ABSTRACT

Approximately half of the familial aggregation of breast cancer remains unexplained. A multiple-case breast cancer family exome-sequencing study identified three likely pathogenic mutations in RINT1 (NM_021930.4) not present in public sequencing databases: RINT1 c.343C>T (p.Q115X), c.1132_1134del (p.M378del), and c.1207G>T (p.D403Y). On the basis of this finding, a population-based case–control mutation-screening study was conducted that identified 29 carriers of rare (minor allele frequency < 0.5%), likely pathogenic variants: 23 in 1,313 early-onset breast cancer cases and six in 1,123 frequency-matched controls [OR, 3.24; 95% confidence interval (CI), 1.29–8.17; \(P = 0.013\)]. RINT1 mutation screening of probands from 798 multiple-case breast cancer families identified four additional carriers of rare genetic variants. Analysis of the incidence of first primary cancers in families of women carrying RINT1 mutations estimated that carriers were at increased risk of Lynch syndrome–spectrum cancers [standardized incidence ratio (SIR), 3.35; 95% CI, 1.7–6.0; \(P = 0.005\)], particularly for relatives diagnosed with cancer under the age of 60 years (SIR, 10.9; 95% CI, 4.7–21; \(P = 0.0003\)).

SIGNIFICANCE: The work described in this study adds RINT1 to the growing list of genes in which rare sequence variants are associated with intermediate levels of breast cancer risk. Given that RINT1 is also associated with a spectrum of cancers with mismatch repair defects, these findings have clinical applications and raise interesting biological questions. Cancer Discov; 4(7): 804–15. ©2014 AACR.

See related commentary by Ngeow and Eng, p. 762.
INTRODUCTION

Large population-based studies have established that women with a family history of breast cancer are at approximately 2- to 3-fold increased risk of the disease (1–3). Although there has been substantial progress in the last 20 years in identifying genetic causes of breast cancer, a considerable proportion of the familial risk remains unexplained. The current estimate of the proportion of familial risk explained by rare mutations in the high-risk breast
cancer susceptibility genes BRCA1 [Mendelian Inheritance in Man (MIM) 113705], BRCA2 (MIM 600185), and PALB2 (MIM 610355) is 20%, and another 28% is explained by associations with common genetic variants (4), although these proportions are highly dependent on age at diagnosis. A further approximately 5% is likely to be explained by mutations in genes in the homologous recombination repair detection and signaling pathways, including MRE11A (MIM 600814), RAD50 (MIM 604040), NBN (MIM 602667), ATM (MIM 607585), and CHEK2 (MIM 604373), and the RAD51 paralogs RAD51C (MIM 602774), RAD51D (MIM 602954), and XRCC2 (MIM 600375), which are all currently under investigation. The potential for intrafamily exome-sequencing approaches to identify additional breast cancer susceptibility genes has been demonstrated recently (5–7). Although large-scale validation of findings from these studies has provided further insight into the possible roles and clinical significance of these susceptibility genes in cancer predisposition, these studies have also illustrated the challenges posed in the context of breast cancer, a disease that has variable phenotypes and is likely to involve a great number of intermediate- to high-risk rare variants in multiple susceptibility genes (5, 6, 8–11).

In many, if not most, cases, mutations in specific cancer susceptibility genes are associated with high risks of cancer at one or two primary sites, but are also associated with more modest increases in risk for a number of other cancer types. For instance, BRCA2 mutations confer high risk of breast and ovarian cancers but are also associated with increased risk of prostate cancer with an estimated relative risk of 2.5 (12) and pancreatic cancer with an increased risk of 6-fold reported in two large independent studies (12, 13). Perhaps more noteworthy examples are the genes in which defects lead to deficiencies in mismatch repair that cause Lynch syndrome (MIM 120435); their defining associations are with colorectal and endometrial cancers, but they are also associated with a spectrum of other tumor types, including ovarian cancer, stomach cancer, bladder cancer, kidney cancer, pancreatic cancer, and brain tumors (14–16). Thus, it is plausible that mutations in other breast cancer susceptibility genes might also predispose to other cancers and should be considered in susceptibility gene discovery efforts.

Here, we conducted whole-exome sequencing (WES) of women affected with breast cancer at an early age from highly selected multiple-case breast cancer families, followed by gene-specific mutation screening using large international breast cancer research resources, and identified RAD50-interacting protein 1 (RINT1) as a putative breast cancer predisposition gene that also seems to be associated with risk for a spectrum of other cancers similar to those that have previously been described for Lynch syndrome.

RESULTS

WES

Bioinformatic analysis of the exome sequences of 89 women with early-onset breast cancer from 47 families identified three mutations in RINT1 (NM_021930.4): c.343C>T (p.Q115X), c.1132_1134del (p.378del), and c.1207G>T (p.D403Y), each occurring in a separate family and not found in the public sequencing databases. The nonsense substitution was carried by both women included in the WES from the family in which it was observed. The in-frame deletion (c.1132_1134del) and missense substitution (c.1207G>T; p.D403Y) were observed in one WES subject each, but both fall at protein positions that are highly evolutionarily conserved and are thus likely to be damaging to protein function (Figs. 1 and 2).

