Supplementary Material

Penetrance of Breast Cancer Susceptibility Genes from the eMERGE III Network

Xiao Fan, PhD[1, 2], Julia Wynn, MS[1], Ning Shang, PhD[3], Cong Liu, PhD[3], Alexander Fedotov, PhD[4], Miranda L.G. Hallquist, MS[14], Adam H. Buchanan, MS[14], Marc S. Williams, MD[14], Maureen E. Smith, MS[5], Christin Hoell, MS[6], Laura J. Rasmussen-Torvik, PhD[7], Josh F. Peterson, MD[8], Georgia L. Wiesner, MD[9], Andrea M. Murad, MS[10], Gail P. Jarvik, MD, PhD[11], Adam S. Gordon, PhD[6], Elisabeth A. Rosenthal, PhD[11], Ian B. Stanaway, PhD[11], David R. Crosslin, PhD[12], Eric B. Larson, MD[13], Kathleen A. Leppig, MD[13], Nora B. Henrikson, PhD[13], Janet L. Williams, MS[14], Rongling Li, PhD[15], Scott Hebbring, PhD[16], Chunhua Weng, PhD[3], Yufeng Shen, PhD[2, 3], Katherine D. Crew, MD[17,18], Wendy K. Chung, MD, PhD[1, 17,18 *]

[1] Department of Pediatrics, Columbia University Irving Medical Center, New York, NY 10032
[2] Department of Systems Biology, Columbia University Irving Medical Center, New York, NY 10032
[3] Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY 10032
[4] Irving Institute for Clinical and Translational Research, Columbia University Irving Medical Center, New York, NY 10032
[5] Department of Medicine, Northwestern University, Chicago Feinberg School of Medicine, IL 60657
[6] Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL 60611
[7] Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL 60611
[8] Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville TN 37203
[9] Department of Medicine, Vanderbilt University Medical Center, Nashville, TN 37203
[10] Division of Genetic Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109
[11] Department of Medicine (Medical Genetics), University of Washington Medical Center, Seattle, WA, 98105
[12] Department of Biomedical Informatics and Medical Education, University of Washington Medical Center, Seattle, WA, 98105
[13] Kaiser Permanente of Washington, Seattle, WA 98112
[14] Genomic Medicine Institute, Geisinger, Danville, PA 17822
[15] Division of Genomic Medicine, National Human Genome Research Institute, National Institutes of Health, Baltimore, MD 21224
[16] Center for Precision Medicine Research, Marshfield Clinic, Marshfield, WI 54449
[17] Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY 10032
[18] Department of Medicine, Columbia University Irving Medical Center, New York, NY 10032
* Corresponding author, email: wkc15@cumc.columbia.edu

Supplementary Table 1. Clinical characteristics of 24,947 eMERGE III participants.

Sex[a]	Ancestry[b]	Count	Current age (SD), y
Female (n=13,458)	African American	1758	38.6 (23.5)
	East Asian	952	52.3 (14.2)
	European American	9122	54.1 (21.3)
	Hispanic	960	49.5 (18)
	South Asian	65	63.1 (13.9)
	unknown	601	48.8 (23.1)
Male (n=11,498)	African American	1994	31.5 (21.8)
	East Asian	588	53.9 (16)
	European American	7892	54.1 (23.8)
	Hispanic	546	46 (20.6)
	South Asian	34	51.3 (20.1)
	unknown	444	41.3 (24.7)

[a]Sex was genetically inferred.
[b]Ancestry was self-reported.

