Supplementary Information

MATERIALS AND METHODS

Study subjects

Exome-sequenced HBOC patients

Clinical characteristics of 24 exome-sequenced HBOC patients are presented in Table 1.

Table 1. Clinical characteristics of 24 exome-sequenced patients.

Family id/ patient id	Gender	Cancer type (age at dg)	Cancer histological type, grade and hormone receptor status	Other cancers in the family (age at dg)
110				
110010	female	breast (26)	ductal, grade II, ER+, PR+, HER2+	prostate (64),
110007	female	breast (48)	ductal, grade II, ER+, PR+, HER2-	unknown (84)
235				
235054	female	breast (52)	lobular, grade II, ER-, PR+, HER2 na	-
235002	female	breast (56)	ductal, grade na, ER+, PR-, HER2+	
235004	female	bil breast (28)	ductal, gr III, ER na, PR na, HER2 na	
236				
236055	female	breast (29)	intraductal, grade na, ER na, PR na, HER2 na	-
236002	female	breast (52)	ductal, grade na, ER na, PR na, HER na	
240				
240059	female	breast (53)	ductal, grade III, ER+, PR-, HER2+	breast (62), thyroid (38)

240010	female	breast (42)	ductal, grade II, ER+, PR+, HER na	
260				
260075	female	breast (29)	ductal, grade III, ER-, PR-, HER2-	breast (65), lung (60)
260002	female	breast (58)	ductal, grade III, ER na, PR na, HER2 na	
271				
271310	female	breast (65)	lobular, grade II, ER+, PR+, HER2-	breast (55)
271009	female	bil breast (44, 50)	ductal in situ, grade I, ER+, PR+, HER2- and ductal, grade na, ER na, PR na, HER2 na	
271010	female	breast (66)	lobular, grade II, ER na, PR na, HER2 na	
TuFamBC21				
904001	female	breast (28)	ductal, grade III, ER-, PR-, HER2-	-
903001	female	breast (45)	na	
TuFamBC22				
906001	female	bil breast (59, 61)	ductal, grade I, ER+, PR+, HER2- and ductal, grade III, ER+, PR+, HER2-	3x breast (49, 50, 56)
TuFamBC23				
913001	female	breast (28)	ductal, grade III, ER-, PR+, HER2+	-
914001	female	bil breast (36, 38)	na	
TuFamBC24				
918001	female	breast (49)	ductal, grade III, ER-, PR-, HER2-	3x breast (-), 2 x ovary (-)
TuFamBC25				
908001	female	bil breast (41, 69), ovary (55)	ductal, grade II, ER+, PR+, HER2- and na	breast (-)
TuFamBC26				
909001	female	breast (64)	ductal, grade II, ER-, PR-, HER2-	3x breast (-)
907001	female	breast (47)	ductal, grade II, ER+, PR+, HER2-	
TuFamBC27				
910001	male	breast (72)	ductal, grade II, ER+, PR+, HER2-	breast (28)

Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

Genotyped female HBOC patients

A cohort of 129 female HBOC patients comprised index individuals from 65 HBOC families recruited from the Genetics Outpatient Clinic of Tampere University Hospital and index individuals from 64 HBOC families recruited from the Department of Clinical Genetics of Turku University Hospital. The clinical characteristics of 65 Tampere patients and 20 Turku patients have been previously described[1, 2], whereas the description of the remaining 44 HBOC patients collected from the Turku region have been presented below (Table 2). The Ethical Committees of Tampere and Turku University Hospitals and the National Authority for Medicolegal Affairs and the Auria Biobank have provided permission for the research project

Table 2. Clinical features of 44 HBOC patients collected from the Turku region.