Additional Genotyping in WES Pedigrees

In the family in which RINT1 c.343C>T (p.Q115X) was identified, the mutation was carried by both women for whom WES was conducted (II-2 and I-4; Fig. 1A), and subsequent testing revealed that the two other female relatives diagnosed with breast cancer in the family (II-1 and II-4) and a male relative (I-2) affected by bladder and lung cancers were also carriers (Fig. 1A). The RINT1 c.1207G>T (p.D403Y) variant was identified in 1 of the 2 affected women selected for WES (III-8 but not III-1; Fig. 1B). This family had an extensive family history of cancer, including 9 diagnoses of breast cancer in 7 female members (2 women were found to be carriers (III-8 and II-10), 2 were found to be noncarriers (III-1 and II-12), and 3 could not be genotyped (I-1, II-7, and I-4) for c.1207G>T). The family also had 4 individuals (II-4, II-9, II-10, and III-10) who had malignant melanoma at early ages (39, 39, 36, and 33 years, respectively). Of these cases, 3 were carriers of c.1207G>T (II-4, II-9, and II-10) and 1 could not be genotyped. The RINT1 c.1132_1134del (p.378del) variant was carried by 1 of the 2 women selected for WES (II-1 but not III-1), whereas the carrier’s mother (I-4), affected by breast cancer at the age of 42 years, could not be genotyped. Of the 8 additional relatives diagnosed with other cancers, 7 could not be genotyped and 1, affected by breast cancer at the age of 32 years, did not carry this variant (Fig. 1C). Figure 1D shows the evolutionary conservation of the relevant amino acid positions for the missense (c.1207G>T; D403Y) and in-frame deletion (c.1132_1134del) variants described above.

On the basis of the findings from exome sequencing and pedigree follow-up, and the Gene Ontology–based plausibility of RINT1 as a breast cancer susceptibility gene, we conducted two further studies in parallel: (i) case-control mutation screening of RINT1 and mutation screening of RINT1 in a series of index cases from multiple-case breast cancer families; and (ii) RINT1 mutation screening in probands of multiple-case breast cancer families, as described in Methods.

Case–Control Mutation Screening

We identified 31 distinct rare variants [minor allele frequency (MAF) < 0.5%] as shown in Table 1. Table 2 shows the results of comparisons of the frequency of the various classes of rare sequence RINT1 variants in cases and controls. There was a statistically significant difference in the frequency of rare sequence variants (independent of in silico assessment) in cases compared with controls (OR, 2.2; P = 0.019). When we consider only those variants that are scored as potentially damaging by PolyPhen2 or Align-GVGD, 6 of 1,123 controls carried such a variant compared with 23 of 1,313 cases [OR, 3.24; 95% confidence interval (CI), 1.29–8.17; P = 0.013]. These variants were distributed across ethnicity as follows:
Caucasians, 15 pathogenic variants in cases versus 4 in controls; East Asians, 0 cases versus 1 control; Latinos, 5 cases versus 1 control; Recent African Ancestry, 3 cases versus 0 controls.

**RINT1 Screening in Probands from Multiple-Case Breast Cancer Families**

Screening in an additional 798 multiple-case breast cancer families identified five rare sequence variants in 6 families (MAF < 0.5% in all public sequencing database populations): the missense variants RINT1 c.413C>T (p.A138V) (2 families); c.1256C>T, p.P419R (1 family); and c.1385C>T, p.S462L (1 family); those latter two families were also included in the case–control mutation screening study through independent ascertainment. All of these variants are predicted to be “probably pathogenic” by PolyPhen2 or have a grade greater than C0 by Align-GVGD in silico analysis. In this set of families, we also identified the variants c.1333+1G>A (1) and c.129 C>A, p.V43V (1), predicted to affect splicing of the RINT1 transcript, as further described below.

**Figure 1.** Pedigrees of families found to carry mutations in RINT1 via WES. Mutation status is indicated for all family members for whom a DNA sample was available. A, pedi with c.343C>T (p.Q115X). B, pedi with c.1207G>T (p.D403Y). C, pedi with c.1132_1134del (p.M378del). Mutation carriers are indicated in black text; obligate carriers are indicated in gray italicized text. Cancer diagnosis and age at onset is indicated for affected members. *, DNA underwent WES; BC, breast cancer (black-filled symbols); UK, unknown age; BIC, bladder cancer; LC, lung cancer; LiC, liver cancer; KC, kidney cancer; BnC, brain cancer; CUP, cancer of unknown primary; NMSC, non-melanoma skin cancer; UC, uterine cancer (all gray-filled symbols); V, verified cancer (via cancer registry or pathology report); wt, wild-type; d., age at death (years). Some symbols represent more than one person as indicated by a numeral. D, RINT1 protein multiple sequence alignment covering positions M378 and D403. Positions L376 and K383 are the nearest invariant positions before and after M378; as the spacing between those positions is invariant, deletion of M378 would be expected to damage the protein. As position D403 is invariant, a nonconservative substitution at that position would also be expected to damage the protein.

**Assessment of Variants Predicted to Cause Aberrant Splicing**

The case–control mutation screening revealed a complex variant, c.1333-5delA;1334+1_1335delGTT, falling within the intron 9 splice acceptor sequence. Application of a minigene splicing assay demonstrated that the variant results in an in-frame delection of the first three nucleotides of RINT1 exon 10 (Supplementary Fig. S1). Screening in multiple-case families revealed two other sequence variants that were predicted to interfere with splicing: c.1333+1G>A and c.129 C>A, p.V43V. RT-PCR assays performed from lymphoblastoid cell line (LCL) cDNA derived from a carrier of 1333+1G>A revealed that this donor variant causes skipping of exon 9; because exon 9 is 226 nucleotides long, the result is a frameshift. RT-PCR assays performed from LCL cDNA
of all non-breast cancers (ii) above. The results of these analyses are presented in Table 3. In general, we see increased risks of Lynch syndrome–spectrum cancers in both the comparison of relatives in RINT1 mutation–carrying families to those in the families of cases without rare variants in RINT1, and in the comparison of risk to RINT1 mutation carriers with expected values based on population incidence rates. There was no evidence for increased risks associated with cancers that fall outside the Lynch syndrome/extended Lynch syndrome spectrum. Analyses with follow-up time restricted to before age 60 and/or incorporating genotyping data resulted in higher estimates of risk for the sites examined, which further strengthens the role of RINT1 in susceptibility to these cancers.

