Supplemental Methods

Ethics Statement

The mutation screening and analyses described here were approved by the Ethics committee of the International Agency for Research on Cancer (IARC), the University of Utah Institutional Review Board (IRB), and the local IRBs of the Breast Cancer Family Registry (BCFR) centers from which we received samples. These local IRBs were the Health Sciences Human Ethics Subcommittee of the University of Melbourne, Australia; the Institutional Review Board of the Northern California Cancer Center (now the Cancer Prevention Institute of California); and the Research Ethics Board of Mount Sinai Hospital, Ontario, Canada. All participants gave written informed consent.

Participants

The design of this study has been described in detail previously[1–6]. Briefly, eligible participants included women ascertained by population-based sampling by the Australian, Northern California, and Ontario sites of the Breast Cancer Family Registry (BCFR)[7]. Participants were recruited between 1995 and 2005. Selection criteria for cases (N = 1,297) were diagnosis of breast cancer at or before age 45 years and self-reported race/ethnicity plus grandparents' country of origin information consistent with Caucasian, East Asian, Hispanic/Latino, or African American racial/ethnic heritage. The controls (N = 1,121) were frequency matched to the cases within each center on racial/ethnic group, with age at selection not more than ± 10 years from the age range at diagnosis of the cases gathered from the same center (Table 1). Because of the shortage

of available controls in some racial/ethnic and age groups, the frequency matching was not one-to-one in all subgroups. Known carriers of pathogenic variants in *BRCA1* and *BRCA2* were excluded. Known *PALB2* carriers were also excluded due to recent results upgrading that gene to high-risk status[8]. *PTEN*, *TP53*, and *CDH1* carrier status was unknown, but would only affect a small number of individuals in the study. In the analysis of the sequence variants, we excluded participants with a PCR failure rate higher than 20% of the coding sequence from the rare variants analyses, and a SNP genotyping failure rate of more than 4 of 18 SNPs from the SNP polygene analyses.

Mutation Screening for Variants in the 9 Moderate-Risk Genes

For mutation screening of the coding exons and proximal splice junction regions of *ATM* (NM_000051.3), *BARD1* (NM_000465.3), *CHEK2* (NM_007194.3), *MRE11A* (NM_005591.3), *NBN* (NM_002485.4), *RAD50* (NM_005732.3), *RAD51* (NM_002875.4), *RINT1* (NM_021930.4), and *XRCC2* (NM_005431.1), we used 30 ng of whole-genome amplified (WGA) DNA obtained by mixing 15 ng of amplified DNA from each of two independent WGA reactions. The laboratory process was as described in detail in the prior studies of the 9 moderate-risk genes used in this study[1–6]. Our semi-automated approach, handled by a Laboratory Information Management System (LIMS)[9], relies on mutation scanning by high-resolution melt curve (HRM) analysis followed by direct Sanger sequencing of the individual samples for which an aberrant melting curve profile is indicative of the presence of a sequence variant.

All exonic sequence variants, plus intronic sequence variants that fell within 20 bp of a splice acceptor or eight bp of a splice donor, and were either unreported or had an

allele frequency of <1% in the large scale reference groups “Caucasian Americans”, “African Americans” and “East Asians” based on exome variant server (EVS) and 1,000 genomes project (1000G) data (http://evs.gs.washington.edu/EVS; http://browser.1000genomes.org/index.html), were confirmed either by independent re-amplification and sequencing from each of the two independent WGA reaction products and concordant variant calls, or, for five variants, by re-amplification and sequencing from genomic DNA.

All samples that failed either at the primary PCR, secondary PCR, or sequencing reaction stage were re-amplified from WGA DNAs or genomic DNAs. Samples that still did not provide satisfactory mutation screening results for at least 80% of the concatenated MRN coding sequence were excluded from further analysis. Primer and probe sequences are available from the authors upon request.

