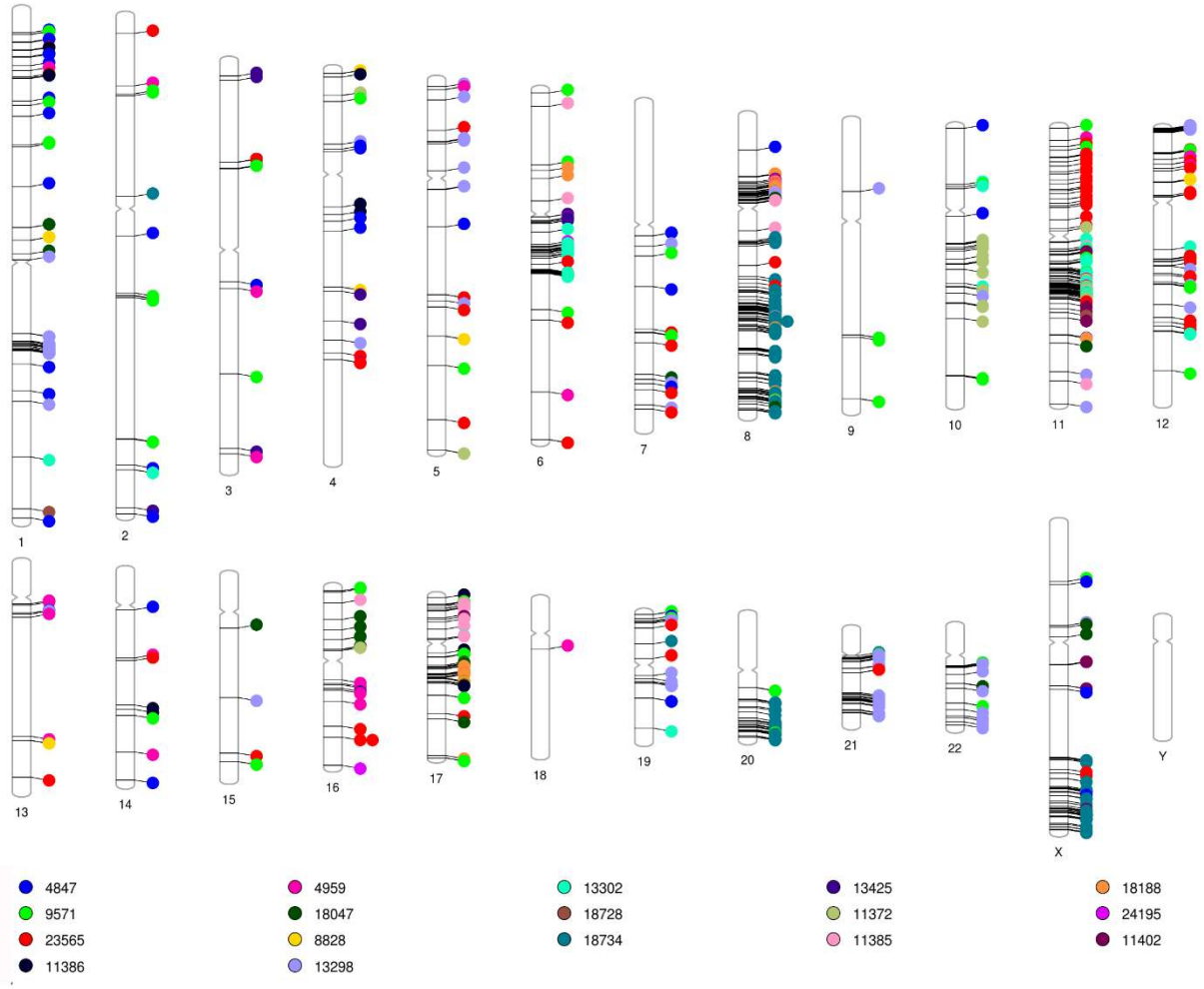
84. Pfeiffer F, Grober C, Blank M, Handler K, Beyer M, Schultze JL, Mayer G: **Systematic evaluation of error rates and causes in short samples in next-generation sequencing**. *Scientific reports* 2018, **8**(1):10950.
85. Friedenson B: **Many Breast Cancer Mutations Parallel Mutations in Known Viral Cancers**. *Journal of Genomes and Exomes* 2014, **3**(4437-JGE-Many-Breast-Cancer-Mutations-Parallel-Mutations-in-Known-Viral-Cancers-.pdf):17-35.
86. Friedenson B: **Mutations in Breast Cancer Exome Sequences Predict Susceptibility to Infections and Converge on the Same Signaling Pathways**. *J Genomes and Exomes* *http://mrcrossreforg/iPage?doi=104137%2FJGES30058* 2015, **4**:1-28.
87. Bhela S, Mulik S, Reddy PB, Richardson RL, Gimenez F, Rajasagi NK, Veiga-Parga T, Osmand AP, Rouse BT: **Critical role of microRNA-155 in herpes simplex encephalitis**. *Journal of immunology* 2014, **192**(6):2734-2743.