Supplementary Table 2. Recurrent P/LP variants in the breast cancer susceptibility genes, sorted by allele count in the eMERGE penetrance cohort.

cDNA	Gene	Function	ClinVar	Allele count
c.470T>C	*CHEK2*	missense	Conflicting interpretations of pathogenicity	26
c.1100delC	*CHEK2*	frameshift	Conflicting interpretations of pathogenicity	7
c.2638+2T>C	*ATM*	splice-donor	Pathogenic/Likely pathogenic by multiple submitters	4
c.5946del	*BRCA2*	frameshift	Pathogenic reviewed by expert panel	3
c.1427C>T	*CHEK2*	missense	Conflicting interpretations of pathogenicity	3
c.1283C>T	*CHEK2*	missense	Conflicting interpretations of pathogenicity	3
c.7271T>G	*ATM*	missense	Pathogenic/Likely pathogenic by multiple submitters	2
c.3049C>T	*ATM*	stop-gain	Pathogenic by multiple submitters	2
c.4876_4877delAA	*BRCA2*	frameshift	Pathogenic reviewed by expert panel	2
c.5217_5223delTTTAAGT	*BRCA2*	frameshift	Pathogenic reviewed by expert panel	2
c.9253dupA	*BRCA2*	frameshift	Pathogenic reviewed by expert panel	2
c.2808_2811delACAA	*BRCA2*	frameshift	Pathogenic reviewed by expert panel	2
c.7558C>T	*BRCA2*	stop-gain	Pathogenic reviewed by expert panel	2
c.2257C>T	*PALB2*	stop-gain	Pathogenic by multiple submitters	2

Supplementary Table 3. Number of carriers and unique P/LP variants for each gene and variant type.

Gene	missense	splice	stop-gain	frameshift	CNV	ClinVar[a]
No. of carriers						
ATM	3	6	7	5	0	18 (86%)
BRCA1	3	2	4	8	0	15 (88%)
BRCA2	0	5	8	26	0	37 (95%)
CHEK2	33	2	1	10	2	7 (15%)
PALB2	0	4	6	5	0	14 (93%)
PTEN	1	0	1	1	0	3 (100%)
TP53	3	0	2	0	0	4 (80%)
No. of unique variants						
ATM	2	3	6	5	0	13 (81%)
BRCA1	3	2	4	8	0	15 (88%)
BRCA2	0	5	7	20	0	30 (94%)
CHEK2	4	2	1	4	2	7 (54%)
PALB2	0	4	5	5	0	13 (93%)
PTEN	1	0	1	1	0	3 (100%)
TP53	3	0	2	0	0	4 (80%)

[a] Number of variants with at least 2-star review status in ClinVar is given with percentage in the parentheses. CNV = copy number variation.

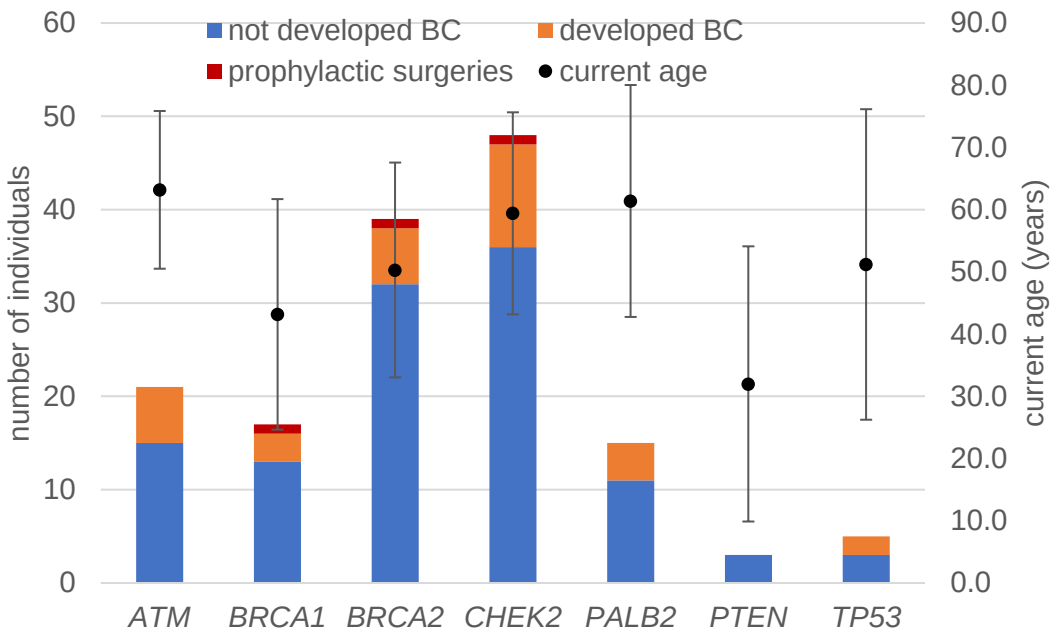
Supplementary Table 4. Prevalence of individuals with two *CHEK2* variants estimated in various European populations[a]