Genotyping for 18 SNPs

Genotyping for 18 of the first breast cancer modest-risk SNPs confirmed by the Breast Cancer Association Consortium(BCAC).[10, 11] Genotyping for rs13387042, rs13281615, rs2981578, rs4973768, rs11249433, rs2046210, rs704010, rs10995190, rs10941679, rs2380205, rs6504950, rs614367, rs1011970, and rs999737) involved a nested and multiplexed polymerase chain reaction (PCR) of whole-genome amplified (WGA) DNA followed by high-resolution melting (HRM) curve analysis to identify major and minor alleles. Our HRM analysis consisted of two assays, either unlabeled probe or small amplicon-based genotyping. Genotyping for rs1045485, rs3803662, rs889312, rs3817198 were genotyped by Taqman. Genotyping began at the World Health

Organization's (WHO) International Agency for Research on Cancer (IARC) before moving to the University of Utah's Huntsman Cancer Institute (HCI). Primer and probe sequences are available from the authors upon request.

Protein Multiple Sequence Alignment Organisms

Orthologs from Human (*Homo sapiens*), either mouse (*Mus musculus*) or rat (*Rattus norvegicus*) from clade Murinae, either pig (*Sus scrofa*), cow (*Bos taurus*), dog (*Canis lupus*) or panda (*Ailuropoda melanoleuca*) from clade Laurasiatheria, elephant (*Loxodonta africana*), armadillo (*Dasypus novemcinctus*), either opossum (*Monodelphis domestica*) or tasmanian devil (*Sarcophilus harrisii*) from clade Metatheria, platypus (*Ornithorhynchus anatinus*), chicken (*Gallus gallus*) and lizard (*Anolis carolinensis*) or painted turtle(*Chrysemys picta bellii*) from clade Sauria, clawed frog (*Xenopus laevis* or *Xenopus tropicalis*), coelacanth (*Latimeria chalumnae*), either zebrafish (*Danio rerio*), Tetraodon (*Tetraodon nigroviridis*), or Fugu (*Takifugu rubripes*) from clade Clupeocephala, lancelet (*Branchiostoma floridae*), and sea urchin (*Strongylocentrotus purpuratus*) were included in our initial alignments. RAD50 and RAD51 did not reach three substitutions per position for their individual pMSA using the preceding organisms, so orthologous protein sequences from the model organisms *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Arabidopsis thaliana* were added until three substitutions per position was achieved.

Ortholog sequences downloaded from Genbank were aligned using the expresso extension of T-Coffee to create the initial pMSA[12, 13]. These initial alignments were checked by hand in Geneious v7.1.4 (http://www.geneious.com) for anomalies that might

be attributed to gene model errors rather than actual sequence divergence. Potential anomalies were corrected by reference to, and gene re-prediction from, genomic DNA sequence available on the UCSC genome browser (http://genome.ucsc.edu).

Protein Multiple Sequence Alignment Depth Determination

We estimated substitutions in each alignment by using Protpars from PHYLIP v3.69[14] with a constrained phylogeny to make a maximum parsimony estimate of the number of substitutions that occurred for each organism in the underlying phylogeny. In order to create a distance table for MAPP, we used Protdist to establish a distance matrix for protein sequences using maximum likelihood estimates with a constrained phylogeny followed by Fitch to generate a phylogenetic tree under the additive tree model. The complete alignments, phylogenetic trees, and distances are available upon request.

Statistics

To assess evidence of risk from the case-control frequency distribution of protein-truncating variants (T), known or very likely spliceogenic splice-junction variants (SJV), and rare missense substitutions (rMS), we constructed a table with one entry per subject; the variants per subject; and annotations for whether the variant was in a key functional domain, its frequency, as well as study center, case-control status, race/ethnicity, and age for the subject. For the subjects who carried more than one rare variant of interest, only the most deleterious score was considered. Data from key functional domain rare silent substitutions, key functional domain rare missense substitutions (including in-frame deletions), and protein truncating variants were used in a ROC analysis. The silent

substitutions were scored using CADD and MAPP, and the missense substitutions were scored using Align-GVGD, CADD, MAPP, and PolyPhen-2[15–18]. ROC curves were drawn (Supplemental Figure 1) and AUC was determined (Supplemental Table 3) using the ROCR library (http://rocr.bioinf.mpi-sb.mpg.de). Correlations between the missense analysis programs were performed using the continuous variables of each missense analysis program for linear regressions via Stata version 12.1 software (StataCorp, College Station, TX, USA). We also created a synthetic variable "consensus" that counted the number of missense analysis programs that considered a variant "above threshold" and preformed similar linear regressions between the variable output of the missense analysis program with our synthetic consensus variable (Supplemental Table 4).