Supplementary Materials: Figures S1 Chromosome distributions of interchromosomal translocations in alternate set of sporadic breast cancers. Fig S2. Distribution of Distributions of EBV variant homologies near breakpoints in nasopharyngeal cancers Fig S3. Distribution of EBV variant homologies in NPC gene fusion breakpoints. Table S1. Repeating sequences are absent at breakpoints in hereditary breast cancers.

Funding: Please add: This research received no external funding

Conflicts of Interest: The author declares no conflict of interest.

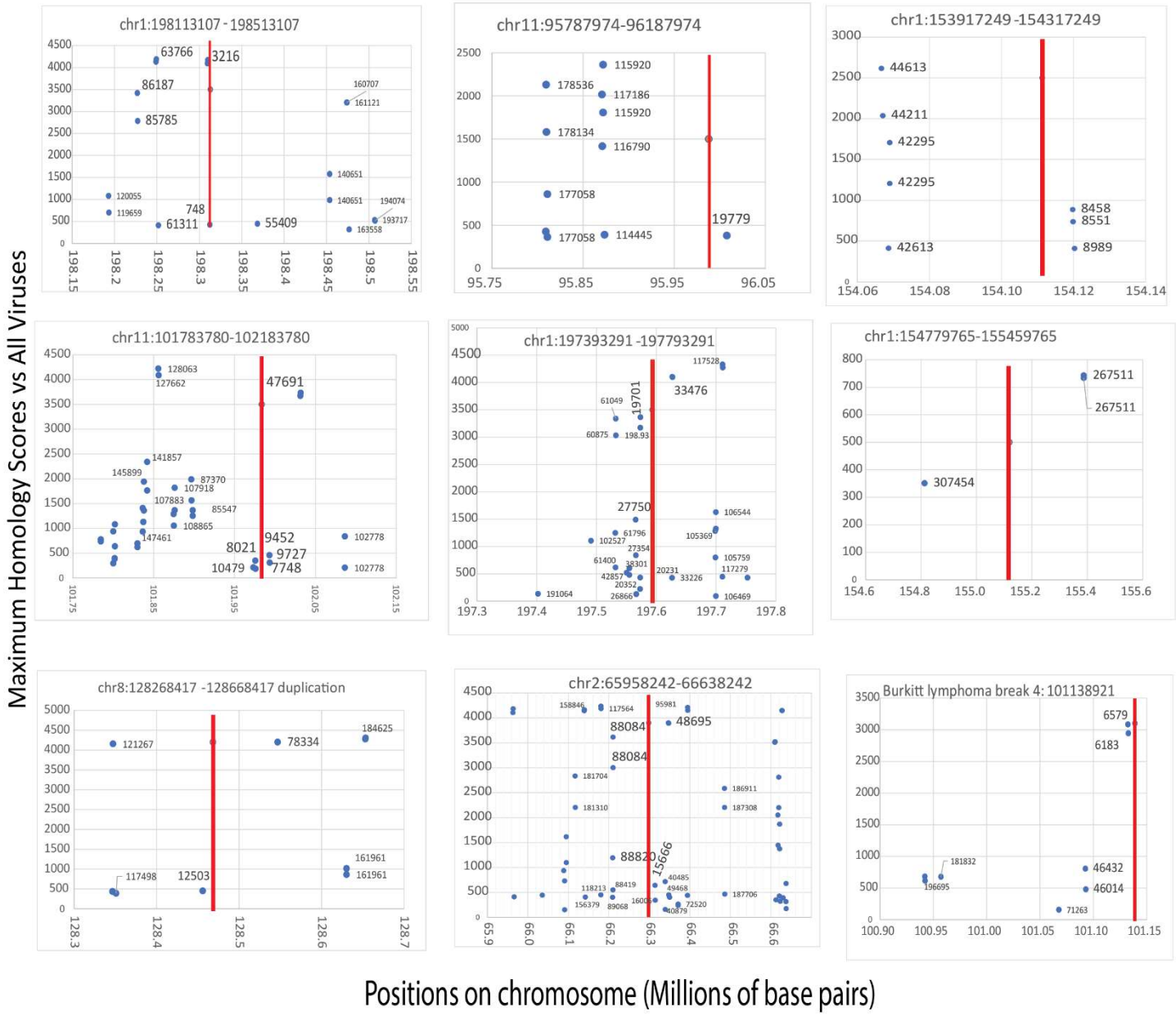
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1