Ancestry	Prevalence	No. of individuals	Reference
CHEK2:c.470T>C			
Byelorussian	1.3%	307	[1]
German	0.6%	486	[1]
German	1.6%	500	[2]
Polish	4.8%	4000	[3]
Finnish	4.9%	12,551	gnomAD (non-cancer)
non-Finish European	1.0%	59,069	gnomAD (non-cancer)
European American	0.2%	17,014	eMERGE III
CHEK2: c.1100delC			
Byelorussian	0.0%	307	[1]
German	0.2%	486	[1]
Czech	0.3%	730	[4]
German	0.5%	1315	[2]
European	0.6%	40063	[5]
non-Finish European	1.7%	12,554	gnomAD (non-cancer)
European American	0.5%	58,451	gnomAD (non-cancer)
European American	0.1%	17,014	eMERGE III

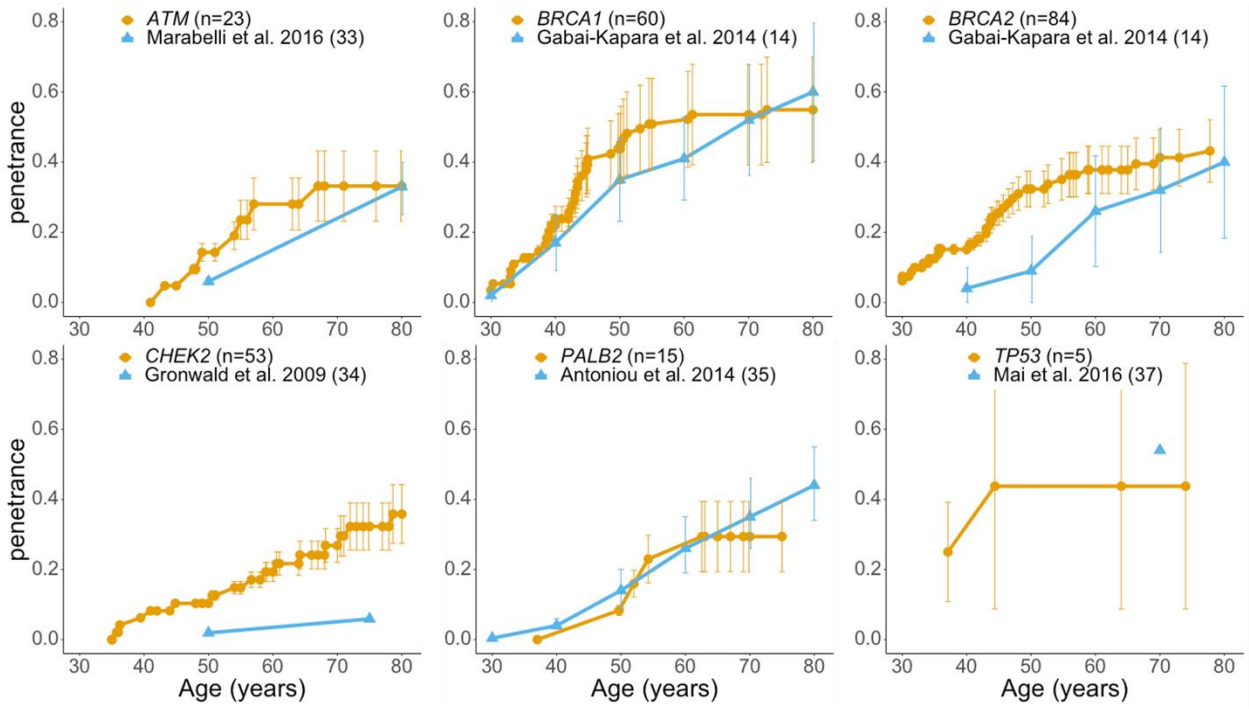
[a]The prevalence in gnomAD was calculated based on the allele frequency.

