Published in final edited form as:

Breast Cancer Res Treat. 2017 December ; 166(3): 937–949. doi:10.1007/s10549-017-4469-0.

Panel sequencing of 264 candidate susceptibility genes and segregation analysis in a cohort of non-*BRCA1*, non-*BRCA2* breast cancer families

Jun Li¹, Hongyan Li², Igor Makunin^{1,3}, kConFab Investigators^{4,5}, Bryony A. Thompson^{2,6}, Kayoko Tao², Erin L. Young², Jacqueline Lopez², Nicola J. Camp², Sean V. Tavtigian², Esther M. John^{7,8}, Irene L. Andrulis^{9,10}, Kum Kum Khanna¹, David Goldgar^{2,11}, and Georgia Chenevix-Trench¹

¹.QIMR Berghofer, Brisbane, QLD 4006, Australia

².Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT 84132, USA

³.Research Computing Centre, The University of Queensland, St Lucia, QLD 4072, Australia

⁴.kConFab, Research Department, Peter MacCallum Cancer Center, Melbourne, Australia

⁵.The Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Australia

⁶.Centre for Epidemiology and Biostatistics, School of Population and Global Health, University of Melbourne, Melbourne, VIC 3000, Australia

⁷.Department of Epidemiology, Cancer Prevention Institute of California, Fremont, CA 94538, USA

⁸.Department of Health Research and Policy (Epidemiology) and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94305, USA

⁹.Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, ON M5G 1X5, Canada

¹⁰.Department of Molecular Genetics, University of Toronto, Toronto, ON M5G 1X5, Canada

¹¹.Department of Dermatology, University of Utah School of Medicine, Salt Lake City, UT 84132, USA

Abstract

Purpose—The main aim of this study was to screen epigenetic modifier genes and known breast cancer driver genes for germline mutations in non-*BRCA1/2* (BRCAx) breast cancer families in order to identify novel susceptibility genes of moderate-high penetrance.

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

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Electronic supplementary material The online version of this article (doi:10.1007/s10549-017-4469-0) contains supplementary material, which is available to authorized users.

Methods—We screened 264 candidate susceptibility genes in 656 index cases from non-*BRCA1/2* families. Potentially pathogenic candidate mutations were then genotyped in all available family members for the assessment of co-segregation of the variant with disease in the family in order to estimate the breast cancer risks associated with these mutations. For 11 of the candidate susceptibility genes, we screened an additional 800 non-*BRCA1/2* breast cancer cases and 787 controls.

Results—Only two genes, *CHD8* and *USH2A* showed any evidence of an increased risk of breast cancer (RR = 2.40 (95% CI 1.0–7.32) and 2.48 (95% CI 1.11–6.67), respectively).

Conclusions—We found no convincing evidence that epigenetic modifier and known breast cancer driver genes carry germline mutations that increase breast cancer risk. *USH2A* is no longer regarded as a breast cancer driver gene and seems an implausible candidate given its association with Usher syndrome. However, somatic mutations in *CHD8* have been recently reported, making it an even more promising candidate, but further analysis of *CHD8* in very large cohorts of families or case-control studies would be required to determine if it is a moderate-risk breast cancer susceptibility gene.

Keywords

Breast cancer; Mutation; Segregation; Susceptibility gene; Chromatin modification genes

Introduction

Breast cancer is the most common cancer and the second leading cause of cancer death in women. Risk increases proportionally with the number of affected relatives with breast cancer and younger age at diagnosis [1]. Highly penetrant mutations in *BRCA1* and *BRCA2* account <40% of all multiple-case breast cancer families [2], and a small number of families are due to moderate-high penetrance genes in *TP53*, *PALB2*, *ATM*, *CHEK2*, *PTEN*, *STK11*, *CDH1* and *NFI* [3]. In addition, over 170 breast cancer risk loci have been identified by genome-wide association studies, which explain ~ 18% of the familial relative risk of breast cancer (Michailidou et al., submitted). The “missing” familial relative risk could be explained by extreme poly-genicity, variants not tagged by genome-wide associations such as short tandem repeats or rare mutations in another gene (or genes) conferring a moderately increased risk of breast cancer [4]. In the past decade, intensive efforts have been made to search for additional breast cancer predisposition genes, with limited success [5, 6].

There is increasing evidence that germline and somatic mutations in cancer target the same genes and pathways [7–9]. Whole-exome sequencing studies have identified somatic driver genes involved in DNA methylation and chromatin modification in a number of different cancers, including breast cancer [10–16]. The main aim of this study was to screen epigenetic modifier genes, including human homologs of modifiers of murine metastable epialleles (MOMME) genes which are modifiers of epigenetic silencing [17–20], for germline mutations in non-*BRCA1/2* (BRCAx) breast cancer families. Another reason for selecting this class of candidate genes was that we have previously reported that *BRCA1*-associated breast tumours have a distinct methylome, suggesting that *BRCA1* may be involved, directly or indirectly, in de novo methylation [21]. A possible mechanism for the

relationship between *BRCA1* mutation status and the methylome is suggested by the finding that *BRCA1* binds to Oct1-binding sites in the *DNMT1* promoter, thereby transcriptionally regulating this maintenance methylation enzyme [22]. Other breast cancer susceptibility genes may share a similar mechanism, thereby altering the methylome. Furthermore, we previously found that BRCAx tumours show two different patterns of methylation, suggesting that mutations in human epigenetic modifying genes may account for some breast cancer families, which share a methylation profile in their tumours. Rare families have been described in which germline epimutations are associated with colorectal cancer [23], but this phenomenon has not been reported in breast cancer families.

We screened 264 candidate susceptibility genes in 656 index cases from non-*BRCA1/2* families. Potentially pathogenic candidate mutations were then genotyped in all available family members for the assessment of co-segregation of the variant with disease in the family in order to estimate the breast cancer risks associated with these mutations. For 11 of the candidate susceptibility genes, we screened an additional 800 non-*BRCA1/2* breast cancer cases and 787 controls.

Methods

Patient cohort

We selected 656 non-*BRCA1/2* families from the Kathleen Cuninghame Foundation Consortium for Research into Familial Breast cancer (kConFab) [24] using the following criteria: (i) no known pathogenic mutation in *BRCA1* or *BRCA2* in any member of the family, and (ii) no known protein-truncating mutation in *PALB2* or the *ATM* V2424G mutation. In addition, we prioritized family selection (i) on the basis of age of diagnosis of the individual to be sequenced, (ii) including any family with a case of pancreatic cancer ($n = 94$), (iii) families who had previously had the most sensitive and complete testing of *BRCA1* and *BRCA2*, (iv) the availability of at least two germline DNA samples from related family members, and (v) ranking based on the probability calculated by BOADICEA [25] of carrying a *BRCA1/2* mutation. Families with multiple cases of breast or ovarian cancer are eligible for recruitment into kConFab under several different criteria [24]. These 656 families had an average of 4.9 breast cancer cases per family (with a range of 0–19 breast cancer cases, and 1–5 ovarian cancer cases). From these families, we selected the youngest breast cancer case for whom germline DNA was available as the index case for sequencing. We also selected 800 early-onset breast cancer cases (mean age of diagnosis 45.5 years; range 20–62) and 787 controls with germline DNA from the Breast Cancer Family Registry [26] (700 cases and 589 controls) and the University of Utah (100 cases and 198 controls). Cases with previously identified known pathogenic variants in *BRCA1* or *BRCA2* were excluded.

Exome sequencing

We performed exome sequencing of germline DNA from 12 members of five kConFab families. The libraries were prepared using NimbleGen SeqCap EZ exome capture kit v.2 and sequenced on the Illumina HiSeq 2000 machine. Paired-end reads were aligned to hg19 genome using BWA.[27] Alignments were de-duplicated with Picard and realigned and

recalibrated using GATK tools [28]. The variants were identified using UnifiedGenotyper followed by variant quality score recalibration and variant filtering according to GATK protocol. Variants were annotated with ANNOVAR and 1000 Genomes data (May 2011). We identified loss of function (LoF; nonsense, at canonical splice sites, and frameshift indels) variants, and all variants shared between fourth- or fifth-degree relatives within each family. We filtered out variants present in the November 2010 release of the 1000 Genomes Project and dbSNP v.132, in UTRs and ncRNA, overlapping known indels, and variants in genes with known LoF variants in the human population [29]. A LoF mutation in *BRCA2* gene was identified in one sample and so this family was excluded from further analysis. For the capture experiment, we selected 26 genes with variants shared between distant relatives in four families, and 79 genes with LoF variants identified in 11 samples.

Candidate gene selection

We selected a list of 103 genes involved in epigenetic regulation (59 selected from the literature—Group 1; 44 using the GO terms ‘chromatin organization’ or ‘chromatin reorganization’—Group 2), and supplemented it with 105 candidate genes from whole-exome sequencing of four non-*BRCA1/2* families (Group 3; 79 loss of function and 26 variants shared between distant family members) and 56 additional breast cancer driver genes (Groups 4 and 5 [30, 31]) known at the time (Table 1).