Supplementary Figure 1.

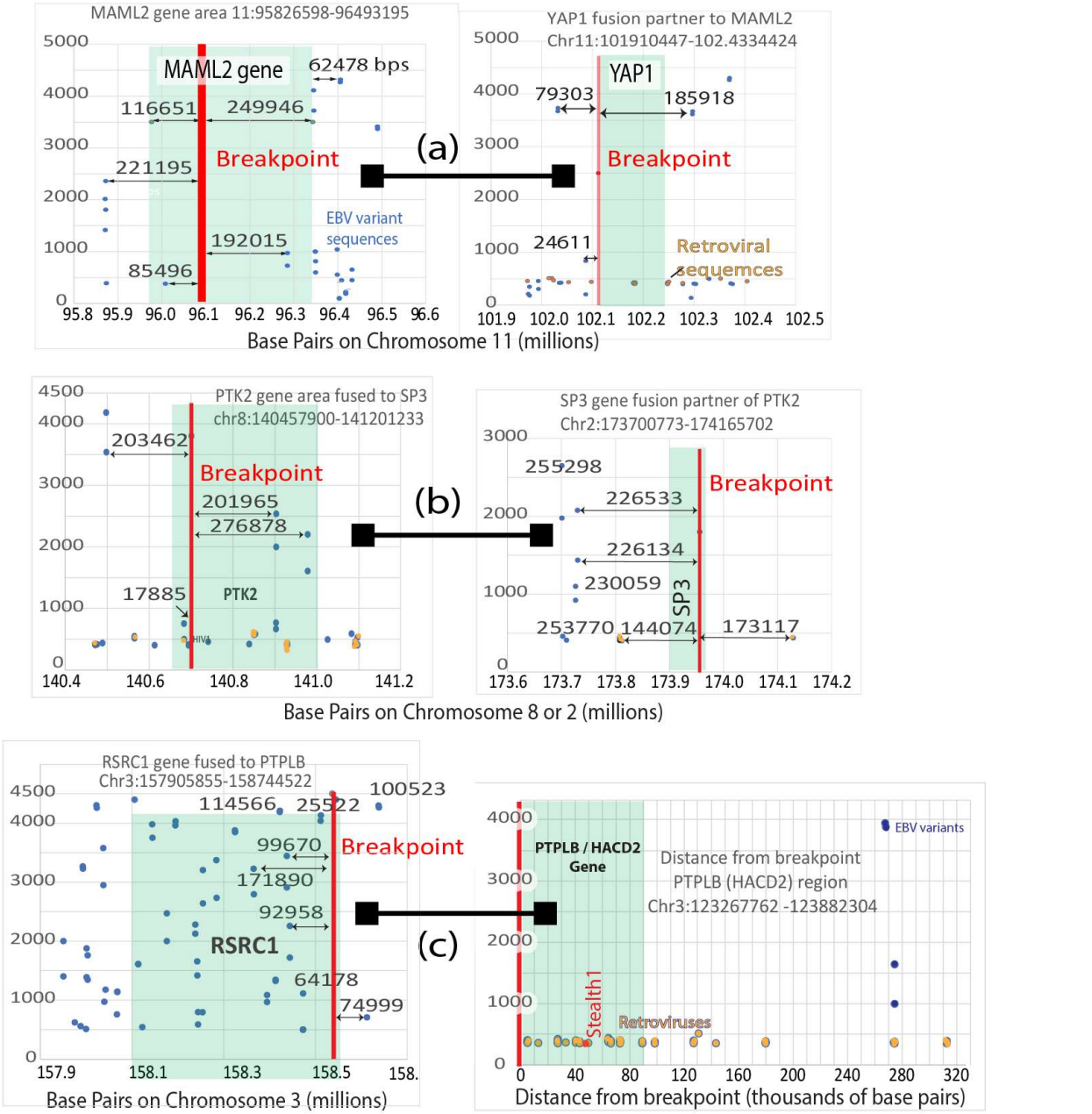
Fig. S1. Distribution of breakpoints involved in inter-chromosomal translocations in sporadic breast cancers. Cancers were selected at random from the list of 560 breast cancers cited in the references. One sporadic cancer (13426) had no breaks of any type, another 8982 had no inter-strand translocations, and 18728 had 2. Inter-chromosome translocation breakpoints in reasonably near telomeres are common. Clustering of breaks on chromosomes 8, 11, and 17 was also consistent with Fig. 1 (see Supplementary Fig 1S).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1