DISCUSSION

Here, we combined WES of women affected with breast cancer from highly selected multiple-case breast cancer families and additional mutation screening in large international breast cancer research resources to identify a new breast cancer predisposition gene, RINT1, that also seems to be associated with the risk for a spectrum of other cancers similar to those that have previously been described for Lynch syndrome.

Ideally, the transition from WES to detailed studies of individual candidate genes in a multistep susceptibility gene identification study such as this one would be driven by exome-wide statistical analysis of the WES data. After most of the analyses described in this work were completed, a new tool for pedigree-based studies, pVAAST (18), was available within the sequence analysis software package VAAST (19, 20). We performed a genome-wide analysis of the breast cancer exomes using pVAAST, accounting for the genealogic relationships between the subjects sequenced and comparing
Table 1. Distribution of RINT1 rare variants (with frequency < 0.5%) identified in the BCFR cases and controls

<table>
<thead>
<tr>
<th>Variant type</th>
<th>Variant</th>
<th>Effect on protein</th>
<th>Control</th>
<th>Case</th>
<th>Align-GVGD</th>
<th>PolyPhen2.1 (Q6NUQ1, Hum Div)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-frame del</td>
<td>c.1329delA; 1334-1_1335delGT</td>
<td>p.(Phe445_Ala446delinsSer)</td>
<td>0 1</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>In-frame del</td>
<td>c.2361G&gt;A</td>
<td>p.W787*</td>
<td>1 1</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.281C&gt;G</td>
<td>p.T94R</td>
<td>1 0</td>
<td>C65</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.376C&gt;T</td>
<td>p.H126Y</td>
<td>1 1</td>
<td>C0</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.388G&gt;A</td>
<td>p.A130T</td>
<td>3 3</td>
<td>C0</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.413C&gt;T</td>
<td>p.A138V</td>
<td>1 3</td>
<td>C65</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.501A&gt;T</td>
<td>p.Q167H</td>
<td>0 1</td>
<td>C15</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.532T&gt;C</td>
<td>p.Y178H</td>
<td>0 1*</td>
<td>C65</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.563C&gt;A</td>
<td>p.P246T</td>
<td>1 1</td>
<td>C0</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.778G&gt;T</td>
<td>p.A260S</td>
<td>0 1</td>
<td>C0</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.1025T&gt;C</td>
<td>p.S424C</td>
<td>1 0</td>
<td>C15</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.1256C&gt;G</td>
<td>p.P419R</td>
<td>0 1</td>
<td>C0</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.1270A&gt;T</td>
<td>p.S414C</td>
<td>0 1</td>
<td>C15</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.1385C&gt;T</td>
<td>p.S456L</td>
<td>0 1</td>
<td>C65</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.1449G&gt;T</td>
<td>p.M483I</td>
<td>1 0</td>
<td>C0</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.1465A&gt;G</td>
<td>p.I489V</td>
<td>0 1</td>
<td>C0</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.1562C&gt;T</td>
<td>p.T521I</td>
<td>0 1</td>
<td>C65</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.2036T&gt;C</td>
<td>p.V679A</td>
<td>0 1</td>
<td>C0</td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.2128C&gt;T</td>
<td>p.R710W</td>
<td>0 1</td>
<td>C65</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.2159G&gt;T</td>
<td>p.C720F</td>
<td>0 1</td>
<td>C65</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.2176T&gt;C</td>
<td>p.L662S</td>
<td>0 1</td>
<td>C65</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>c.2179T&gt;C</td>
<td>p.F727L</td>
<td>1 0</td>
<td>C15</td>
<td>Prob. Dam.</td>
<td></td>
</tr>
</tbody>
</table>

Total Rare Variants: 12, 31
Total Likely Pathogenic: 6, 23

*This case carries both p.Y178H and p.P788S. Only p.P788S was considered in the analyses. Variants highlighted in bold are deemed likely pathogenic, defined as A-GVGD score > C0 or PolyPhen “probably damaging.”

Abbreviation: BCFR, Breast Cancer Family Registry.

against 88 CEU [Utah Residents (CEPH) with Northern and Western European ancestry] and GBR (British in England and Scotland) control exomes from the 1000 Genomes Project. Post hoc, there was a significant excess of RINT1 variants (P = 0.029 based on a permutation test of cases and controls).

RINT1 was originally identified as a RAD50-interacting protein, and overexpression of truncated RINT1 was found to result in a defective radiation-induced G2–M checkpoint (21). Subsequently, RINT1 was also shown to be involved in several aspects of Golgi apparatus dynamics. The protein seems to play a role in pericentriolar positioning of the Golgi, Golgi assembly, and partitioning of the Golgi apparatus between daughter cells during mitosis (22–24). RINT1 is also a component of the NAG, RINT1, and ZW10 (NRZ) complex, which
likely acts as a tethering complex that harnesses dynein for retrograde trafficking of coat protein I (COPI)-coated vesicles from the Golgi to the endoplasmic reticulum (25, 26).

Nonetheless, neither the distribution of protein-truncating mutations nor the distribution of missense substitutions that we observed at conserved RINT1 residues provide immediate insight into which RINT1 function(s) are important to its role as a cancer susceptibility gene. It is noteworthy that 81% of heterozygous rint1 knockout mouse developed tumors during their lifespan (22).