Targeted sequencing and selection of putative mutations

We performed custom-designed targeted sequencing covering the coding exons of all 264 candidate genes (Table 1). Targets were captured using the Agilent Target Enrichment kit by Axeq Technologies Inc. according to the manufacturer’s instructions. For the 11 candidate genes (*ATR*, *AURKC*, *BAZ1B*, *CASP8*, *CHD5*, *DNMT3A*, *FSHR*, *KDM5B*, *MAP3KI*, *RBL1*, *TET2*) screened in the additional cases and controls, we captured the targets using the Roche SeqCap EZ Choice Library (cat# 06266339001) and Seq-Cap EZ Reagent Kit Plus v2 (NimbleGen #06-953-247-001) according to the manufacturer’s instructions. All sequencing were done by 100 bp paired-end reads on the Illumina HiSeq 2000 platform. High-quality sequencing data with an average depth of 359.4-fold (ranging from 31.97 to 1215.27 per exon) were obtained. Sequences were aligned to human reference genome (GRCh37) using BWA [27]. Reads marked as PCR or optical duplicates were removed from consideration by Picard (<http://picard.sourceforge.net>). Variants were called from the targeted exons, and also from the off-target reads up to 25 nucleotides into the introns to capture potential splice site mutations. Single nucleotide variants (SNVs) and small indels were called by GATK [28]. After calling, to eliminate possible falsepositive variants, we applied strict quality control filtering of excluding variants with <30 reads or variants with reference: alternative reads ratio <0.2 or >5.0. We used ANNOVAR [32] to annotate the functional consequences of variants, as well as to retrieve allele frequencies from the public databases (dbSNP137, 1000 Genomes Project April 2012 release, complete genomes, and NHLBI-ESP 6500 exomes). Variants were removed from our analysis if they have been reported in those public databases. The potential functional importance of non-synonymous variants was predicted by the following tools: SIFT [33], Polyphen2 [34], LRT [35], MutationTaster [36] and CADD [37]. These tools are not totally independent as most include the analysis of evolutionary conservation. Any non-synonymous variant that was described

by at least three of SIFT [33], Polyphen2 [34], LRT [35], or MutationTaster [36] as damaging or probably damaging, or had a CADD score >15, was regarded as putatively deleterious, and selected for segregation and casey-control analysis. Protein-truncating mutations within the last exon of each gene were removed from our candidate list since they are predicted not to activate nonsense-mediated mRNA decay (NMD). The impact on splicing was predicted for intronic variants within 25 bp from the exon/intron boundary using stand alone perl scripts of MaxEnt [38] (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html). Intronic variants were classified as putatively pathogenic if they fulfilled the following criteria: (a) the predicted value of either 5' or 3' splicing motif was >3, and (b) the variation between MaxEnt values of variant and wildtype sequences was 30%.

Family sequencing and segregation analysis

Segregation analysis for all putative pathogenic mutations was carried out by either Sanger sequencing, using the standard sequencing procedure of BigDye Terminator v3.1 Cycle Sequencing Kit in the PE Applied Biosystem (PE Applied Biosystem), or by Fluidigm 96.96 Dynamic Array integrated fluidic circuits (IFCs) on the BioMark™ HD System (Fluidigm Corporation) according to the manufacturer's instructions, or by iPLEX genotyping [39] on the MassARRAY system (Agena Bioscience). The primers can be provided on request.

Statistical analyses

We calculated the penetrance for breast cancer in carriers of the susceptibility genes assuming a constant relative risk with age. Models were fitted under maximum likelihood theory using a modified version of the LINKAGE genetic analysis package [40]. Non-carriers were assumed to be at population risks specific to Australia with incidence rates taken from cancer registry data obtained from Cancer Incidence in Five Continents, VIII (IARC, Lyon). We estimated risk ratios (RR, the age-specific breast cancer incidence rate in carriers divided by the relevant population rate). Reported breast cancers with unknown age at diagnosis were excluded from all analyses. Individuals with cancers other than breast (including ovarian cancer) were treated as unaffected at the age of their cancer diagnosis. To test for an association of pooled variants in a gene with breast cancer risk in the case-control study, we estimated odds ratios (OR) and corresponding 95% CIs based on the Fisher's exact test.

Results

We identified 420 putative pathogenic mutations in 313 index cases using the following criteria: (a) not reported in any of these public databases (dbSNP137, 1000 Genomes Project April 2012 release, Complete Genomes, and NHLBI-ESP 6500 exomes); (b) not present in more than five index cases, since these were not likely to be high risk variants; (c) protein-truncating mutations (PTC, including frameshift insertion/deletion and stopgain) not in the last exon; (d) missense variants predicted as "deleterious" by three out of four tools, or with a CADD score >15, and were present in genes in which we also identified PTCs; (e) intronic variants called from ± 25 bp region of the exon-intron boundary had to be predicted by MaxEntScan to possibly create a new splice site and in genes in which we also identified PTCs or predicted deleterious missense mutations. From these 313 families, there were 2178

germline DNAs from family members available for segregation analysis. Of the 420 putative pathogenic mutations identified, we could not design assays for 36, and failed to validate 9 variants in the index case. However, the remaining 375 variants were successfully validated by Fluidigm or iPLEX and 259 families (some of which had more than one putatively pathogenic variant), were informative for segregation analysis (Supplementary Table 1).

In the single gene analysis, assuming the constant relative risk with age, only two genes, *CHD8* and *USH2A* showed any evidence of an increased risk of breast cancer (RR = 2.40 (95% CI 1.0–7.32) and 2.48 (95% CI 1.11–6.67), respectively) (Table 2). We identified seven families with putative pathogenic mutations in *CHD8* (six missense variants and one in-frame deletion) and 10 families with putative pathogenic mutations in *USH2A* (two protein truncating, seven missense and one in-frame deletion). We also performed group analyses for Groups 1 and 2, Group 3, and Groups 4 and 5, but this did not identify any groups in which variants were associated with breast cancer risk (Table 3). In the case–control screening, we identified a small number of putatively pathogenic mutations, but found no evidence of an association with breast cancer risk.

Discussion

We carried out a large screen of candidate BRCAx genes, mostly epigenetic modifier genes and known breast cancer driver genes, by analysing 264 genes in 656 multiple-case families, followed by segregation analysis of 375 putative, pathogenic missense, splice site and truncating variants in 259 informative families. One of the limitations of this study is that since we began the project, the list of breast cancer driver genes has been extended [16, 41–43], so there are at least 90 known driver genes that we did not screen. We found some evidence that mutations in *CHD8* (seven families) and *USH2A* (10 families) are associated with breast cancer risk. These genes were selected as candidates as an epigenetic modifier [44] and a putative cancer driver [45], respectively. Mutations in *USH2A* cause Usher syndrome type IIa and retinitis pigmentosa. *USH2A* is no longer considered a breast cancer driver gene [16, 41–43], but has been shown to have a high false discovery rate in mutation calling [46]. *CHD8* encodes a member of the chromodomain-helicase-DNA binding protein family and plays a role in transcriptional regulation, epigenetic remodelling, promotion of cell proliferation, and regulation of RNA synthesis [47]. Mutations in *CHD8* have been recently reported in over 5% of breast tumours [46], making it a very plausible candidate for breast cancer susceptibility. However, the *p* value for testing the association between variants in *CHD8* and breast cancer risk is only nominally significant ($p = 0.05$), and given that we tested 66 genes, clearly would not be significant after correction for multiple testing. Eleven of the genes, *ATR*, *AURKC*, *BAZ1B*, *CASP8*, *CHD5*, *DNMT3A*, *FSHR*, *KDM5B*, *MAP3KI*, *RBL1* and *TET2*, were also screened in approximately 800 cases and controls, but none were found to be associated with breast cancer risk. In conclusion, we have some evidence that mutations in *CHD8* might be associated with an elevated risk of breast cancer, warranting further investigation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We wish to thank Jonathan Ellis for help with variant calling and Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow-Up Study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab. kConFab is supported by a Grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. This project was funded by the Susan G. Komen Foundation, the NHMRC and NIH Grant R01 CA155767 to DEG & SVT and R01 CA163353 to NJC. The Breast Cancer Family Registry (BCFR) was supported by Grant UM1 CA164920 from the USA National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centres in the BCFR, nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR.