Supplementary Figure S2. Distributions of EBV variant homologies around eight breakpoints (indicated by red lines) in nasopharyngeal cancer NPC-5989 [29] resemble hereditary breast cancers. Burkitt's lymphoma break coordinates are also near DNA sequences with EBV variant homologies (lower right panel). The numbers near each of the blue dots are the number of base pairs between the breakpoint and the start of the homologous sequences. The numbers on the horizontal axis are the numbers of base pairs on the chromosome in millions. The vertical axes are again the maximum homology scores vs. all viruses

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1



Supplementary Figure S3. Areas around most gene fusions in nasopharyngeal cancers resemble breast cancer chromosome break regions. Homology to EBV tumor variants (blue dots) is similar to that found near breast cancer chromosome breaks. Gene fusions are shown as left to right pairs (a) MAML2-YAP1, (b) PTK2-SP3, and (c) RSRC1-PTPLB). Maximum homology scores are plotted against chromosome locations for the first five panels. The PTLB panel (lower right) graphs maximum homology vs. absolute values for distance from viral homologies to the breakpoint. Orange dots indicate the start point of retroviral homologies.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 July 2021 doi:10.20944/preprints202107.0430.v1

Supplementary Table S1 Absence of inverted repeats at breakpoints in BRCA2 associated breast cancers. Chromosome coordinates for breaks vs nearby unbroken sequences were assayed for repeats within 100 base pairs in either direction using "RepeatAround".

Status	Breakpoint or non-Breakpoint	Direct Repeat	Inverted Repeat	Mirror Repeat	Complementary Repeat
No Breaks	1:105993316	2 (1 8bps, 1 9bps)	0	0	0
Breaks	1:102731470	1 (8 bps)	0	0	0
Breaks	1:104326329	1	0	0	0
Breaks	1:145685562	1	0	0	0
No Breaks	1:143999000	2 (8,10bps)	0	2 (8 bps)	0
No Breaks	2:23000000	4 (8,8,10,10 bps)	0	0	0
Break	2: 25,505,554	1 (10 bps)	0	0	0
Breaks	2:100,035,528	1(10 bps)	0	0	0
Breaks	4:101,009,819	2(8,8 bps)	0	0	0
No breaks	4:102,639,952	2(9, 20 bps)	0	0	0
Breaks	8:80,687,247	2 (8,14 bps)	0	0	0
No Breaks	8:83,500,000	2 (8,8 bps)	0	0	0
Breaks	11:94,386,526	2(9, 13 bps)	0	0	0
No Breaks	11:91,962,848	6(8,8,8,9,12)	0	0	0
Breaks	12:88,191,644	1(14 bps)	0	0	0
No Breaks	12:90,932,271	1(12 bps)	0	0	0