The analyses of non-breast cancers in the RINT1 mutation–carrying families further extend our understanding of breast cancer susceptibility in the broader context of cancer susceptibility syndromes that include increased risk to breast cancer. The pattern of cancer susceptibility observed in the families included in this report is very similar to that in Lynch syndrome, which is associated with DNA mismatch repair deficiency and characterized by high increased risks of colorectal and endometrial cancers and more moderate increased risks of a number of other cancers, including ovarian, small intestine, and skin cancers (17).

Although not part of any formal analysis, Fig. 1 provides additional, anecdotal support of the hypothesis that RINT1 mutations are associated with the spectrum of Lynch syndrome–like cancers. These families were recruited and included in our exome-sequencing project because they have

---

**Table 2. Risk estimates and \( P \) values from RINT1 case-control mutation screening**

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Crude OR (95% CI)</th>
<th>Adj. OR (95% CI)</th>
<th>Adj. OR (95% CI)</th>
<th>Adj. OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All rare variants (incl. C0)</td>
<td>2.24 (1.14–4.38)</td>
<td>2.31 (1.17–4.55)</td>
<td>2.31 (1.16–4.62)</td>
<td>2.33 (1.16–4.65)</td>
</tr>
<tr>
<td>PolyPhen2 (prob damaging)</td>
<td>3.73 (1.06–13.13)</td>
<td>4.10 (1.3–14.18)</td>
<td>3.10 (0.85–11.30)</td>
<td>3.19 (0.88–11.59)</td>
</tr>
<tr>
<td>Align-GVGD (&gt;C0)(^a)</td>
<td>3.17 (1.28–7.85)</td>
<td>2.96 (2.23–3.92)</td>
<td>3.00 (1.18–7.63)</td>
<td>3.08 (1.22–7.81)</td>
</tr>
<tr>
<td>PolyPhen2 or Align-GVGD(^b)</td>
<td>3.32 (1.35–8.18)</td>
<td>3.53 (1.42–8.74)</td>
<td>3.16 (1.25–7.98)</td>
<td>3.24 (1.29–8.17)</td>
</tr>
<tr>
<td>PolyPhen2 and Align-GVGD(^c)</td>
<td>3.44 (0.97–12.23)</td>
<td>3.66 (1.02–13.12)</td>
<td>2.80 (0.76–10.36)</td>
<td>2.89 (0.78–10.64)</td>
</tr>
</tbody>
</table>

\(a\)In the binary analysis, only carriers of a missense substitution with grade > C0 or of an in-frame deletion (IFR) were considered. 
\(b\)Carriers of an IFR or carriers of a missense substitution with grade > C0 and predicted as probably damaging by PolyPhen2.

---

**Table 3. Analysis of incidence of first primary cancers in RINT1 mutation–carrying families**

<table>
<thead>
<tr>
<th>Site group(^d)</th>
<th>Cancers observed(^e)</th>
<th>RINT1 vs. wild-type</th>
<th>RINT1 vs. population rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RINT1</td>
<td>WT</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>All non-breast</td>
<td>51</td>
<td>2,117</td>
<td>1.67 (1.2–2.3)</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>13</td>
<td>363</td>
<td>2.51 (1.4–4.4)</td>
</tr>
<tr>
<td>Female</td>
<td>12</td>
<td>251</td>
<td>2.89 (1.6–5.3)</td>
</tr>
<tr>
<td>Lynch syndrome</td>
<td>18</td>
<td>375</td>
<td>3.24 (1.9–5.7)</td>
</tr>
<tr>
<td>Lynch syndrome dx &lt; 60</td>
<td>13</td>
<td>191</td>
<td>3.79 (2.2–6.5)</td>
</tr>
<tr>
<td>Lynch syndrome spectrum</td>
<td>30</td>
<td>823</td>
<td>2.44 (1.5–3.9)</td>
</tr>
<tr>
<td>Lynch syndrome spectrum dx &lt; 60</td>
<td>22</td>
<td>444</td>
<td>2.71 (1.8–4.0)</td>
</tr>
<tr>
<td>All but Lynch syndrome</td>
<td>33</td>
<td>1,742</td>
<td>1.33 (0.8–1.8)</td>
</tr>
<tr>
<td>All but Lynch syndrome spectrum</td>
<td>21</td>
<td>1,294</td>
<td>1.16 (0.9–1.9)</td>
</tr>
</tbody>
</table>

\(^a\)Number of cancers of each type observed in 16,038 person-years in family members of RINT1 probands and 914,178 person-years of observations in family members of wild-type probands.

\(^e\)See text for definition.

**Abbreviations:** HR, hazard ratio; SIR, standardized incidence ratio; dx, diagnosis.
multiple cases of breast cancer, but these families also have a number of other cancers, including bladder [age at diagnosis (dx) 68, 40, and 86 years] and uterine (dx 30 years) cancers, consistent with a Lynch syndrome–like (mismatch repair deficiency) phenotype. These families also have a number of other cancers that are not consistent with Lynch syndrome, including malignant melanoma (dx 39, 40, 36, and 33 years) and lung cancer (dx 81, 61, 43, and 66 years), for which our broader family-based analyses did not demonstrate significant enrichment in \( \text{RINT1} \) mutation–carrier families. In the 5 families from the targeted resequencing study of candidate genes in probands from the Kathleen Cunningham Foundation Consortium for research into Familial Breast Cancer (kConFab) resource, we also identified a number of cancers in relatives of the probands, including cancers of the colon (dx 47 years), small intestine (dx 63 years), bladder (dx 73 and 72 years), cervix (dx 26 years), and pancreas (dx 79 years).