References

1. Lalloo F, Evans DG (2012) Familial breast cancer. *Clin Genet* 82(2):105–114. doi:10.1111/j.1399-0004.2012.01859.x [PubMed: 22356477]
2. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J, Sobol H, Teare MD, Struewing J, Arason A, Scherneck S, Peto J, Rebbeck TR, Tonin P, Neuhausen S, Barkardottir R, Eyfjord J, Lynch H, Ponder BA, Gayther SA, Zelada-Hedman M et al. (1998) Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 62(3): 676–689 [PubMed: 9497246]
3. Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, Chenevix-Trench G, Rahman N, Robson M, Domchek SM, Foulkes WD (2015) Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* 372(23):2243–2257. doi:10.1056/NEJMs1501341 [PubMed: 26014596]
4. Wen H, Kim YC, Snyder C, Xiao F, Fleissner EA, Becirovic D, Luo J, Downs B, Sherman S, Cowan KH, Lynch HT, Wang SM (2014) Family-specific, novel, deleterious germline variants provide a rich resource to identify genetic predispositions for BRCAx familial breast cancer. *BMC cancer* 14:470. doi:10.1186/1471-2407-14-470 [PubMed: 24969172]
5. Sun J, Wang Y, Xia Y, Xu Y, Ouyang T, Li J, Wang T, Fan Z, Fan T, Lin B, Lou H, Xie Y (2015) Mutations in RECQL gene are associated with predisposition to breast cancer. *PLoS Genet* 11(5):e1005228. doi:10.1371/journal.pgen.1005228 [PubMed: 25945795]
6. Cybulski C, Carrot-Zhang J, Kluzniak W, Rivera B, Kashyap A, Wokolorczyk D, Giroux S, Nadaf J, Hamel N, Zhang S, Huzarski T, Gronwald J, Byrski T, Swiec M, Jakubowska A, Rudnicka H, Lener M, Masojc B, Tonin PN, Rousseau F, Gorski B, Debniak T, Majewski J, Lubinski J, Foulkes WD, Narod SA, Akbari MR (2015) Germline RECQL mutations are associated with breast cancer susceptibility. *Nat Genet* 47(6):643–646. doi:10.1038/ng.3284 [PubMed: 25915596]
7. Carter JM, O'Hara C, Dundas G, Gilchrist D, Collins MS, Eaton K, Judkins AR, Biegel JA, Folpe AL (2012) Epithelioid malignant peripheral nerve sheath tumor arising in a schwannoma, in a patient with “neuroblastoma-like” schwannomatosis and a novel germline SMARCB1 mutation. *Am J Surg Pathol* 36(1):154–160. doi: 10.1097/PAS.0b013e3182380802 [PubMed: 22082606]
8. Glodzik D, Morganella S, Davies H, Simpson PT, Li Y, Zou X, Diez-Perez J, Staaf J, Alexandrov LB, Smid M, Brinkman AB, Rye IH, Russnes H, Raine K, Purdie CA, Lakhani SR, Thompson AM, Birney E, Stunnenberg HG, van de Vijver MJ, Martens JW, Borresen-Dale AL, Richardson AL, Kong G, Viari A, Easton D, Evan G, Campbell PJ, Stratton MR, Nik-Zainal S (2017) A somatic-mutational process recurrently duplicates germline susceptibility loci and tissue-specific super-enhancers in breast cancers. *Nat Genet* 49(3):341–348. doi:10.1038/ng.3771 [PubMed: 28112740]
9. Kanchi KL, Johnson KJ, Lu C, McLellan MD, Leiserson MD, Wendl MC, Zhang Q, Koboldt DC, Xie M, Kandoth C, McMichael JF, Wyczalkowski MA, Larson DE, Schmidt HK, Miller CA, Fulton RS, Spellman PT, Mardis ER, Druley TE, Graubert TA, Goodfellow PJ, Raphael BJ, Wilson RK, Ding L (2014) Integrated analysis of germline and somatic variants in ovarian cancer. *Nat Commun* 5:3156. doi:10.1038/ncomms4156 [PubMed: 24448499]

10. Liao L, Testa JR, Yang H (2015) The roles of chromatin-remodelers and epigenetic modifiers in kidney cancer. *Cancer Genet* 208(5):206–214. doi:10.1016/j.cancergen.2015.02.008 [PubMed: 25873528]
11. Jhanwar SC (2015) Genetic and epigenetic pathways in myelodysplastic syndromes: a brief overview. *Adv Biol Regul* 58:28–37. doi:10.1016/j.jbior.2014.11.002 [PubMed: 25499150]
12. Cazzola M, Della Porta MG, Malcovati L (2013) The genetic basis of myelodysplasia and its clinical relevance. *Blood* 122(25):4021–4034. doi:10.1182/blood-2013-09-381665 [PubMed: 24136165]
13. Lee RS, Roberts CW (2013) Rhabdoid tumors: an initial clue to the role of chromatin remodeling in cancer. *Brain Pathol* 23(2):200–205. doi:10.1111/bpa.12021 [PubMed: 23432645]
14. Ratnakumar K, Bernstein E (2013) ATRX: the case of a peculiar chromatin remodeler. *Epigenetics* 8(1):3–9. doi:10.4161/epi.23271 [PubMed: 23249563]
15. Koschmann C, Nunez FJ, Mendez F, Brosnan-Cashman JA, Meeker AK, Lowenstein PR, Castro MG (2017) Mutated chromatin regulatory factors as tumor drivers in cancer. *Cancer Res* 77(2):227–233. doi:10.1158/0008-5472.CAN-16-2301 [PubMed: 28062403]
16. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC, Van Loo P, Ju YS, Smid M, Brinkman AB, Morganella S, Aure MR, Lingjaerde OC, Langerod A, Ringner M, Ahn SM, Boyault S, Brock JE, Broeks A, Butler A, Desmedt C, Dirix L, Dronov S, Fatima A, Foekens JA, Gerstung M, Hooijer GK, Jang SJ, Jones DR, Kim HY, King TA, Krishnamurthy S, Lee HJ, Lee JY, Li Y, McLaren S, Menzies A, Mustonen V, O'Meara S, Pauporte I, Pivrot X, Purdie CA, Raine K, Ramakrishnan K, Rodriguez-Gonzalez FG, Romieu G, Sieuwerts AM, Simpson PT, Shepherd R, Stebbings L, Stefansson OA, Teague J, Tommasi S, Treilleux I, Van den Eynden GG, Vermeulen P, Vincent-Salomon A, Yates L, Caldas C, van't Veer L, Tutt A, Knappskog S, Tan BK, Jonkers J, Borg A, Ueno NT, Sotiriou C, Viari A, Futreal PA, Campbell PJ, Span PN, Van Laere S, Lakhani SR, Eyfjord JE, Thompson AM, Birney E, Stunnenberg HG, van de Vijver MJ, Martens JW, Borresen-Dale AL, Richardson AL, Kong G, Thomas G, Stratton MR (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534(7605):47–54. doi:10.1038/nature17676 [PubMed: 27135926]
17. Blewitt ME, Gendrel AV, Pang Z, Sparrow DB, Whitelaw N, Craig JM, Apedaile A, Hilton DJ, Dunwoodie SL, Brockdorff N, Kay GF, Whitelaw E (2008) SmcHd1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* 40(5):663–669. doi:10.1038/ng.142 [PubMed: 18425126]
18. Ashe A, Morgan DK, Whitelaw NC, Bruxner TJ, Vickaryous NK, Cox LL, Butterfield NC, Wicking C, Blewitt ME, Wilkins SJ, Anderson GJ, Cox TC, Whitelaw E (2008) A genome-wide screen for modifiers of transgene variegation identifies genes with critical roles in development. *Genome Biol* 9(12):R182. doi:10.1186/gb-2008-9-12-r182 [PubMed: 19099580]
19. Whitelaw NC, Chong S, Morgan DK, Nestor C, Bruxner TJ, Ashe A, Lambley E, Meehan R, Whitelaw E (2010) Reduced levels of two modifiers of epigenetic gene silencing, Dnmt3a and Trim28, cause increased phenotypic noise. *Genome Biol* 11(11):R111. doi:10.1186/gb-2010-11-11-r111 [PubMed: 21092094]
20. Daxinger L, Oey H, Apedaile A, Sutton J, Ashe A, Whitelaw E (2012) A forward genetic screen identifies eukaryotic translation initiation factor 3, subunit H (eIF3h), as an enhancer of variegation in the mouse. *G3 (Bethesda)* 2(11):1393–1396. doi:10.1534/g3.112.004036 [PubMed: 23173090]
21. Flanagan JM, Cocciardi S, Waddell N, Johnstone CN, Marsh A, Henderson S, Simpson P, da Silva L, kConFab I, Khanna K, Lakhani S, Boshoff C, Chenevix-Trench G (2010) DNA methylome of familial breast cancer identifies distinct profiles defined by mutation status. *Am J Hum Genet* 86(3):420–433. doi:10.1016/j.ajhg.2010.02.008 [PubMed: 20206335]
22. Shukla V, Coumoul X, Lahusen T, Wang RH, Xu X, Vassilopoulos A, Xiao C, Lee MH, Man YG, Ouchi M, Ouchi T, Deng CX (2010) BRCA1 affects global DNA methylation through regulation of DNMT1. *Cell Res* 20(11):1201–1215. doi:10.1038/cr.2010.128 [PubMed: 20820192]
23. Hitchins MP, Wong JJ, Suthers G, Suter CM, Martin DI, Hawkins NJ, Ward RL (2007) Inheritance of a cancer-associated MLH1 germ-line epimutation. *N Engl J Med* 356(7):697–705. doi:10.1056/NEJMoa064522 [PubMed: 17301300]