For MAPP, Align-GVGD, and Polyphen2, logistic regression trend tests were formatted such that participants who did not carry any rare variant were given a score of 0, carriers of rMS outside of a domain were given a 1, carriers of a "below threshold" rMS were given a 2, carriers of an "above threshold" rMS were given a 3, carriers of an in-frame deletion were given a 4 and T+SJVs were given a 5. CADD gives severity scores for more than rMS and was treated differently as follows: carriers of a non-rMS variant that would have been in the "above threshold" group were given a 2, carriers of a "below threshold" rMS were given a 3, carriers of an "above threshold" rMS were given a 4, carriers of an in-frame deletion were given a 5 and T+SJVs were given a 6.

These row labels were then used as a categorical variable in the logistic regressions. The reference non-carrier group (assigned logistic regression row label 0 for each classifier) comprised the participants who were not reported to carry an rMS, an in-frame deletion, or a T+SJV variant in a domain of one of the 9 genes. The same reference

group of non-carriers was used for MAPP, Align-GVGD and Polyphen2 analyses. The participants in CADD that carried a non-rMS variant that would have been in the "above threshold" group would have been classified as non-carriers in the alternative methods, but were not considered as part of the reference group for the analyses involving CADD. Differences in the case–control ratio between racial/ethnic groups and study center were accounted for by including categorical variables for each racial/ethnic group and each study center. Adjustment for racial/ethnic group should also capture confounding of genetic and social factors with interaction terms, allowing that this confounding effect may be different for the broadly labeled racial/ethnic groups in different centers.

To determine the correlation between NPS and the observed OR, we divided our subjects into deciles based on the percentiles observed in the controls. The middle quintile (40-60) was used as reference group and given a value of 0. The remaining deciles were given values 1-8 respectively, which were used as these categorical variables in logistic regressions. For the threshold analysis, we used the middle quintile as the reference group and gave this group a value of 0. The "below threshold, outside of reference" group was given a value of 1 and the "above threshold" group was given a value of 2. To create the histogram for the moderate-risk gene combined with the SNP polygene, the NPS and OR from the moderate-risk gene were multiplied together. The middle quintile was excluded for the linear regression equation, and the NPS was treated as the independent variable and the resultant OR of each group as the dependent variable. P-values were found by testing the regression coefficient equal to 0 or 1.

Supplementary Methods Sources

1 Tavtigian S V, Oefner PJ, Babikyan D, Hartmann A, Healey S, Le Calvez-Kelm F, Lesueur F, Byrnes GB, Chuang S-C, Forey N, Feuchtinger C, Gioia L, Hall J, Hashibe M, Herte B, McKay-Chopin S, Thomas A, Vallée MP, Voegele C, Webb PM, Whiteman DC, Sangrajrang S, Hopper JL, Southey MC, Andrulis IL, John EM, Chenevix-Trench G. Rare, evolutionarily unlikely missense substitutions in ATM confer increased risk of breast cancer. *Am J Hum Genet* 2009;85:427–46.

2 Damiola F, Pertesi M, Oliver J, Le Calvez-Kelm F, Voegele C, Young EL, Robinot N, Forey N, Durand G, Vallée MP, Tao K, Roane TC, Williams GJ, Hopper JL, Southey MC, Andrulis IL, John EM, Goldgar DE, Lesueur F, Tavtigian S V. Rare key functional domain missense substitutions in MRE11A, RAD50, and NBN contribute to breast cancer susceptibility: results from a Breast Cancer Family Registry case-control mutation-screening study. *Breast Cancer Res* 2014;16:R58.