Supplementary Table 5. Summary of recent penetrance studies that were used to compare with these population-based penetrance estimates for the six breast cancer genes.

Genes	No. of women with P/LP variants	Source of the study populations	Ancestry	Reference
ATM	NA	Meta case-control studies	multiple	Marabelli et al. 2016 [6]
BRCA1/2	211	General population	Ashkenazi Jewish	Gabai-Kapara et al. 2014 [7]
CHEK2	533	Cancer clinic	European	Gronwald et al. 2009 [8]
PALB2	311	Cancer clinic	European	Antoniou et al. 2014 [9]
PTEN	368	Breast Cancer Association Consortium	Asian	Han et al. 2017 [10]
TP53	189	Cancer history	multiple	Mai et al. 2016 [11]



Supplementary Figure 1. Breast cancer (BC) status by the last review including not developed breast cancer (blue), developed breast cancer (orange) and prophylactic mastectomy (red). The average age and standard deviation were shown in the secondary y axis.



Supplementary Figure 2. Penetrance (orange) of six breast cancer genes compared with penetrance (blue) in the literature from Table 2. The sample size of the entire eMERGE III including those who were aware of their breast cancer genetic results prior to enrollment in eMERGE. Error bars indicate 95% confidence intervals (CI). CI of penetrance for *CHEK2* and *TP53* are not available from the literatures. Genes are sorted alphabetically.

References:

1. Bogdanova, N., et al., *Association of two mutations in the CHEK2 gene with breast cancer.* Int J Cancer, 2005. **116**(2): p. 263-6.
2. Dufault, M.R., et al., *Limited relevance of the CHEK2 gene in hereditary breast cancer.* Int J Cancer, 2004. **110**(3): p. 320-5.
3. Szymanska-Pasternak, J., et al., *CHEK2 variants predispose to benign, borderline and low-grade invasive ovarian tumors.* Gynecol Oncol, 2006. **102**(3): p. 429-31.
4. Kleibl, Z., et al., *The CHEK2 c.1100delC germline mutation rarely contributes to breast cancer development in the Czech Republic.* Breast Cancer Res Treat, 2005. **90**(2): p. 165-7.
5. Muranen, T.A., et al., *Genetic modifiers of CHEK2*1100delC-associated breast cancer risk.* Genet Med, 2017. **19**(5): p. 599-603.
6. Marabelli, M., S.C. Cheng, and G. Parmigiani, *Penetrance of ATM Gene Mutations in Breast Cancer: A Meta-Analysis of Different Measures of Risk.* Genet Epidemiol, 2016. **40**(5): p. 425-31.
7. Gabai-Kapara, E., et al., *Population-based screening for breast and ovarian cancer risk due to BRCA1 and BRCA2.* Proc Natl Acad Sci U S A, 2014. **111**(39): p. 14205-10.
8. Gronwald, J., et al., *Cancer risks in first-degree relatives of CHEK2 mutation carriers: effects of mutation type and cancer site in proband.* Br J Cancer, 2009. **100**(9): p. 1508-12.
9. Antoniou, A.C., et al., *Breast-cancer risk in families with mutations in PALB2.* N Engl J Med, 2014. **371**(6): p. 497-506.
10. Han, M.R., et al., *Evaluating genetic variants associated with breast cancer risk in high and moderate-penetrance genes in Asians.* Carcinogenesis, 2017. **38**(5): p. 511-518.
11. Mai, P.L., et al., *Risks of first and subsequent cancers among TP53 mutation carriers in the National Cancer Institute Li-Fraumeni syndrome cohort.* Cancer, 2016. **122**(23): p. 3673-3681.