In our sensitivity analyses, we observed some variation with regard to the choice of the year of start of follow-up. In general, results derived from beginning follow-up before 1920 showed smaller/less significant effects than those yielded from follow-up after 1920, until 1940 when the smaller sample sizes led to more unstable estimates. When we included control families in the analysis, the HR estimates were in general slightly lower, but in many cases were associated with lower \( P \) values. Finally, upon relaxing our criteria for inclusion of variants, we found that when comparing all rare \( \text{RINT1} \) variants, the HR estimates were broadly similar to prior analyses (in some cases higher), indicating that among the variants excluded from the main analysis, there are likely to be some that are associated with an increased risk. Functional analyses may be required to identify the relevant pathogenic variants from the entire set; this refinement would be expected to yield increased estimates of risk for associated cancers.

Because many of the cancers (particularly those at sites other than breast or ovary) are taken from reports of relatives (most often the family proband), we also repeated some analyses using only cancers verified by pathology report, death certificate, medical record, or self-report. Although this reduced the number of events by 50% to 70% depending on the site, the effects we observed remained statistically significant, albeit with smaller effect sizes. For example, the main Lynch syndrome HR went from 3.24 to 2.64 and the associated \( P \) value increased from 4 × 10\(^{-5}\) to 0.004. However, we note that a similar effect was seen in the analysis of cancers outside the Lynch syndrome spectrum, so that the relative magnitude of the estimates remained quite similar. Furthermore, the accuracy of reporting of relatives’ cancers is unlikely to be related to \( \text{RINT1} \) sequence variant status, as such genetic status was not known to the proband at the time of the reporting. Finally, the pattern of increased risk in Lynch syndrome–spectrum cancers compared with all other cancer sites was remarkably similar when the analyses were restricted to verified or self-report cancers only. However, the validation of \( \text{RINT1} \) as a susceptibility gene for other sites requires the sequencing of the gene in large series of cases and controls, as well as families who meet the Lynch syndrome–spectrum criteria but without mutations in the known mismatch repair genes.

Our report further illustrates the capacity of massively parallel sequencing in the discovery of additional breast cancer susceptibility genes when used with an appropriate study design. Our report also gives weight to the concept that information from mouse gene–phenotype relationships and other similar resources could complement a Gene Ontology–driven analysis and potentially be very helpful in identifying genes of interest from sequencing datasets acquired during studies of cancer susceptibility. On the basis of the data presented here, we would recommend that \( \text{RINT1} \) be added to genes currently being tested by targeted gene panels that are now increasingly being used in commercial/diagnostic genetic testing of patients with cancer and their families.

**METHODS**

**Subjects**

**WES** DNA samples from 89 women with breast cancer from 47 families were selected for exome capture followed by massively parallel sequencing (WES). These samples were contributed by a number of breast cancer research resources (27–29). Selection criteria for inclusion were three cases of breast cancer in the family diagnosed under the age of 50 years, with two cases who were more distant than first-degree relatives of each other having a DNA sample available for WES analysis. All participants provided written informed consent for participation in genetic studies. WES studies were approved by the University of Utah (Salt Lake City, UT) Institutional Review Board (IRB) and The University of Melbourne (Melbourne, VIC, Australia), Human Research Ethics Committee (HREC).

**Case–Control Mutation Screening** Eligible participants included women ascertained by population-based sampling by the Australia, Northern California, and Ontario sites of the Breast Cancer Family Registry (BCFR) between 1995 and 2005 (27). For the Ontario and Northern California sites, cases diagnosed at the age of 35 years and older or 36 years and older, respectively, that met high-risk criteria (bilaterality, Ashkenazi origin, family history) were preferentially included; and cases from the Australian site oversampled cases with early age at diagnosis. For the present study, selection criteria for cases (\( n = 1,131 \)) were diagnosis of breast cancer at or before 45 years of age 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 (\( n = 1,123 \)) were frequency-matched to the cases within each center by racial/ethnic group, with age at selection not more than 10 years older or younger than the age range at diagnosis of the cases ascertained at the same center. Because of the shortage of available controls in some ethnic and age groups, the frequency matching was not one-to-one in all subgroups. The design of this study has been described in detail previously (6, 30–32). Recruitment and genetic studies were approved by the Ethics Committee of the International Agency for Research on Cancer (IARC; Lyon, France), the University of Utah IRB, and the local IRBs of the BCFR centers from which we received samples. Written informed consent was obtained from each participant. Distribution of cases and controls by center and ethnicity is shown in Supplementary Table S1.

**Multiple-Case Breast Cancer Families** Index cases from 798 multiple-case breast or breast/ovarian cancer families were selected for mutation screening of \( \text{RINT1} \). These consisted of 114 youngest affected members of families participating in the Australian Breast Cancer Family Study (ABCFS; ref. 27) and 684 youngest affected members of families participating in the kConFab familial breast cancer resource (28), selected on the basis of strength of family history and availability of DNA samples. Informed consent was as described for the women whose samples were selected for exome sequencing described above.
Laboratory and Bioinformatic Methods