3 Le Calvez-Kelm F, Oliver J, Damiola F, Forey N, Robinot N, Durand G, Voegele C, Vallée MP, Byrnes G, Registry BCF, Hopper JL, Southey MC, Andrulis IL, John EM, Tavtigian S V, Lesueur F. RAD51 and Breast Cancer Susceptibility: No Evidence for Rare Variant Association in the Breast Cancer Family Registry Study. *PLoS One* 2012;7:e52374.

4 Le Calvez-Kelm F, Lesueur F, Damiola F, Vallée M, Voegele C, Babikyan D, Durand G, Forey N, McKay-Chopin S, Robinot N, Nguyen-Dumont T, Thomas A, Byrnes GB, Hopper JL, Southey MC, Andrulis IL, John EM, Tavtigian S V. Rare, evolutionarily unlikely missense substitutions in CHEK2 contribute to breast cancer susceptibility: results from a breast cancer family registry case-control mutation-screening study. *Breast Cancer Res* 2011;13:R6.

5 Park DJ, Tao K, Le Calvez-Kelm F, Nguyen-Dumont T, Robinot N, Hammet F, Odefrey F, Tsimiklis H, Teo ZL, Thingholm LB, Young EL, Voegele C, Lonie A, Pope BJ, Roane TC, Bell R, Hu H, Shankaracharya, Huff CD, Ellis J, Li J, Makunin I V., John EM, Andrulis IL, Terry MB, Daly M, Buys SS, Snyder C, Lynch HT, Devilee P, Giles GG, Hopper JL, Feng B-J, Lesueur F, Tavtigian S V., Southey MC, Goldgar DE. Rare Mutations in RINT1 Predispose Carriers to Breast and Lynch Syndrome-Spectrum Cancers. *Cancer Discov* 2014;4:804–15.

6 Park DJ, Lesueur F, Nguyen-Dumont T, Pertesi M, Odefrey F, Hammet F, Neuhausen SL, John EM, Andrulis IL, Terry MB, Daly M, Buys S, Le Calvez-Kelm F, Lonie A, Pope BJ, Tsimiklis H, Voegele C, Hilbers FM, Hoogerbrugge N, Barroso A, Osorio A, Giles GG, Devilee P, Benitez J, Hopper JL, Tavtigian S V, Goldgar DE, Southey MC. Rare mutations in XRCC2 increase the risk of breast cancer. *Am J Hum Genet* 2012;90:734–9.

7 John EM, Hopper JL, Beck JC, Knight J a, 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. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Res* 2004;6:R375–89.

8 Antoniou AC, Casadei S, Heikkinen T, Barrowdale D, Pylkäs K, Roberts J, Lee A, Subramanian D, De Leeneer K, Fostira F, Tomiak E, Neuhausen SL, Teo ZL, Khan S, Aittomäki K, Moilanen JS, Turnbull C, Seal S, Mannermaa A, Kallioniemi A, Lindeman GJ, Buys SS, Andrulis IL, Radice P, Tondini C, Manoukian S, Toland AE, Miron P, Weitzel JN, Domchek SM, Poppe B, Claes KBM, Yannoukakos D, Concannon P, Bernstein JL, James P a., Easton DF, Goldgar DE, Hopper JL, Rahman N, Peterlongo P, Nevanlinna H, King M-C, Couch FJ, Southey MC, Winqvist R, Foulkes WD, Tischkowitz M. Breast-cancer risk in families with mutations in PALB2. *N Engl J Med* 2014;371:497–506.

9 Voegele C, Tavtigian S V, de Silva D, Cuber S, Thomas a, Le Calvez-Kelm F. A Laboratory Information Management System (LIMS) for a high throughput genetic platform aimed at candidate gene mutation screening. *Bioinformatics* 2007;23:2504–6.