24. Mann GJ, Thorne H, Balleine RL, Butow PN, Clarke CL, Edkins E, Evans GM, Fereday S, Haan E, Gattas M, Giles GG, Goldblatt J, Hopper JL, Kirk J, Leary JA, Lindeman G, Niedermayr E, Phillips KA, Picken S, Pupo GM, Saunders C, Scott CL, Spurdle AB, Suthers G, Tucker K, Chenevix-Trench G, Kathleen Cunningham Consortium for Research in Familial Breast C (2006) Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Res BCR* 8(1):R12. doi:10.1186/bcr1377 [PubMed: 16507150]
25. Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, Narod SA, Risch HA, Eyfjord JE, Hopper JL, Southey MC, Olsson H, Johannsson O, Borg A, Pasini B, Radice P, Manoukian S, Eccles DM, Tang N, Olah E, Anton-Culver H, Warner E, Lubinski J, Gronwald J, Gorski B, Tryggvadottir L, Syrjakoski K, Kallioniemi OP, Eerola H, Nevanlinna H, Pharoah PD, Easton DF (2008) The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* 98(8):1457–1466. doi:10.1038/sj.bjc.6604305 [PubMed: 18349832]
26. John EM, Hopper JL, Beck JC, Knight JA, Neuhausen SL, Senie RT, Ziogas A, Andrulis IL, Anton-Culver H, Boyd N, Buys SS, Daly MB, O'Malley FP, Santella RM, Southey MC, Venne VL, Venter DJ, West DW, Whittemore AS, Seminara D, Breast Cancer Family R (2004) The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Research: Bcr* 6(4):R375–R389. doi:10.1186/bcr801 [PubMed: 15217505]
27. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760. doi:10.1093/bioinformatics/btp324 [PubMed: 19451168]
28. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303. doi:10.1101/gr.107524.110 [PubMed: 20644199]
29. MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, Albers CA, Zhang ZD, Conrad DF, Lunter G, Zheng H, Ayub Q, DePristo MA, Banks E, Hu M, Handsaker RE, Rosenfeld JA, Fromer M, Jin M, Mu XJ, Khurana E, Ye K, Kay M, Saunders GI, Suner MM, Hunt T, Barnes IH, Amid C, Carvalho-Silva DR, Bignell AH, Snow C, Yngvadottir B, Bumpstead S, Cooper DN, Xue Y, Romero IG, Wang J, Li Y, Gibbs RA, McCarroll SA, Dermitzakis ET, Pritchard JK, Barrett JC, Harrow J, Hurler ME, Gerstein MB, Tyler-Smith C, Genomes Project C (2012) A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 335(6070):823–828. doi:10.1126/science.1215040 [PubMed: 22344438]
30. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, Nik-Zainal S, Martin S, Varela I, Bignell GR, Yates LR, Papaemmanuil E, Beare D, Butler A, Cheverton A, Gamble J, Hinton J, Jia M, Jayakumar A, Jones D, Latimer C, Lau KW, McLaren S, McBride DJ, Menzies A, Mudie L, Raine K, Rad R, Chapman MS, Teague J, Easton D, Langerod A, Oslo Breast Cancer C, Lee MT, Shen CY, Tee BT, Huimin BW, Broeks A, Vargas AC, Turashvili G, Martens J, Fatima A, Miron P, Chin SF, Thomas G, Boyault S, Mariani O, Lakhani SR, van de Vijver M, van 't Veer L, Foekens J, Desmedt C, Sotiriou C, Tutt A, Caldas C, Reis-Filho JS, Aparicio SA, Salomon AV, Borresen- Dale AL, Richardson AL, Campbell PJ, Futreal PA, Stratton MR (2012) The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486(7403):400–404. doi: 10.1038/nature11017 [PubMed: 22722201]
31. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, Turashvili G, Ding J, Tse K, Haffari G, Bashashati A, Prentice LM, Khattra J, Burleigh A, Yap D, Bernard V, McPherson A, Shumansky K, Crisan A, Giuliany R, Heravi-Moussavi A, Rosner J, Lai D, Birol I, Varhol R, Tam A, Dhalla N, Zeng T, Ma K, Chan SK, Griffith M, Moradian A, Cheng SW, Morin GB, Watson P, Gelmon K, Chia S, Chin SF, Curtis C, Rueda OM, Pharoah PD, Damaraju S, Mackey J, Hoon K, Harkins T, Tadigotla V, Sigaroudinia M, Gascard P, Tlsty T, Costello JF, Meyer IM, Eaves CJ, Wasserman WW, Jones S, Huntsman D, Hirst M, Caldas C, Marra MA, Aparicio S (2012) The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* 486(7403):395–399. doi:10.1038/nature10933 [PubMed: 22495314]
32. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38(16):e164. doi:10.1093/nar/gkq603 [PubMed: 20601685]

33. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4(7):1073–1081. doi:10.1038/nprot.2009.86 [PubMed: 19561590]
34. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7(4):248–249. doi:10.1038/nmeth0410-248 [PubMed: 20354512]
35. San Lucas FA, Wang G, Scheet P, Peng B (2012) Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. *Bioinformatics* 28(3):421–422. doi:10.1093/bioinformatics/btr667 [PubMed: 22138362]
36. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 7(8):575–576. doi:10.1038/nmeth0810-575 [PubMed: 20676075]
37. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46(3):310–315. doi:10.1038/ng.2892 [PubMed: 24487276]
38. Yeo G, Burge CB (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol J Comput Mol Cell Biol* 11(2–3):377–394. doi:10.1089/1066527041410418
39. Johnatty SE, Spurdle AB, Beesley J, Chen X, Hopper JL, Duffy DL, Chenevix-Trench G, Kathleen Cunningham Consortium for Research in Familial Breast C (2008) Progesterone receptor polymorphisms and risk of breast cancer: results from two Australian breast cancer studies. *Breast Cancer Res Treat* 109(1):91–99. doi:10.1007/s10549-007-9627-3 [PubMed: 17592773]
40. Lathrop GM, Lalouel JM, Julier C, Ott J (1985) Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am J Hum Genet* 37(3):482–498 [PubMed: 3859205]
41. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C, Bowlby R, Shen H, Hayat S, Fieldhouse R, Lester SC, Tse GM, Factor RE, Collins LC, Allison KH, Chen YY, Jensen K, Johnson NB, Oesterreich S, Mills GB, Cherniack AD, Robertson G, Benz C, Sander C, Laird PW, Hoadley KA, King TA, Network TR, Perou CM (2015) Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* 163(2):506–519. doi:10.1016/j.cell.2015.09.033 [PubMed: 26451490]
42. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, Pugh M, Jones L, Russell R, Sammut SJ, Tsui DW, Liu B, Dawson SJ, Abraham J, Northen H, Peden JF, Mukherjee A, Turashvili G, Green AR, McKinney S, Oloumi A, Shah S, Rosenfeld N, Murphy L, Bentley DR, Ellis IO, Purushotham A, Pinder SE, Borresen-Dale AL, Earl HM, Pharoah PD, Ross MT, Aparicio S, Caldas C (2016) The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* 7:11479. doi:10.1038/ncomms11479 [PubMed: 27161491]
43. Cancer Genome Atlas N (2012) Comprehensive molecular portraits of human breast tumours. *Nature* 490(7418):61–70. doi:10.1038/nature11412 [PubMed: 23000897]
44. Ishihara K, Oshimura M, Nakao M (2006) CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol Cell* 23(5):733–742. doi:10.1016/j.molcel.2006.08.008 [PubMed: 16949368]
45. Kim N, Hong Y, Kwon D, Yoon S (2013) Somatic mutaoe profile in human cancer tissues. *Genom Inform* 11(4):239–244. doi:10.5808/GI.2013.11.4.239
46. Pongor L, Kormos M, Hatzis C, Pusztai L, Szabo A, Gyorffy B (2015) A genome-wide approach to link genotype to clinical outcome by utilizing next generation sequencing and gene chip data of 6,697 breast cancer patients. *Genome Med* 7:104. doi:10.1186/s13073-015-0228-1 [PubMed: 26474971]
47. Subtil-Rodriguez A, Vazquez-Chavez E, Ceballos-Chavez M, Rodriguez-Paredes M, Martin-Subero JL, Esteller M, Reyes JC (2014) The chromatin remodeller CHD8 is required for E2F-dependent transcription activation of S-phase genes. *Nucleic Acids Res* 42(4):2185–2196. doi:10.1093/nar/gkt1161 [PubMed: 24265227]
48. Jones S, Wang TL, Shih Ie M, Mao TL, Nakayama K, Roden R, Glas R, Slamon D, Diaz LA, Jr, Vogelstein B, Kinzler KW, Velculescu VE, Papadopoulos N (2010) Frequent mutations of

- chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science* 330(6001):228–231. doi:10.1126/science.1196333 [PubMed: 20826764]
49. Robinson DR, Wu YM, Vats P, Su F, Lonigro RJ, Cao X, Kalyana-Sundaram S, Wang R, Ning Y, Hodges L, Gursky A, Siddiqui J, Tomlins SA, Roychowdhury S, Pienta KJ, Kim SY, Roberts JS, Rae JM, Van Poznak CH, Hayes DF, Chugh R, Kunju LP, Talpaz M, Schott AF, Chinnaiyan AM (2013) Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet* 45(12):1446–1451. doi:10.1038/ng.2823 [PubMed: 24185510]
 50. Yan Z, Cui K, Murray DM, Ling C, Xue Y, Gerstein A, Parsons R, Zhao K, Wang W (2005) PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes. *Genes Dev* 19(14):1662–1667. doi:10.1101/gad.1323805 [PubMed: 15985610]
 51. Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, Schlick RD, Tang LH, Wolfgang CL, Choti MA, Velculescu VE, Diaz LA, Jr, Vogelstein B, Kinzler KW, Hruban RH, Papadopoulos N (2011) DAXX/ATRAX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* 331(6021):1199–1203. doi:10.1126/science.1200609 [PubMed: 21252315]
 52. Zekri A, Lesan V, Ghaffari SH, Tabrizi MH, Modarressi MH (2012) Gene amplification and overexpression of Aurora-C in breast and prostate cancer cell lines. *Oncol Res* 20(5-6):241–250 [PubMed: 23581231]
 53. Mobasher MB, Shirkoobi R, Zendehele K, Jahanzad I, Talebi S, Afsharpad M, Modarressi MH (2015) Transcriptome analysis of the cancer/testis genes, DAZ1, AURKC, and TEX101, in breast tumors and six breast cancer cell lines. *Tumour Biol* 36(10):8201–8206. doi:10.1007/s13277-015-3546-4 [PubMed: 25994570]
 54. Nishikawa H, Wu W, Koike A, Kojima R, Gomi H, Fukuda M, Ohta T (2009) BRCA1-associated protein 1 interferes with BRCA1/BARD1 RING heterodimer activity. *Cancer Res* 69(1):111–119. doi:10.1158/0008-5472.CAN-08-3355 [PubMed: 19117993]
 55. Chen X, Arciero CA, Wang C, Broccoli D, Godwin AK (2006) BRCC36 is essential for ionizing radiation-induced BRCA1 phosphorylation and nuclear foci formation. *Cancer Res* 66(10):5039–5046. doi:10.1158/0008-5472.CAN-05-4194 [PubMed: 16707425]
 56. Peng C, Zhou J, Liu HY, Zhou M, Wang LL, Zhang QH, Yang YX, Xiong W, Shen SR, Li XL, Li GY (2006) The transcriptional regulation role of BRD7 by binding to acetylated histone through bromodomain. *J Cell Biochem* 97(4):882–892. doi:10.1002/jcb.20645 [PubMed: 16265664]
 57. Kim MS, Chung NG, Kang MR, Yoo NJ, Lee SH (2011) Genetic and expressional alterations of CHD genes in gastric and colorectal cancers. *Histopathology* 58(5):660–668. doi:10.1111/j.1365-2559.2011.03819.x [PubMed: 21447119]
 58. Mullighan CG, Zhang J, Kasper LH, Lerach S, Payne-Turner D, Phillips LA, Heatley SL, Holmfeldt L, Collins-Underwood JR, Ma J, Buetow KH, Pui CH, Baker SD, Brindle PK, Downing JR (2011) CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* 471(7337):235–239. doi:10.1038/nature09727 [PubMed: 21390130]
 59. Ozdag H, Batley SJ, Forsti A, Iyer NG, Daigo Y, Boutell J, Arends MJ, Ponder BA, Kouzarides T, Caldas C (2002) Mutation analysis of CBP and PCAF reveals rare inactivating mutations in cancer cell lines but not in primary tumours. *Br J Cancer* 87(10):1162–1165. doi:10.1038/sj.bjc.6600554 [PubMed: 12402157]
 60. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M, Hovestadt V, Albrecht S, Kool M, Nantel A, Konermann C, Lindroth A, Jager N, Rausch T, Ryzhova M, Korbel JO, Hielscher T, Hauser P, Garami M, Klekner A, Bogner L, Ebinger M, Schuhmann MU, Scheurlen W, Pekrun A, Fruhwald MC, Roggendorf W, Kramm C, Durken M, Atkinson J, Lepage P, Montpetit A, Zakrzewska M, Zakrzewski K, Liberski PP, Dong Z, Siegel P, Kulozik AE, Zapatka M, Guha A, Malkin D, Felsberg J, Reifemberger G, von Deimling A, Ichimura K, Collins VP, Witt H, Milde T, Witt O, Zhang C, Castelo-Branco P, Lichter P, Faury D, Tabori U, Plass C, Majewski J, Pfister SM, Jabado N (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482(7384):226–231. doi:10.1038/nature10833 [PubMed: 22286061]
 61. Chong S, Vickaryous N, Ashe A, Zamudio N, Youngson N, Hemley S, Stopka T, Skoultschi A, Matthews J, Scott HS, de Kretser D, O'Bryan M, Blewitt M, Whitelaw E (2007) Modifiers of

- epigenetic reprogramming show paternal effects in the mouse. *Nat Genet* 39(5):614–622. doi: 10.1038/ng2031 [PubMed: 17450140]
62. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, Harris CC, Lichti CF, Townsend RR, Fulton RS, Dooling DJ, Koboldt DC, Schmidt H, Zhang Q, Osborne JR, Lin L, O’Laughlin M, McMichael JF, Delehaunty KD, McGrath SD, Fulton LA, Magrini VJ, Vickery TL, Hundal J, Cook LL, Conyers JJ, Swift GW, Reed JP, Alldredge PA, Wylie T, Walker J, Kalicki J, Watson MA, Heath S, Shannon WD, Varghese N, Nagarajan R, Westervelt P, Tomasson MH, Link DC, Graubert TA, DiPersio JF, Mardis ER, Wilson RK (2010) DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* 363(25):2424–2433. doi:10.1056/NEJMoa1005143 [PubMed: 21067377]
 63. Yan XJ, Xu J, Gu ZH, Pan CM, Lu G, Shen Y, Shi JY, Zhu YM, Tang L, Zhang XW, Liang WX, Mi JQ, Song HD, Li KQ, Chen Z, Chen SJ (2011) Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet* 43(4):309–315. doi:10.1038/ng.788 [PubMed: 21399634]
 64. Bachman KE, Rountree MR, Baylin SB (2001) Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *J Biol Chem* 276(34):32282–32287. doi:10.1074/jbc.M104661200 [PubMed: 11427539]
 65. Bryan EJ, Jokubaitis VJ, Chamberlain NL, Baxter SW, Dawson E, Choong DY, Campbell IG (2002) Mutation analysis of EP300 in colon, breast and ovarian carcinomas. *Int J Cancer* 102(2): 137–141. doi: 10.1002/ijc.10682 [PubMed: 12385008]
 66. Ernst T, Chase AJ, Score J, Hidalgo-Curtis CE, Bryant C, Jones AV, Waghorn K, Zoi K, Ross FM, Reiter A, Hochhaus A, Drexler HG, Duncombe A, Cervantes F, Oscier D, Boulwood J, Grand FH, Cross NC (2010) Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet* 42(8):722–726. doi: 10.1038/ng.621 [PubMed: 20601953]
 67. Nikoloski G, Langemeijer SM, Kuiper RP, Knops R, Massop M, Tonnissen ER, van der Heijden A, Scheele TN, Vandenbergh P, de Witte T, van der Reijden BA, Jansen JH (2010) Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nat Genet* 42(8):665–667. doi:10.1038/ng.620 [PubMed: 20601954]
 68. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, Paul JE, Boyle M, Woolcock BW, Kuchenbauer F, Yap D, Humphries RK, Griffith OL, Shah S, Zhu H, Kimbara M, Shashkin P, Charlot JF, Tcherpakov M, Corbett R, Tam A, Varhol R, Smailus D, Moksa M, Zhao Y, Delaney A, Qian H, Birol I, Schein J, Moore R, Holt R, Horsman DE, Connors JM, Jones S, Aparicio S, Hirst M, Gascoyne RD, Marra MA (2010) Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet* 42(2):181–185. doi: 10.1038/ng.518 [PubMed: 20081860]
 69. Harten SK, Bruxner TJ, Bharti V, Blewitt M, Nguyen TM, Whitelaw E, Epp T (2014) The first mouse mutants of D14Abbl1 (Fam208a) show that it is critical for early development. *Mamm Genome* 25(7–8):293–303. doi:10.1007/s00335-014-9516-0 [PubMed: 24781204]
 70. Roberts AR, Blewitt ME, Youngson NA, Whitelaw E, Chong S (2011) Reduced dosage of the modifiers of epigenetic reprogramming Dnmt1, Dnmt3L, SmcHD1 and Foxo3a has no detectable effect on mouse telomere length in vivo. *Chromosoma* 120(4):377–385. doi:10.1007/s00412-011-0318-9 [PubMed: 21553025]
 71. Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, Van Tine BA, Hoog J, Goiffon RJ, Goldstein TC, Ng S, Lin L, Crowder R, Snider J, Ballman K, Weber J, Chen K, Koboldt DC, Kandoth C, Schierding WS, McMichael JF, Miller CA, Lu C, Harris CC, McLellan MD, Wendl MC, DeSchryver K, Allred DC, Esserman L, Unzeitig G, Margenthaler J, Babiera GV, Marcom PK, Guenther JM, Leitch M, Hunt K, Olson J, Tao Y, Maher CA, Fulton LL, Fulton RS, Harrison M, Oberkell B, Du F, Demeter R, Vickery TL, Elhammali A, Piwnica-Worms H, McDonald S, Watson M, Dooling DJ, Ota D, Chang LW, Bose R, Ley TJ, Piwnica-Worms D, Stuart JM, Wilson RK, Mardis ER (2012) Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* 486(7403):353–360. doi:10.1038/nature11143 [PubMed: 22722193]
 72. Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ, Shen D, Boca SM, Barber T, Ptak J, Silliman N, Szabo S, Dezso Z, Ustyanksky V, Nikolskaya T, Nikolsky Y, Karchin R, Wilson PA, Kaminker JS, Zhang Z, Croshaw R, Willis J, Dawson D, Shipitsin M, Willson JK, Sukumar S, Polyak K, Park BH, Pethiyagoda CL, Pant PV, Ballinger DG, Sparks AB, Hartigan J, Smith DR,