Sample id	Cancer (age at dg)	Histology, grade	Receptor status	Family history of cancer (age at dg if known)
TuFam28	Breast (43)	Lobular, 2	ER+, PR+, HER2-	2x Breast (48, -)
TuFam29	Breast (62)	Ductal, 3	ER+, PR+, HER2-	5x Breast (32, 38, 39, 41, 45), Pancreas (46)
TuFam30	Bil. Breast (42, 64)	na and DCIS, 3	na	Breast (48)
TuFam31	Breast (67)	Ductal	ER+, PR+, HER2 na	7x Breast (48, 52, 61, 61, 70, 72, 75)
TuFam32	Breast (42)	Lobular, 2	ER+, PR+, HER2-	6x Breast (44, 47, 50, -, -, -)
TuFam33	Breast (46)	na	na	2x Breast (60, 80)
TuFam34	Breast (52)	Ductal, 2	ER-, PR+, HER2 na	Ovarian (42)
TuFam35	Bil. Breast (40, 40)	Ductal, 1 and Ductal, 3	ER+, PR+, HER2 -	3x Breast (43, 49, 70), Uterus (22)
TuFam36	Breast (50, 60)	Ductal tubular, 1 and DCIS, 2-3	na	4x Breast (35, 40, 45, 45)
TuFam37	Breast (48)	Lobular, 1	ER+, PR+, HER2-	6x Breast (50, 50, 51, 55, 60, 65)
TuFam38	Breast (28)	Ductal, 3	ER+, PR+, HER2+	Breast (78)
TuFam39	Breast (49)	Ductal, 2	ER+, PR+, HER2-	2x Breast (53, 65), Ovarian (86)
TuFam40	Breast (34)	Ductal, 3	ER+, PR-, HER2-	3x Breast (54, 61, -)
TuFam41	Breast (59)	Ductal, 1	ER+, PR+, HER2-	2x Breast (56, 58), Ovarian (27)
TuFam42	Breast (49)	Ductal, 3	ER+, PR+, HER2+	4x Breast (43, 54, 70, -)
TuFam43	Breast (49), Ovarian (77)	na and endometrioid, 2	na and ER-, PR+, HER2 na	3x Breast (62, 69, 69)
TuFam44	Breast (53, 62)	Lobular, 2 and Ductal, 3	ER-, PR-, HER2- and ER+, PR+, HER2+	6x Breast (49, 50, 50, 50, 52, 59)
TuFam45	Breast (58, 65)	Ductal tubular and Ductal, 2	ER+, PR+, HER2 na and ER+, PR+, HER2-	5x Breast (41, 48, 62, 69, 70)
TuFam46	Breast (41)	Ductal, 3	ER-, PR-, HER2-	3x Breast (59, 63), Ovarian (53, 74)
TuFam47	Breast (55)	Mucinous, 2	ER+, PR+, HER2-	5x Breast (45, 51, 55, 57, 81)
TuFam48	Breast (42)	Ductal, 2	ER+, PR+, HER2-	Breast (38)
TuFam49	Breast (44, 55)	Ductal, 1 and Ductal, 1	ER+, PR-, HER2-	2x Breast (47, 70)
TuFam50	Breast (51)	Lobular, 2	ER+, PR+, HER2-	2x Breast (46, 64)
TuFam51	Breast (36)	DCIS, na	na	2x Breast (65, -)
TuFam52	Breast (50)	Ductal, 2	ER+, PR+, HER2-	4x Breast (42, 50, 54, 54)