WES Capture was performed using SeqCap EZ Exome v2.0 (Roche Nimblegen; n = 86) and TrueSeq Exome Enrichment Kit (Illumina; n = 3). Sequencing was performed using SOLiD 3.5 (n = 8), SOLiD 4 (n = 28), 5500xl SOLiD (n = 12; Life Technologies), GAIX (n = 2), and HiSeq (n = 39) instruments (Illumina). Mapping to the human reference genome build hg19 was performed using BioScope v1.2 (SOLiD 3.5 data), BioScope v1.3 (SOLiD 4 data), LifeScope 2.5.1 (5500xl SOLiD data), bwa 0.6.2 (GAIX data; ref. 33), Novoalign (HiSeq data; n = 40), and bwa 0.7.5a (HiSeq data; n = 3). Local realignment was performed using the GATK 1.5-25 RealignerTargetCreator and IndelRealigner modules, duplicates were removed using Picard MarkDuplicates 1.29, and sorting and indexing were performed using SAMtools 0.1.8. Variant calling was conducted using the GATK 1.5-25 UnifiedGenotyper module. Variant calls were filtered to the capture target regions using BEDTools 2.16.2 (34) and bed files were provided by the manufacturers. Further filtering was performed using Filtering and Annotation of Variants that are Rare (FAVR; ref. 35). Annotation was performed using ANNOVAR (36).

Case-Control Mutation Screening
For mutation screening of the coding exons and proximal splice junction regions of RINT1 (NM021930), we used 30 ng of mixed whole-genome amplified (WGA) DNA obtained from two independent WGA reactions. The laboratory process was as described in detail for our recent studies of ATM, CHEK2, XRCC2, and RAD51 (6, 30-32). Our semi-automated approach, handled by a Laboratory Information Management System (LIMS; refs. 37, 38), relies on mutation scanning by High-Resolution Melt (HRM) curve analysis followed by direct Sanger sequencing of the individual samples for confirmation of the sequence variant. In a previous work, we showed that the HRM technique performs with high sensitivity and specificity (1.0 and 0.8, respectively, for amplifications of <400 bp) for mutation screening by comparing the results with those obtained with Sanger sequencing (39).

All exonic sequence variants, plus intronic sequence variants that fell within 20 bp of a splice acceptor or 8 bp of a splice donor and were either unreported or had an allele frequency of <0.5% in the large-scale reference groups “Caucasian Americans,” “African Americans,” and “East Asians” based on exome variant server (EVS) and 1000 Genomes Project (1000G) data, were re-amplified from genomic DNA or from the two independent WGA reaction products to confirm the presence of the variant.

All samples that failed at either the primary PCR, secondary PCR, or sequencing reaction stage were re-amplified from WGA DNAs or genomic DNAs. Primer and probe sequences are available in Supplementary Table S2.

Sequence Analysis of Multiple-Case Breast Cancer Families
The 114 probands from the ABCFS study were screened by HRM followed by Sanger sequencing confirmation of aberrant melt curves as described above. The 684 cases from the kConFab study were screened by LRMS, followed by Sanger sequencing of the 15 coding and flanking exonic regions of RINT1 (chr07:105,172,532-105,208,124; hg19) followed by 100 bp paired-end sequencing on HiSeq 2000 (Illumina) at AxaX SEQ Technologies, Inc./Macrogen. Intron 9, exon 10, and 924 bp of intron 10 were amplified from the human RINT1-bearing BAC clone RP11-96R13 (BACPAC Resources Center, Children’s Hospital, Oakland, CA) and inserted into the exon trapping vector pSPL3 (a generous gift from Dr. Bernd Wissinger and Dr. Nicole Weisschuh, Centre for Ophthalmology University Clinics Tuebingen, Germany). The mutation was introduced into an intron 9 PCR product by transformation into E. coli and the DNA was then amplified by PCR using primers designed to recognize the mutation. The resulting PCR product was used as the template for the amplification of genomic DNA from the proband.

Alignment of Phylogenetic Data
We used M-Coffee, which is part of the Tree-based Consistency Objective Function for alignment Evaluation (T-Coffee) software suite of alignment tools, to prepare a protein multiple sequence alignment for RINT1 to predict the effect of missense substitutions. The alignment included sequences from 17 species: Homo sapiens, Macaca mulatta, Callithrix jacchus, Mus musculus, Bos taurus, Loxodonta africana, Dasypus novemcinctus, Monodelphis domestica, Ornithorhyncus anatinus, Gallus gallus, Xenopus tropicalis, Latimeria chalumnae, Danio rerio, Branchiostoma lanceolatum, Strongylocentrotus purpuratus, Nematomorpha venustus, and Trichoplax adhaerens. The alignment was characterized using the Protpars routine of Phylogeny Inference Package version 3.2 software (PHYLIB; ref. 40) to make a maximum parsimony estimate of the number of substitutions that occurred along each clade of the underlying phylogeny. The sequence alignment, or updated versions thereof, is available at the Align-GVGD website. Missense substitutions observed during our mutation screening of RINT1 were scored using Align-GVGD with our curated alignment, and with PolyPhen2 using its precompiled alignments.

In Silico Prediction of Splicing
All exonic sequence variants, plus intronic variants detected in the vicinity of the splice junction sequences, with allele frequencies <0.5%, were scored for their potential impact on splicing using MaxEntScan (MES), which computes the maximum entropy score of a given sequence using splice site models trained on human data (41). We calibrated MES by calculating the average and standard deviation of MES scores for the wild-type splice junctions in BRCA1, BRCA2, and ATM, allowing us to convert raw MES scores into z-scores. On the basis of BRCA1 and BRCA2 mutation screening data used previously to calibrate Align-GVGD (42, 43), we found that rare variants that fall within the acceptor or donor region and reduce the MES score for the splice signal in which they fall show an approximately 95% probability to damage splice junction function when they result in a calibrated MES score of z < -2, or approximately 40% probability when they result in a calibrated MES score of -2 < z < -1. In addition, exonic rare variants that increase the MES donor score of their sequence context and result in a calibrated MES donor score of z > 0 have an approximately 64% probability to create a de novo donor, whereas if they result in a calibrated MES score of -2 ≤ z ≤ 0, the probability to create a de novo donor is approximately 30% (Vallee and colleagues, manuscript in preparation). These MES-based rules were used to identify rare sequence variants that are likely to alter RINT1 gene mRNA splicing.