- Suh E, Papadopoulos N, Buckhaults P, Markowitz SD, Parmigiani G, Kinzler KW, Velculescu VE, Vogelstein B (2007) The genomic landscapes of human breast and colorectal cancers. *Science* 318(5853):1108–1113. doi:10.1126/science.1145720 [PubMed: 17932254]
73. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, Wolniak K, Peeters JK, Liu W, Choe SE, Fantin VR, Paietta E, Lowenberg B, Licht JD, Godley LA, Delwel R, Valk PJ, Thompson CB, Levine RL, Melnick A (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* 18(6):553–567. doi: 10.1016/j.ccr.2010.11.015 [PubMed: 21130701]
74. Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, Davies H, Jones D, Lin ML, Teague J, Bignell G, Butler A, Cho J, Dalgliesh GL, Galappaththige D, Greenman C, Hardy C, Jia M, Latimer C, Lau KW, Marshall J, McLaren S, Menzies A, Mudie L, Stebbings L, Largaespada DA, Wessels LF, Richard S, Kahnoski RJ, Anema J, Tuveson DA, Perez-Mancera PA, Mustonen V, Fischer A, Adams DJ, Rust A, Chan-on W, Subimerb C, Dykema K, Furge K, Campbell PJ, Teh BT, Stratton MR, Futreal PA (2011) Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* 469(7331):539–542. doi:10.1038/nature09639 [PubMed: 21248752]
75. Makinen N, Aavikko M, Heikkinen T, Taipale M, Taipale J, Koivisto-Korander R, Butzow R, Vahteristo P (2016) Exome Sequencing of Uterine Leiomyosarcomas Identifies Frequent Mutations in TP53, ATRX, and MED12. *PLoS Genet* 12(2):e1005850. doi:10.1371/journal.pgen.1005850 [PubMed: 26891131]
76. Ng SB, Bigam AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 42(9):790–793. doi:10.1038/ng.646 [PubMed: 20711175]
77. Parsons DW, Li M, Zhang X, Jones S, Leary RJ, Lin JC, Boca SM, Carter H, Samayoa J, Bettegowda C, Gallia GL, Jallo GI, Binder ZA, Nikolsky Y, Hartigan J, Smith DR, Gerhard DS, Fuets DW, VandenBerg S, Berger MS, Marie SK, Shinjo SM, Clara C, Phillips PC, Minturn JE, Biegel JA, Judkins AR, Resnick AC, Storm PB, Curran T, He Y, Rasheed BA, Friedman HS, Keir ST, McLendon R, Northcott PA, Taylor MD, Burger PC, Riggins GJ, Karchin R, Parmigiani G, Bigner DD, Yan H, Papadopoulos N, Vogelstein B, Kinzler KW, Velculescu VE (2011) The genetic landscape of the childhood cancer medulloblastoma. *Science* 331(6016):435–439. doi: 10.1126/science.1198056 [PubMed: 21163964]
78. Balakrishnan A, Bleeker FE, Lamba S, Rodolfo M, Daniotti M, Scarpa A, van Tilborg AA, Leenstra S, Zanon C, Bardelli A (2007) Novel somatic and germline mutations in cancer candidate genes in glioblastoma, melanoma, and pancreatic carcinoma. *Cancer Res* 67(8):3545–3550. doi: 10.1158/0008-5472.CAN-07-0065 [PubMed: 17440062]
79. Li S, Yen L, Pastor WA, Johnston JB, Du J, Shew CJ, Liu W, Ho J, Stender B, Clark AT, Burlingame AL, Daxinger L, Patel DJ, Jacobsen SE (2016) Mouse MORC3 is a GHKL ATPase that localizes to H3K4me3 marked chromatin. *Proc Natl Acad Sci USA* 113(35):E5108–E5116. doi:10.1073/pnas.1609709113 [PubMed: 27528681]
80. Desmedt C, Fumagalli D, Pietri E, Zoppoli G, Brown D, Nik-Zaina S, Gundem G, Rothe F, Majaj S, Garuti A, Carminati E, Loi S, Van Brussel T, Boeckx B, Maetens M, Mudie L, Vincent D, Kheddoudi N, Serra L, Massa I, Ballestrero A, Amadori D, Salgado R, de Wind A, Lambrechts D, Piccart M, Larsimont D, Campbell PJ, Sotiriou C (2015) Uncovering the genomic heterogeneity of multifocal breast cancer. *J Pathol* 236(4):457–466. doi: 10.1002/path.4540 [PubMed: 25850943]
81. Wagner BL, Norris JD, Knotts TA, Weigel NL, McDonnell DP (1998) The nuclear corepressors NCoR and SMRT are key regulators of both ligand- and 8-bromo-cyclic AMP-dependent transcriptional activity of the human progesterone receptor. *Mol Cell Biol* 18(3):1369–1378 [PubMed: 9488452]
82. Karhemo PR, Rivinoja A, Lundin J, Hyvonen M, Chernenko A, Lammi J, Sihto H, Lundin M, Heikkilä P, Joensuu H, Bono P, Laakkonen P (2011) An extensive tumor array analysis supports tumor suppressive role for nucleophosmin in breast cancer. *Am J Pathol* 179(2):1004–1014. doi: 10.1016/j.ajpath.2011.04.009 [PubMed: 21689627]

83. Daxinger L, Harten SK, Oey H, Epp T, Isbel L, Huang E, Whitelaw N, Apedaile A, Sorolla A, Yong J, Bharti V, Sutton J, Ashe A, Pang Z, Wallace N, Gerhardt DJ, Blewitt ME, Jeddeloh JA, Whitelaw E (2013) An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. *Genome Biol* 14(9):R96. doi:10.1186/gb-2013-14-9-r96 [PubMed: 24025402]
84. Wu Q, Madany P, Akech J, Dobson JR, Douthwright S, Browne G, Colby JL, Winter GE, Bradner JE, Pratap J, Sluder G, Bhargava R, Chiosea SI, van Wijnen AJ, Stein JL, Stein GS, Lian JB, Nickerson JA, Imbalzano AN (2015) The SWI/SNF ATPases Are Required for Triple Negative Breast Cancer Cell Proliferation. *J Cell Physiol* 230(11):2683–2694. doi:10.1002/jcp.24991 [PubMed: 25808524]
85. Rodriguez-Nieto S, Canada A, Pros E, Pinto AI, Torres-Lanzas J, Lopez-Rios F, Sanchez-Verde L, Pisano DG, Sanchez-Cespedes M (2011) Massive parallel DNA pyrosequencing analysis of the tumor suppressor BRG1/SMARCA4 in lung primary tumors. *Hum Mutat* 32(2):E1999–E2017. doi:10.1002/humu.21415 [PubMed: 21280140]
86. Bochar DA, Wang L, Beniya H, Kinev A, Xue Y, Lane WS, Wang W, Kashanchi F, Shiekhhattar R (2000) BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. *Cell* 102(2):257–265 [PubMed: 10943845]
87. Ring HZ, Vameghi-Meyers V, Wang W, Crabtree GR, Francke U (1998) Five SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin (SMARC) genes are dispersed in the human genome. *Genomics* 51(1):140–143. doi:10.1006/geno.1998.5343 [PubMed: 9693044]
88. Legare S, Cavallone L, Mamo A, Chabot C, Sirois I, Magliocco A, Klimowicz A, Tonin PN, Buchanan M, Keilty D, Hassan S, Laperriere D, Mader S, Aleynikova O, Basik M (2015) The estrogen receptor cofactor SPEN functions as a tumor suppressor and candidate biomarker of drug responsiveness in hormone-dependent breast cancers. *Cancer Res* 75(20):4351–4363. doi: 10.1158/0008-5472.CAN-14-3475 [PubMed: 26297734]
89. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324(5929):930–935. doi: 10.1126/science.1170116 [PubMed: 19372391]
90. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466(7310):1129–1133. doi:10.1038/nature09303 [PubMed: 20639862]