TuFam53	Breast (56)	Ductal, 2	ER+, PR+, HER2-	4x Breast (61, - , -, -), and Ovarian 2x (67, 78)
TuFam54	Breast (49)	Ductal, na	ER+, PR+, HER2 na	2x Breast (35, 59), Pancreas (59)
TuFam55	Breast (50, 62)	DCIS, na and Ductal, 3	ER+, PR-, HER2+	2x Breast (42, 56)
TuFam56	Breast (37, 46)	Ductal, 2 and Ductal, 2	na and ER+, PR+, HER2-	-
TuFam57	Breast (58)	Ductal, na	ER+, PR+, HER2-	3x Breast (49, 65, 65)
TuFam58	Breast (67)	Ductal, 3	ER+, PR+, HER2-	2x Breast (30, 58)
TuFam59	Breast (39)	Ductal, 3	ER+, PR+, HER2-	Breast (55)
TuFam60	Breast (46)	Ductal, 3	ER+, PR+, HER2-	Breast (36)
TuFam61	Breast (50, 68)	Ductal, 2	ER+, PR na and HER2 na	3x Breast (55, 63, 82)
TuFam62	Breast (53)	Ductal, 2	ER+, PR+, HER2-	3x Breast (33, 70, 70)
TuFam63	Breast (64)	Ductal, 1	ER+, PR+, HER2-	4x Breast (49, 51, 71, 81)
TuFam64	Breast (25)	Ductal, 3	ER+, PR-, HER2+	Breast (-)
TuFam65	Breast (48)	Ductal, 2	ER+, PR+, HER2-	3x Breast (39, -, -)
TuFam66	Breast (52)	Ductal, 1	ER+, PR+, HER2-	4x Breast (49, 52, 58, 60) and Ovarian (45)
TuFam67	Breast (34, 37)	Ductal, 3 and LCIS, na	ER-, PR-, HER2 na	Breast (61)
TuFam68	Bil. Breast (42, 44)	Ductal, 1	ER+, PR+, HER2-	2x Breast (32, 51)
TuFam69	Breast (25)	Ductal, 3	ER-, PR-, HER2-	2x Breast (50, 70)
TuFam70	Breast (33)	Ductal, 3	ER+, PR+, HER2-	Ovarian (60)
TuFam71	Breast (53)	Ductal, 3	ER+, PR+, HER2-	2x Breast (47, 60)

Abbreviations: DCIS, ductal carcinoma in situ; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor

Genotyped male BC patients

A cohort of 49 male BC patients comprised 45 Finnish male BC patients from the previously described cohort[3, 4] and 5 additional male BC patients collected from the Turku region (described below, Table 3).

Table 3. Clinical feature of 5 male breast cancer patients.

Sample id	Cancer (age at dg)	Histology, grade	Receptor status	Family history of cancer (age at dg if known)
MaleBC1	Breast (59)	na	na	Pancreas
MaleBC2	Breast (70)	Ductal, 2	ER+, PR+, HER2-	Breast (75)
MaleBC3	Breast (60)	Ductal, 2	ER+, PR+, HER2-	Pancreas (42)
MaleBC4	Breast (72)	Ductal, 3	ER+, PR+, HER2-	Breast (49)
MaleBC5	Bil. Breast (72, 72)	Ductal, 2 and Ductal, 2	ER+, PR+, HER2-	-

Abbreviations: BC, breast cancer; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor

Genotyped female and male population controls

Population controls were female and male blood donors whose blood samples were obtained from the Finnish Red Cross. The blood samples were collected from the Tampere, Turku, and Kuopio region during years 1997-1998.

Genotyped breast or breast and ovarian cancer patients

Breast or breast and ovarian cancer patients were *BRCA1/2*-negative Finnish females whose FFPE tumour and normal tissue samples were obtained from the Auria Biobank (Turku, Finland). The patient cohort was selected from the Auria Biobank according the high-risk hereditary breast cancer criteria.[1] DNA was extracted from FFPE tissue using a GeneRead DNA FFPE Kit (Qiagen Inc., Valencia, CA, USA).

Data-analysis

Reads were aligned with Bowtie2 against the human reference genome (hg19) using the default parameters.[5] Quality control was conducted using FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/). As a preprocessing step for variant calling, PCR duplicates and reads with a mapping quality of less than 10 were filtered out using SAMtools.[6] Variant calling was conducted using the Pypette bioinformatics toolkit, which was developed in house (https://github.com/annalam/pypette). Variants were annotated with Annovar[7] using the RefSeq genes as the reference gene set. Mapping details are presented in Table 4.

Table 4. Mapping details.