10 Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, Lee A, Turnbull C, Rahman N, Fletcher O, Peto J, Gibson L, Dos Santos Silva I, Nevanlinna H, Muranen T a, Aittomäki K, Blomqvist C, Czene K, Irwanto A, Liu J, Waisfisz Q, Meijers-Heijboer H, Adank M, van der Luijt RB, Hein R, Dahmen N, Beckman L, Meindl A, Schmutzler RK, Müller-Myhsok B, Lichtner P, Hopper JL, Southey MC, Makalic E, Schmidt DF, Uitterlinden AG, Hofman A, Hunter DJ, Chanock SJ, Vincent D, Bacot F, Tessier DC, Canisius S, Wessels LF a, Haiman C a, Shah M, Luben R, Brown J, Luccarini C, Schoof N, Humphreys K, Li J, Nordestgaard BG, Nielsen SF, Flyger H, Couch FJ, Wang X, Vachon C, Stevens KN, Lambrechts D, Moisse M, Paridaens R, Christiaens M-R, Rudolph A, Nickels S, Flesch-Janys D, Johnson N, Aitken Z, Aaltonen K, Heikkinen T, Broeks A, Veer LJV, van der Schoot CE, Guénel P, Truong T, Laurent-Puig P, Menegaux F, Marme F, Schneeweiss A, Sohn C, Burwinkel B, Zamora MP, Perez JIA, Pita G, Alonso MR, Cox A, Brock IW, Cross SS, Reed MWR, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Henderson BE, Schumacher F, Le Marchand L, Andrulis IL, Knight J a, Glendon G, Mulligan AM, Lindblom A, Margolin S, Hooning MJ, Hollestelle A, van den Ouweland AMW, Jager A, Bui QM, Stone J, Dite GS, Apicella C, Tsimiklis H, Giles GG, Severi G, Baglietto L, Fasching P a, Haeberle L, Ekici AB, Beckmann MW, Brenner H, Müller H, Arndt V, Stegmaier C, Swerdlow A, Ashworth A, Orr N, Jones M, Figueroa J, Lissowska J, Brinton L, Goldberg MS, Labrèche F, Dumont M, Winqvist R, Pylkäs K, Jukkola-Vuorinen A, Grip M, Brauch H, Hamann U, Brüning T, Radice P, Peterlongo P, Manoukian S, Bonanni B, Devilee P, Tollenaar R a EM, Seynaeve C, van Asperen CJ, Jakubowska A, Lubinski J, Jaworska K, Durda K, Mannermaa A, Kataja V, Kosma V-M, Hartikainen JM, Bogdanova N V, Antonenkova NN, Dörk T,

Kristensen VN, Anton-Culver H, Slager S, Toland AE, Edge S, Fostira F, Kang D, Yoo K-Y, Noh D-Y, Matsuo K, Ito H, Iwata H, Sueta A, Wu AH, Tseng C-C, Van Den Berg D, Stram DO, Shu X-O, Lu W, Gao Y-T, Cai H, Teo SH, Yip CH, Phuah SY, Cornes BK, Hartman M, Miao H, Lim WY, Sng J-H, Muir K, Lophatananon A, Stewart-Brown S, Siriwanarangsan P, Shen C-Y, Hsiung C-N, Wu P-E, Ding S-L, Sangrajrang S, Gaborieau V, Brennan P, McKay J, Blot WJ, Signorello LB, Cai Q, Zheng W, Deming-Halverson S, Shrubsole M, Long J, Simard J, Garcia-Closas M, Pharoah PDP, Chenevix-Trench G, Dunning AM, Benitez J, Easton DF. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013;45:353–61, 361e1–2.