Table 1

Candidate genes screened in kConFab index cases

RefSeq_gene	Group	References
<i>ABCA13</i>	3	–
<i>AHSP</i>	3	–
<i>AKT1</i>	4	[30]
<i>AKT2</i>	4	[30]
<i>ANKRD17</i>	3	–
<i>ANKRD30A</i>	3	–
<i>APC</i>	4	[30]
<i>ARHGAP31</i>	3	–
<i>ARHGAP5</i>	5	[31]
<i>ARID1A</i>	1	[48]
<i>ARID1B</i>	1	[16, 49]
<i>ARID2</i>	1	[50]
<i>ASB4</i>	3	–
<i>ASH1L</i>	3	–
<i>ASIC3</i>	3	–
<i>ASXL1</i>	2	–
<i>ATR</i>	5	[31]
<i>ATRX</i>	1	[51, 5]
<i>AURKC</i>	1	[52, 53]
<i>BABAM1</i>	2	–
<i>BAP1</i>	1	[54]
<i>BAZ1B</i>	1	[18]
<i>BBS10</i>	3	–
<i>BCOR</i>	2	–
<i>BMPER</i>	3	–
<i>BRCC3</i>	1	[55]
<i>BRD7</i>	1	[56]
<i>C11orf30</i>	5	[31]
<i>C14orf21</i>	3	–
<i>C15orf42</i>	3	–
<i>C17orf50</i>	3	–
<i>C17orf78</i>	3	–
<i>CABIN1</i>	2	–
<i>CACNA1B</i>	3	–
<i>CACNA2D2</i>	3	–
<i>CADPS2</i>	3	–
<i>CALHM3</i>	3	–
<i>CAND2</i>	3	–
<i>CASC5</i>	2	–

RefSeq_gene	Group	References
<i>CASP8</i>	4	[30]
<i>CCND1</i>	4	[30]
<i>CD163L1</i>	3	–
<i>CDC73</i>	2	–
<i>CDHR2</i>	3	–
<i>CDKN1B</i>	4	[30]
<i>CELF2</i>	3	–
<i>CELSR1</i>	5	[31]
<i>CEP192</i>	5	[31]
<i>CHD1</i>	1	[57]
<i>CHD2</i>	1	[57]
<i>CHD3</i>	1	[57]
<i>CHD4</i>	1	[57]
<i>CHD5</i>	2	–
<i>CHD6</i>	2	–
<i>CHD7</i>	1	[57]
<i>CHD8</i>	1	[57]
<i>CHD9</i>	1	[57]
<i>CLIC6</i>	3	–
<i>CREBBP</i>	1	[58, 59]
<i>DAXX</i>	1	[51, 60]
<i>DAZAP2</i>	3	–
<i>DGKH</i>	3	–
<i>DLEC1</i>	3	–
<i>DLG5</i>	5	[31]
<i>DNAH5</i>	5	[31]
<i>DNMT1</i>	1	[61]
<i>DNMT3A</i>	1	[62, 63]
<i>DNMT3B</i>	1	[64]
<i>DPYD</i>	3	–
<i>DSCR6</i>	3	–
<i>EP300</i>	1	[65]
<i>EP400</i>	2	–
<i>EPX</i>	3	–
<i>ERBB2</i>	4	[30]
<i>EYA1</i>	2	–
<i>EZH2</i>	1	[66–68]
<i>FAM104B</i>	3	–
<i>FAM208A</i>	1	[69]
<i>FGFR1</i>	4	[30]
<i>FIGNL1</i>	3	–
<i>FLJ43860</i>	3	–

RefSeq_gene	Group	References
<i>FMN2</i>	5	[31]
<i>FOXO3</i>	1	[70]
<i>FREM1</i>	3	–
<i>FSHR</i>	2	–
<i>GATA2</i>	2	–
<i>GATA3</i>	4	[30]
<i>GGT1</i>	3	–
<i>GH1</i>	5	[31]
<i>GPR112</i>	5	[31]
<i>GPR125</i>	3	–
<i>H3F3A</i>	1	[60]
<i>HDAC1</i>	1	[18]
<i>HDAC4</i>	2	–
<i>HDAC9</i>	2	–
<i>HECW1</i>	5	[31]
<i>HMGA1</i>	1	[71]
<i>HMGA2</i>	1	[72]
<i>HNF1A</i>	2	–
<i>HRNR</i>	5	[31]
<i>HSPB1</i>	3	–
<i>HUWE1</i>	2	–
<i>IDH1</i>	1	[73]
<i>IDH2</i>	1	[73]
<i>IER2</i>	3	–
<i>IGSF3</i>	3	–
<i>IKZF1</i>	2	–
<i>IL17B</i>	3	–
<i>ITPR2</i>	3	–
<i>JAK2</i>	2	–
<i>JMJD1C</i>	2	–
<i>JMJD1C</i>	2	–
<i>KCNK15</i>	3	–
<i>KDM2A</i>	2	–
<i>KDM3A</i>	2	–
<i>KDM3B</i>	2	–
<i>KDM5A</i>	2	–
<i>KDM5B</i>	2	–
<i>KDM5C</i>	1	[74]
<i>KDM6A</i>	1	[74]
<i>KIAA0556</i>	5	[31]
<i>KIAA2018</i>	5	[31]
<i>KRAS</i>	4	[30]

RefSeq_gene	Group	References
<i>LAMA2</i>	5	[31]
<i>LILRA1</i>	3	–
<i>LIX1</i>	3	–
<i>LPHN2</i>	5	[31]
<i>LRP2</i>	5	[31]
<i>LYRM5</i>	3	–
<i>MAML1</i>	3	–
<i>MAP2K4</i>	4	[30]
<i>MAP3K1</i>	4	[30]
<i>MAP3K13</i>	4	[30]
<i>MDM2</i>	4	[30]
<i>MDN1</i>	5	[31]
<i>MED12</i>	1	[75]
<i>MEN1</i>	1	[51]
<i>MESP1</i>	3	–
<i>MLL</i>	2	–
<i>MLL2</i>	1	[76, 77]
<i>MLL3</i>	1	[78, 77]
<i>MLL4</i>	2	–
<i>MLL5</i>	2	–
<i>MORC3</i>	1	[79]
<i>MRPL37</i>	3	–
<i>MTOR</i>	1	[51]
<i>MYB</i>	1	[80]
<i>MYC</i>	4	[30]
<i>NCOR1</i>	1	[81]
<i>NEB</i>	3	–
<i>NELL1</i>	5	[31]
<i>NET1</i>	3	–
<i>NF1</i>	4	[30]
<i>NIPBL</i>	2	–
<i>NLRP2</i>	3	–
<i>NOS1</i>	2	–
<i>NOTCH1</i>	3	–
<i>NPM1</i>	1	[82]
<i>NSD1</i>	2	–
<i>NUMBL</i>	3	–
<i>NUP210</i>	3	–
<i>OPRM1</i>	3	–
<i>PAM</i>	3	–
<i>PAX5</i>	2	–
<i>PBRM1</i>	1	[74]

RefSeq_gene	Group	References
<i>PCDHB11</i>	3	–
<i>PDE4DIP</i>	5	[31]
<i>PDGFRL</i>	3	–
<i>PGLYRP4</i>	3	–
<i>PIK3C2G</i>	3	–
<i>PIK3CA</i>	5	[31]
<i>PKD1L1</i>	5	[31]
<i>PLEKHB1</i>	3	–
<i>PLIN4</i>	3	–
<i>POC5</i>	3	–
<i>PPARGC1A</i>	2	–
<i>PRDM6</i>	3	–
<i>PRKAA2</i>	2	–
<i>PRKCA</i>	2	–
<i>PRKCB</i>	2	–
<i>PRKCQ</i>	5	[31]
<i>PUM2</i>	3	–
<i>RALGAPA1</i>	5	[31]
<i>RANBP17</i>	3	–
<i>RASA3</i>	3	–
<i>RB1</i>	1	[71, 80]
<i>RBL1</i>	2	–
<i>RBM44</i>	3	–
<i>REST</i>	2	–
<i>RIF1</i>	1	[83]
<i>RLF</i>	1	[83]
<i>RNF180</i>	3	–
<i>RP1L1</i>	3	–
<i>RPTN</i>	3	–
<i>SALL1</i>	2	–
<i>SAMD14</i>	3	–
<i>SATL1</i>	5	[31]
<i>SCAMP2</i>	3	–
<i>SETD1A</i>	2	–
<i>SETD2</i>	1	[74]
<i>SETX</i>	3	–
<i>SLC22A5</i>	3	–
<i>SLC25A2</i>	3	–
<i>SLC37A4</i>	3	–
<i>SMAD4</i>	4	[30]
<i>SMARCA1</i>	1	[71]
<i>SMARCA2</i>	1	[84]