Sample	Total clean reads	Mapped reads	Percent of mapped reads	Uniquely mapped reads	Covered Refseq exon bases by at least by 1x(%)	Covered Refseq exon bases by at least 10x(%)	Covered Refseq exon bases by at least 20x(%)	Average Refseq exon base coverage
110007	60491188	60012949	99,20940716	47010716	66,33453428	48,78094263	43,30561723	24,76236669
110008	58209420	57771123	99,24703424	44867326	64,63549373	48,46726826	43,00549764	24,46697238
110010	60558184	60101188	99,24536046	47325260	66,61659597	49,01802351	43,52106517	24,69498302
235002	59193138	58759238	99,26697584	45379527	64,78921476	48,56253537	43,20337894	24,69769455
235004	61760640	61132611	98,9831242	45500133	65,30118782	48,68657899	42,90494705	25,080616
235054	58195244	57798309	99,31792536	45212375	64,95049156	48,46133148	42,96553328	24,6047982
236002	60808422	60334723	99,22099771	46484579	65,60290957	48,96960119	43,49112982	24,71913152
236055	58814280	58442017	99,36705338	45721740	64,71632852	48,07209154	42,6231848	24,46985691
240006	58810690	58175447	98,91985114	45531552	65,50264934	48,90654304	43,34547659	24,66973333
240010	88504896	82074829	92,7347895	58716602	66,91307831	46,59222692	42,0466292	28,50163429
240059	59980372	59552145	99,28605478	46347885	65,75528639	48,97443746	43,40291684	24,75347328
260002	59693288	59250280	99,25785961	46298341	66,16488766	48,80095397	43,20766301	24,84404285
260075	59088498	58659812	99,27450178	44972059	65,54688999	48,53900603	42,42204475	24,78177083
271009	59733374	59107365	98,95199458	45571318	65,40931233	48,90224989	43,33818006	24,5334661
271010	60092600	59393952	98,83738098	46510011	67,20396868	48,8489589	42,74072649	24,62669157
271310	61779310	61165179	99,00592771	47312682	66,28092182	49,10908821	43,64587747	25,17704649

901001	69302282	68682551	99,10575672	54382071	69,50733416	49,20009977	44,19181589	26,19568664
902001	69309726	68724458	99,15557594	54739562	70,10789093	48,74817629	43,43654648	25,23277051
903001	80511072	79808420	99,12726041	60853055	70,17648532	49,79009089	45,4032773	29,29334984
904001	71546050	70931520	99,14107068	58053244	72,16295398	48,59816769	43,34950919	25,4020256
905001	67551268	66920005	99,06550533	53830482	69,62274613	49,24161313	44,15172191	25,87519494
906001	64475984	63786899	98,93125322	53085265	71,84213604	48,25745248	42,56034182	23,4224397
907001	64794038	64194634	99,07490871	50641261	66,0475793	49,0222311	44,24810638	26,39046334
908001	71505520	70977311	99,26130318	58128806	72,42381826	48,60090276	43,28278916	25,32205747
909001	56773622	55311763	97,42510879	47131927	70,41996444	47,56020014	40,84741077	20,94776499
910001	71315452	70005927	98,16375699	54494257	67,43354952	49,3134559	44,76439463	27,73496163
911001	68414044	67127410	98,11934228	52915217	68,36603819	49,09700596	44,24888283	26,58449736
912001	78198362	76794384	98,20459411	58551580	68,63834697	49,82562729	45,72116445	29,61127415
913001	87259344	85998302	98,55483443	64659281	71,59798417	50,07045191	46,09756384	31,0629815
914001	94334968	93467950	99,08091557	70579887	74,7986169	50,08224379	46,00202322	30,93105567
915001	83917780	83235674	99,18717345	63516628	72,17614451	50,05195846	45,86124904	29,78729119
916001	73148646	72286833	98,82183328	57706345	72,31226131	49,60215822	44,6189475	26,29584592
917001	86584818	85865504	99,16923773	65202084	72,60745779	50,24442148	46,06804329	30,4530826
918001	68555520	67968552	99,14380636	54758769	70,4076385	49,80081079	44,83297995	26,23011572
919001	68544808	67915845	99,08240607	53985423	69,5044941	49,59081742	44,64440045	26,17306783
920001	70275238	69642454	99,09956335	55811498	71,09376132	49,4860748	44,39464494	26,0225829
921001	72088582	71519283	99,21027854	57521350	71,48915192	49,7356955	44,77809717	26,47399846