Assessment of c.129C>A (p.V43M) and c.1333+1G>A RINT1 Transcripts
Epstein–Barr virus-transformed lymphoblastoid cell lines (LCL) bearing either reference sequence RINT1, RINT1 c.129C>A or RINT1 c.1333+1G>A were cultured without or with cycloheximide (to stabilize transcripts sensitive to decay) and prepared for RNA extraction as described previously (44, 45). To guard against the amplification of genomic DNA, the reverse transcription primer and all the PCR primers except one (due to restrictions in length) were designed to span RINT1 exon–exon boundaries. cDNA was synthesized using the Thermoscript RT-PCR System Kit (Life Technologies), and RINT1 products were PCR-amplified using AmpliTaq Gold DNA Polymerase (Life Technologies). Non–reverse-transcribed templates (extracted total RNA added in place of cDNA during PCR) were single-stranded to confirm the presence of the variant.

Minigene Assay Characterization of RINT1 c.1334-5delA (852)
We used a minigene assay to characterize the effect of c.1334-5delA (852) on RINT1 RNA splicing. A region spanning some or all of RINT1 intron 9, exon 10, and 924 bp of intron 10 was amplified from the human RINT1-bearing BAC clone RP11-96R13 (BACPAC Resources Center, Children’s Hospital, Oakland, CA) and inserted into the exon trapping vector pSPL3 (a generous gift from Dr. Bernd Wissinger and Dr. Nicole Weisschuh, Centre for Ophthalmology University Clinics Tuebingen, Germany). The mutation was introduced into an intron 9 PCR product by transformation into E. coli and the DNA was then amplified by PCR using primers designed to recognize the mutation. The resulting PCR product was used as the template for the amplification of genomic DNA from the proband.
incorporating c.1334-5delA;1334-1, 1335delGTG) into an intron 9 reverse primer; the mutation was then inserted into the minigene construct by in vitro homologous recombination using the Cold Fusion Cloning Kit (System Biosciences).

For expression assays, the minigene constructs were transfected into both COS-7 and HEK293 cells using TurboFect (Thermo Scientific) for COS-7, or FuGene 6 (Promega) for HEK293, following the manufacturers’ recommended protocols. Transfected cells were cultured for 44 hours and then cultured for an additional 4 hours without or with cycloheximide (to stabilize transcripts bearing frameshift mutations). Total RNA was isolated using RNeasy kits (Qiagen), treated with Turbo DNase (Life Technologies) to remove contaminating genomic DNA, and then reverse-transcribed using SuperScript VILO (Life Technologies). Minigene transcripts were PCR-amplified using NEBNext High-Fidelity 2x PCR Master Mix (New England Biolabs) and then separated by agarose gel electrophoresis. PCR products from the mutation-building construct were cloned into the pCR-BluntII-TOPO vector (Life Technologies), and clone sequences were evaluated by Sanger sequencing. Primer sequences for all transcript assessments are shown in Supplementary Table S3.

Statistical Analyses

Case–Control Analysis Missense variants in RINT1 identified by HRM/Sanger sequencing were assessed using the in silico analysis tools Align-GVGD and PolyPhen2 (see above) for missense variants. Sequence variants were classified as pathogenic by Align-GVGD if they were any grade higher than C0 (equivalent to being conserved in mammals), and by PolyPhen2 if assessed as “probably damaging.” In addition, sequence variants predicted to result in a truncated protein were considered pathogenic (although all potential splice variants were validated experimentally as described above). The frequencies of different groupings of variants among cases and controls based on these definitions were compared by unconditional logistic regression, adjusting for ethnicity and study center.

Analysis of Cancers in RINT1 Mutation–Carrier Families We compared the incidence of cancers at all sites in the family members of the 23 case probands with likely pathogenic RINT1 rare variants to that in the family members of 1,290 case probands without any such variants. From these 1,313 families, we excluded: three families who did not have pedigree data; 19 families in which the proband had been subsequently found to carry pathogenic mutations in BRCA1 (6 families), BRCA2 (6), and the specific variants PALB2 c.31133113G>A (p.W1038* (4) and ATM c.7271T>G (p.V2424G) (3), known to be associated with high risk of breast cancer (46, 47), and 8 families in which the proband was found to carry a rare RINT1 missense variant that was not scored as pathogenic by either in silico prediction tool.

Probands and their relatives were censored at the earliest of their age at last contact/vital status information, age at death, or age at first invasive cancer (at any site). We excluded from this latter comparison set individuals with missing year of birth or who were born before 1920, as diagnostic and censoring information was likely to be less reliable in this cohort. With these restrictions, the dataset consisted of 373 members of 23 RINT1 mutation-carrying families with a total of 16,037 person-years, and 20,918 members of 1,260 families without RINT1 mutations comprising 914,178 person-years. In these analyses, cancers at sites other than breast diagnosed in probands were included if they had occurred before the breast cancer diagnosis. For these analyses and those described below, we used the definition of “likely pathogenic” as any rare variant that was scored by either of the two in silico methods used for variant assessment (corresponding to line 4 of Table 2).