RefSeq_gene	Group	References
<i>SMARCA4</i>	1	[85]
<i>SMARCA5</i>	1	[61]
<i>SMARCB1</i>	1	[7]
<i>SMARCC2</i>	1	[86]
<i>SMARCD1</i>	1	[87]
<i>SMCHD1</i>	1	[17]
<i>SMG5</i>	5	[31]
<i>SPEN</i>	1	[88]
<i>SPHKAP</i>	5	[31]
<i>SRA1</i>	3	–
<i>SRCAP</i>	2	–
<i>SSH2</i>	3	–
<i>ST8SIA5</i>	3	–
<i>SUPT3H</i>	3	–
<i>SYNE1</i>	5	[31]
<i>SYNE2</i>	5	[31]
<i>TAF1</i>	2	–
<i>TAF1L</i>	2	–
<i>TBX3</i>	4	[30]
<i>TCTN1</i>	3	–
<i>TDO2</i>	5	[31]
<i>TEK</i>	3	–
<i>TET1</i>	1	[89]
<i>TET2</i>	1	[90]
<i>TMEM106C</i>	3	–
<i>TMEM131</i>	3	–
<i>TMEM144</i>	3	–
<i>TMEM37</i>	3	–
<i>TPR</i>	3	–
<i>TPTE2</i>	3	–
<i>TRIM28</i>	1	[19]
<i>TRRAP</i>	2	–
<i>TSHZ3</i>	5	[31]
<i>TXNDC2</i>	3	–
<i>TYRO3</i>	3	–
<i>UBAP2L</i>	5	[31]
<i>UBR5</i>	5	[31]
<i>UGT1A8</i>	5	[31]
<i>USH2A</i>	5	[31]
<i>UTP14C</i>	3	–
<i>UTS2R</i>	3	–
<i>WDR36</i>	3	–

RefSeq_gene	Group	References
<i>XRN1</i>	5	[31]
<i>ZBTB34</i>	3	–
<i>ZNF142</i>	5	[31]
<i>ZNF217</i>	3	–
<i>ZNF227</i>	3	–
<i>ZNF256</i>	3	–
<i>ZNF418</i>	3	–
<i>ZNF438</i>	3	–
<i>ZNF514</i>	3	–
<i>ZNF518B</i>	3	–
<i>ZNF558</i>	3	–
<i>ZNF570</i>	3	–
<i>ZNF703</i>	4	[30]
<i>ZNF764</i>	3	–
<i>ZNF850</i>	3	–

Table 2

Results of segregation analysis in kConFab families

Gene	Group	# Families	Log likelihood
<i>ABCA13</i>	3	15	0.85 (0.32, 1.95)
<i>ANKRD30A</i>	3	1	0.01 ^a (0.01 ^a , 2.17)
<i>APC</i>	4	8	2.96 (0.86, 26.0 ^a)
<i>ARID1A</i>	1	7	0.58 (0.05, 1.83)
<i>ASH1L</i>	3	5	1.24 (0.43, 2.86)
<i>ATR</i>	5	4	1.00 (0.17, 5.65)
<i>BAZ1B</i>	1	3	>26.0
<i>C15orf42</i>	3	5	0.74 (0.06, 2.55)
<i>CASC5</i>	2	2	1.49 (0.10, 8.82)
<i>CASP8</i>	4	1	>26.0
<i>CDHR2</i>	3	1	<0.01
<i>CELSR1</i>	5	5	0.50 (0.04, 1.73)
<i>CEP192</i>	5	9	1.22 (0.35, 3.39)
<i>CHD5</i>	2	3	0.96 (0.16, 3.28)
<i>CHD7</i>	1	5	1.17 (0.42, 2.87)
<i>CHD8</i>	1	7	2.40 (1.0, 7.32)
<i>CHD9</i>	1	9	0.62 (0.22, 1.34)
<i>DGKH</i>	3	4	1.75 (0.80, 4.57)
<i>DLEC1</i>	3	5	0.01 ^a (0.01 ^a , 3.86)
<i>DNAH5</i>	5	16	0.86 (0.36, 1.79)
<i>DNMT3A</i>	1	1	>26.0
<i>EP400</i>	2	5	1.65 (0.66, 4.21)
<i>HECW1</i>	5	3	1.41 (0.42, 4.13)
<i>HUWE1</i>	2	9	2.9 (0.8, 15.0)
<i>IL17B</i>	3	2	0.85 (0.07, 2.52)
<i>ITPR2</i>	3	9	1.59 (0.58, 4.61)
<i>KDM5B</i>	2	6	7.50 (0.24, 26.0 ^a)
<i>LAMA2</i>	5	19	1.32 (0.75, 2.29)
<i>LPHN2</i>	5	5	1.60 (0.76, 3.33)
<i>LRP2</i>	5	7	0.73 (0.17, 2.06)
<i>MAP3K1</i>	4	6	1.16 (0.09, 16.17)
<i>MDN1</i>	5	11	0.01 ^a (0.01 ^a , 0.03)
<i>NCOR1</i>	1	5	1.59 (0.06, 26.0 ^a)
<i>NEB</i>	3	25	0.87 (0.40, 1.65)
<i>NELL1</i>	5	1	1.0 (0.04, 22.26)
<i>NET1</i>	3	3	1.32 (0.45, 3.30)
<i>NIPBL</i>	2	5	1.22 (0.36, 2.86)

Gene	Group	# Families	Log likelihood
<i>NOS1</i>	2	6	0.97 (0.14, 4.39)
<i>NOTCH1</i>	3	7	1.82 (0.67, 5.22)
<i>NUP210</i>	3	6	0.58 (0.04, 2.14)
<i>PAX5</i>	2	3	1.52 (0.68, 3.21)
<i>PDE4DIP</i>	5	1	>26.0
<i>PIK3C2G</i>	3	1	2.78 (0.64, 25.47)
<i>PKD1L1</i>	5	3	0.94 (0.10, 2.61)
<i>PLIN4</i>	3	1	0.78 (0.05, 7.54)
<i>PUM2</i>	3	3	1.77 (0.38, 8.95)
<i>RALGAPA1</i>	5	5	1.22 (0.39, 3.29)
<i>RBM44</i>	3	3	0.30 (0.05, 1.02)
<i>RIF1</i>	1	6	0.39 (0.03, 1.52)
<i>RNF180</i>	3	2	0.01 ^a (0.01 ^a , 1.85)
<i>SETX</i>	3	7	1.37 (0.48, 3.80)
<i>SLC22A5</i>	3	3	6.66 (0.03, 26.0 ^a)
<i>SMG5</i>	5	2	0.01 ^a (0.01 ^a , 6.44)
<i>SPHKAP</i>	5	2	0.67 (0.09, 4.44)
<i>SUPT3H</i>	3	1	0.01 ^a (0.01 ^a –2.3)
<i>SYNE1</i>	5	22	0.88 (0.45, 1.50)
<i>SYNE2</i>	5	9	2.14 (0.94, 5.93)
<i>TDO2</i>	5	1	0.01 ^a (0.01 ^a , 1.73)
<i>TET2</i>	1	3	2.21 (0.07, 26.0 ^a)
<i>TMEM131</i>	3	9	1.61 (0.77, 3.78)
<i>TPR</i>	3	4	1.06 (0.31, 3.05)
<i>UBAP2L</i>	5	4	0.70 (0.15, 2.73)
<i>UBR5</i>	5	1	0.01 ^a (0.01 ^a –6.6)
<i>USH2A</i>	5	10	2.48 (1.11, 6.67)
<i>XRN1</i>	5	2	<0.01
<i>ZNF558</i>	3	2	1.28 (0.33, 3.22)
TOTAL		366	

Group 1: genes involved in epigenetic regulation selected from the literature

Group 2: using the GO terms 'chromatin organization' or 'chromatin reorganization'

Group 3: candidate genes from whole-exome sequencing of five non-*BRCA1/2* families

Groups 4 and 5: breast cancer driver genes [30, 31]

^aEstimate at boundary

Table 3

Segregation analysis: candidate genes stratified by group

Groups	Description	Included genes	# Families	Log likelihood (95% CI)
1 + 2	Genes involved in epigenetic regulation selected from the literature and using GO terms	<i>ARID1A, BAZ1B, CHD7, CHD8, CHD9, DNMT3A, NCOR1, RIF1, TET2, CASC5, CHD5, EP400, HUWE1, KDM5B, NIPBL, NOS1, PAX5</i>	85	1.17 (0.86, 1.56)
3	Candidate genes from whole-exome sequencing	<i>ABCA13, ANK4D30A, C15orf42, CDHR2, DGKH, DLEC1, IL17B, ITPR2, NEB, NET1, NOTCH1, PIK3C2G, PLIN4, PUM2, RBM44, RNF180, SLC22A5, SUPT3H</i>	124	1.17 (0.94, 1.46)
4 + 5	Driver genes [30, 31]	<i>APC, CASP8, MAP3K1, ATR, CELSR1, CEP192, DNAH5, HECW1, LAMA2, LPHN2, LRP2, MDN1, NELL1, PDE4DIP, PKDILL1, RALGAPA1, SMG5, SPHKAP, SYNE1, SYNE2, TDO2, UBAP2L, UBR5, USH2A, XRN1</i>	157	1.10 (0.89, 1.35)
1 + 2 + 4 + 5			242	1.11 (0.94, 1.32)