Based on the variant annotation, functional variants (splice-site Single Nucleotide Variants (SNVs)/indels, non-synonymous SNVs, stop gain/loss variants, and insertions/deletions (indels) inducing frameshift) were prioritized. Moreover, rare variants (Minor Allele Frequency (MAF) ≤ 0.05) were selected from this set of variants by screening them against the 1000 Genomes project[8] (August 2014 version) and Exome SequencingProject 6500 (ESP6500) (included in Annovar) (https://esp.gs.washington.edu/drupal/). In addition, the Sequencing Initiative Suomi (SISU) database was utilized to obtain information about allele frequencies in the Finnish population (http://www.sisuproject.fi/). From the remaining set of variants, only those variants for which the host genes participate in the DNA damage response pathway were retained. The pathway data were retrieved from the InPath database, which includes data integrated from several sources: KEGG, Wikipathways and BioCyc.[9]

To further narrow down the list of candidates, the deleteriousness of the remaining variants was evaluated by utilizing a precompiled set of pathogenicity predictions included in the Annovar ljb26_all dataset. This dataset covers predictions of the deleteriousness of all possible SNVs for 11 pathogenicity predictor methods and 3 methods evaluating the evolutionary conservation of the variant locus (Table 5). Only variants that were predicted to be deleterious by at least one of the predictors or that were frameshift indels were selected for further assessment. Furthermore, variants present only in healthy relatives were excluded.

Table 5. List of software included in the Annovar ljb26_all dataset.

Software/Method	Predictor type	Website / data source
SIFT	Pathogenicity predictor	http://sift.jcvi.org/
PolyPhen2 HVAR	Pathogenicity predictor	http://genetics.bwh.harvard.edu/pph2/
PolyPhen2 HDIV	Pathogenicity predictor	http://genetics.bwh.harvard.edu/pph2/
Mutation Assessor	Pathogenicity predictor	http://mutationassessor.org/
Mutation Taster	Pathogenicity predictor	http://www.mutationtaster.org/
LR	Pathogenicity predictor	http://genomics.usc.edu/members/15-member-detail/36-coco-dong
LRT	Pathogenicity predictor	Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19:1553–1561.
RadialSVM (MetaSVM)	Pathogenicity predictor	http://genomics.usc.edu/members/15-member-detail/36-coco-dong
VEST3	Pathogenicity predictor	http://karchinlab.org/apps/appVest.html
CADD	Pathogenicity predictor	http://cadd.gs.washington.edu/
FATHMM	Pathogenicity predictor	http://fathmm.biocompute.org.uk/
GERP++	Evolutionary conservation score	http://mendel.stanford.edu/SidowLab/downloads/gerp/

phyloP	Evolutionary conservation score	http://compgen.bscb.cornell.edu/phast/
SiPhy	Evolutionary conservation score	http://www.broadinstitute.org/genome_bio/siphy/

Variant validation

Sanger sequencing

Genomic DNA was amplified by PCR. PCR primers were designed with Primer3 software.[10, 11] Primer sequences and PCR conditions are available upon request. PCR products were visualized with 1.5% agarose gel. PCR products were purified by using rAPid Alkaline Phosphatase (Roche Diagnostics GmbH, Mannheim, Germany) and Exonuclease I (New England Biolabs, Ipswich, MA, USA). Sequencing was conducted by using a Big Dye Terminator v3.1 Cycle Sequencing Kit and and ABIPRISM 3130 × l Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) as indicated by the manufacturer. Sequences were analysed with Sequencher™ v5.1 software (Gene Codes Corporation, Ann Arbor, MI, USA).