11 Meyer KB, O'Reilly M, Michailidou K, Carlebur S, Edwards SL, French JD, Prathalingham R, Dennis J, Bolla MK, Wang Q, de Santiago I, Hopper JL, Tsimiklis H, Apicella C, Southey MC, Schmidt MK, Broeks A, Van 't Veer LJ, Hogervorst FB, Muir K, Lophatananon A, Stewart-Brown S, Siriwanarangsan P, Fasching P a, Lux MP, Ekici AB, Beckmann MW, Peto J, Dos Santos Silva I, Fletcher O, Johnson N, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Marme F, Schneeweiss A, Sohn C, Burwinkel B, Guénel P, Truong T, Laurent-Puig P, Menegaux F, Bojesen SE, Nordestgaard BG, Nielsen SF, Flyger H, Milne RL, Zamora MP, Arias JI, Benitez J, Neuhausen S, Anton-Culver H, Ziogas A, Dur CC, Brenner H, Müller H, Arndt V, Stegmaier C, Meindl A, Schmutzler RK, Engel C, Ditsch N, Brauch H, Brüning T, Ko Y-D, Nevanlinna H, Muranen T a, Aittomäki K, Blomqvist C, Matsuo K, Ito H, Iwata H, Yatabe Y, Dörk T, Helbig S, Bogdanova N V, Lindblom A, Margolin S, Mannermaa A, Kataja V, Kosma V-M, Hartikainen JM, Chenevix-Trench G, Wu AH, Tseng C-C, Van Den Berg D, Stram DO, Lambrechts D, Thienpont B, Christiaens M-R, Smeets A, Chang-Claude J, Rudolph A, Seibold P, Flesch-Janys D, Radice P, Peterlongo P, Bonanni B, Bernard L, Couch FJ, Olson JE, Wang X, Purrington K, Giles GG, Severi G, Baglietto L, McLean C, Haiman C a, Henderson BE, Schumacher F, Le Marchand L, Simard J, Goldberg MS, Labrèche F, Dumont M, Teo S-H, Yip C-H, Phuah S-Y, Kristensen V, Grenaker Alnæs G, Børresen-Dale A-L, Zheng W, Deming-Halverson S, Shrubsole M, Long J, Winqvist R, Pylkäs K, Jukkola-Vuorinen A, Kauppila S, Andrulis IL, Knight J a, Glendon G, Tchatchou S, Devilee P, Tollenaar R a EM, Seynaeve CM, García-Closas M, Figueroa J, Chanock SJ, Lissowska J, Czene K, Darabi H, Eriksson K, Hooning MJ, Martens JWM, van den Ouweland AMW, van Deurzen CHM, Hall P, Li J, Liu J, Humphreys K, Shu X-O, Lu W, Gao Y-T, Cai H, Cox A, Reed MWR, Blot W, Signorello LB, Cai Q, Pharoah PDP, Ghoussaini M, Harrington P, Tyrer J, Kang D, Choi J-Y, Park SK, Noh D-Y, Hartman M, Hui M, Lim W-Y, Buhari S a, Hamann U, Försti A, Rüdiger T, Ulmer H-U, Jakubowska A, Lubinski J, Jaworska K, Durda K, Sangrajrang S, Gaborieau V, Brennan P, McKay J, Vachon C, Slager S, Fostira F, Pilarski R, Shen C-Y, Hsiung C-N, Wu P-E, Hou M-F, Swerdlow A, Ashworth A, Orr N, Schoemaker MJ, Ponder B a J, Dunning AM, Easton DF. Fine-scale mapping of the FGFR2 breast cancer risk locus: putative functional variants differentially bind FOXA1 and E2F1. *Am J Hum Genet* 2013;93:1046–60.

12 Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaeli B, Keduas V, Notredame C. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res* 2006;34:W604–8.

13 Taly J-F, Magis C, Bussotti G, Chang J-M, Di Tommaso P, Erb I, Espinosa-Carrasco J, Kemena C, Notredame C. Using the T-Coffee package to build multiple sequence alignments of protein, RNA, DNA sequences and 3D structures. *Nat Protoc* 2011;6:1669–82.

14 Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.6. *Distrib by author Dep Genome Sci Univ Washington, Seattle* 2005.

15 Tavtigian S V, Deffenbaugh a M, Yin L, Judkins T, Scholl T, Samollow PB, de Silva D, Zharkikh a, Thomas a. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* 2006;43:295–305.

16 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–5.

17 Stone E a, Sidow A. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Res* 2005;15:978–86.

18 Adzhubei I, Jordan DM, Sunyaev SR. Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. *Curr Protoc Hum Genet* 2013;Chapter 7:Unit7.20.