TaqMan SNP genotyping

Genotyping was performed by using TaqMan® SNP Genotyping Assays (Applied Biosystems, Foster City, CA, USA) run on an ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) according to the manufacturers' instructions. The following assays were used: C_176281813_10 (*CDKN2A* c.496C>T), C___1244825_20 (*RAD52* c.538G>A), C___2283286_20 (*ATM* c.2572T>C), C__45273750_10 (*ATM* c.3161C>G), C_166902853_10 (*RPA2* c.122C>T), C__30585831_10 (*ATM* c.5558A>T), C__63879305_20 (*ATM* c.4424A>G), C__32333045_10 (*RAD50* c.280A>C), C_153129907_10 (*BRCA1* c.3904A>C), C__32376001_10 (*MYC* c.77A>G), C_102161615_10 (*RBL2*

c.1723G>C), C_167350917_10 (*NCOA3* c.3353A>C), C__15956024_20 (*PLAU* c.43G>T), C__15760210_10 (*RAD1* c.341G>A), C_168146154_10 (*WNT10A* c.337C>T), and C_190555357_10 (*WNT3A* c.277G>A). Genotypes were called with SDS v2.2.2 software (Applied Biosystems).

RESULTS

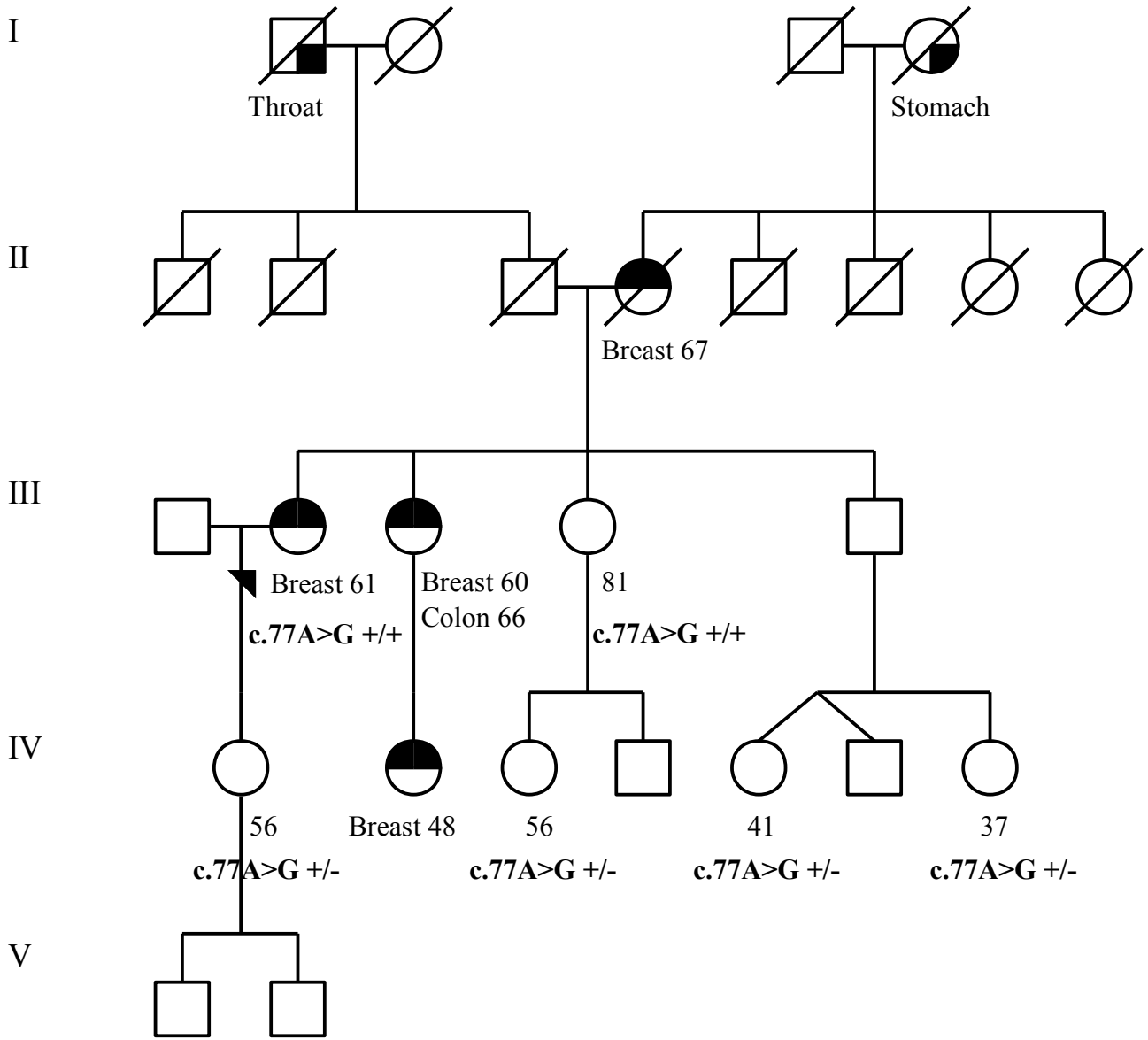
I
II
III
IV
V
Throat
Stomach
Breast 67
Breast 61
c.77A>G +/+
Breast 60
Colon 66
81
c.77A>G +/+
56
c.77A>G +/-
Breast 48
56
c.77A>G +/-
41
c.77A>G +/-
37
c.77A>G +/-

Figure 1. The *MYC* c.77A>G variant in pedigree 256. Genotyping identified the *MYC* c.77A>G variant altogether in a total of six individuals in family 256. The variant was present in its homozygous form (+/+) in the index patient and a healthy sister and in its heterozygous form (+/-) in four other healthy relatives in generation IV. No DNA samples were available from the affected relatives. Females are marked with circles, and males are marked with squares. The index patient is marked with an arrow. Breast cancer is marked with a half black circle and with the age at diagnosis. Other cancers are indicated with quarter black circles or squares. The current ages of healthy females are presented. Deceased individuals are indicated with a slash. Generations are presented in Roman numerals.

REFERENCES

1 Kuusisto KM, Bebel A, Vihinen M, Schleutker J, Sallinen SL: Screening for BRCA1, BRCA2, CHEK2, PALB2, BRIP1, RAD50, and CDH1 mutations in high-risk Finnish BRCA1/2-founder mutation-negative breast and/or ovarian cancer individuals. *Breast Cancer Res* 2011; **13:** R20.

2 Kuusisto KM, Akinrinade O, Vihinen M, Kankuri-Tammilehto M, Laasanen S, Schleutker J: Copy Number Variation Analysis in Familial *BRCA1/2*-Negative Finnish Breast and Ovarian Cancer. *PLoS ONE* 2013; **8:** e71802.

3 Syrjakoski K, Hyytinen ER, Kuukasjarvi T *et al:* Androgen receptor gene alterations in Finnish male breast cancer. *Breast Cancer Res Treat* 2003; **77:** 167-170.

4 Syrjakoski K, Kuukasjarvi T, Waltering K *et al:* BRCA2 mutations in 154 finnish male breast cancer patients. *Neoplasia* 2004; **6:** 541-545.

5 Langmead B, Salzberg SL: Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; **9:** 357-359.

6 Li H, Handsaker B, Wysoker A *et al:* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25:** 2078-2079.

7 Wang K, Li M, Hakonarson H: ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; **38:** e164.

8 1000 Genomes Project Consortium, Abecasis GR, Auton A *et al:* An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491:** 56-65.

9 Zhou H, Jin J, Zhang H, Yi B, Wozniak M, Wong L: IntPath--an integrated pathway gene relationship database for model organisms and important pathogens. *BMC Syst Biol* 2012; **6 Suppl 2:** S2-0509-6-S2-S2. Epub 2012 Dec 12.

10 Untergasser A, Cutcutache I, Koressaar T *et al:* Primer3--new capabilities and interfaces. *Nucleic Acids Res* 2012; **40:** e115.

11 Koressaar T, Remm M: Enhancements and modifications of primer design program Primer3. *Bioinformatics* 2007; **23:** 1289-1291.