In a first exploratory analysis, we used a simple incidence rate ratio approach to assess differences in incidences between members of families in which the proband carried a likely pathogenic RINT1 variant and members of families not harboring rare variants in RINT1 to examine the overall incidences of cancers other than breast cancers, according to the following topologic site groups as defined by the International Classification of Diseases for Oncology (ICD-O): oral cavity and pharynx (C00–14), digestive organs (C15–26), respiratory organs (C30–39), bone (C40–41), skin (C43–44), soft tissue (C45–49), female genital organs (C51–58), male genital organs (C60–63), urinary tract (C64–68), eye/brain/CNS (C69–72), thyroid/endocrine (C73–75), and lymphoid tissues (C81–96).

In more detailed analyses, we examined the incidence of first primary cancers using two different approaches. First, we used the Cox proportional hazards model in a retrospective cohort analysis to estimate the hazard ratio (HR) for occurrence of cancer by site group among members of families of probands found to carry a RINT1 likely pathogenic sequence variant compared with members of families not carrying a rare RINT1 variant. These analyses included adjustment for study center and used a robust variance estimate to account for clustering of cancers within families. We obtained DNA samples from 36 family members of probands carrying a RINT1 rare variant from the relevant study centers to enable genotyping of these individuals for the variant identified in the index case of the family. These genotypes were used to inform the estimation of carrier probabilities for our second set of analyses as described below.

In this analysis, we estimated carrier probabilities for each individual using the program MENDEL (48) based on the observed phenotypes, the family genotypes, and assuming the HR for breast cancer as estimated from the case-control analysis. We then calculated a weighted standardized incidence ratio (SIR) by comparing age-, sex-, site-, and registry-based incidence rates averaged over the period 1988–2002 using data from Cancer Incidence in Five Continents versions 7, 8, and 9 (IARC). Cancer registries were matched to the location/ethnicity of the cases as follows: Canada: Ontario; Australia: Victoria and New South Wales; California: Hispanic, East Asian, Black, with weights equal to the probability of being in the given cancer type/group, in the nth individual, \( E_n \) is the cumulative incidence for that individual based on their censoring age, gender, and residence, and \( w \) is the estimated carrier probability for the nth individual.

Exact \( P \) values were calculated assuming the (estimated) observed number of cases in carriers; \( \sum O_i \) follows a Poisson distribution with mean \( \sum E_i \), 95% CIs were calculated using the accurate Poisson-based approximation as detailed in Rothman and Boice (49).

We also performed several sensitivity analyses by varying the inclusion criterion of birth year after 1920 using a range from 1900 to 1940, using a broader definition of pathogenicity for the observed sequence variants than that in the main analyses, and incorporating the families of controls with and without RINT1 variants. All statistical analyses were performed using STATA v 12.0 (StataCorp).

Disclosure of Potential Conflicts of Interest

S.V. Tavtigian is an inventor on BRCA1 and BRCA2 patents and has received royalties on these patents from the NIH. No potential conflicts of interest were disclosed by the other authors.

Disclaimer

The content of this article does not necessarily reflect the views or policies of the NCI or any of the collaborating centers in the BCPR, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government or the BCPR.

Authors’ Contributions


S.V. Tavtigian

M.C. Southey

D.E. Goldgar

D.J. Park

G.G. Giles

J.L. Hopper

Downloaded from cancerdiscovery.aacrjournals.org on October 16, 2017. © 2014 American Association for Cancer Research.

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D.J. Park, F. Le Calvez-Kelm, H. Tsimitkis, Z.L. Teo, L.B. Thingholm, E.L. Young, J. Li, E.M. John, I.L. Andruleis, M.B. Terry, M. Daly, S.S. Buys, C. Snyder, P. Devilee, J.L. Hopper, M.C. Southey


Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D.J. Park, K. Tao, N. Robinot, F. Hammet, F. Odefrey, Z.L. Teo, C. Voegele, A. Lomie, J. Ellis, C. Snyder, P. Devilee, G.G. Giles, J.L. Hopper, F. Lesueur, S.V. Tavtigian, M.C. Southey

Study supervision: F. Le Calvez-Kelm, J.L. Hopper, S.V. Tavtigian, M.C. Southey, D.E. Goldgar

Carried out the laboratory experiments: N. Robinot

Acknowledgments

The Australian BCFR (ABCFS; 1992–1995) was supported by the Australian National Health and Medical Research Council (NHMRC), the New South Wales Cancer Council, and the Victorian Health Promotion Foundation (Australia). The authors thank Margaret McCredie for her key role in the establishment and leadership of the ABCFR in Sydney, Australia, and the families who donated their time, information, and biospecimens. The Genetic Epidemiology Laboratory at the University of Melbourne has also received generous support from B. Hovey and Dr. R.W. Brown, to whom the authors are most grateful. The authors also thank G. Keough for assistance with article preparation.

The work of the BCFR centers [BCFR-AU (ABCFS), BCFR-NC, BCFR-NY, BCFR-ON, BCFR-PA (FCCCS), and BCFR-UT] was supported by grant UM1 CA164920 from the National Cancer Institute.

The work of BCFR-ON was additionally supported by the Canadian Institutes of Health Research “CIHR Team in Familial Risks of Breast Cancer” program.

The authors thank Heather Thorne, Eveline Niedermaier, all the research nurses and staff of the Kathleen Cunningham Foundation Consortium for research into Familial Breast Cancer (kComFab), and the heads and staff of the Family Cancer Clinics and the Clinical Follow Up Study (funded by NHMRC grants 145684, 288704, and 454508) for their contributions to this resource, and the many families who contribute to kComFab. kComFab has been supported by grants from the National Breast Cancer Foundation (Australia), the NHMRC, the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, the Cancer Foundation of Western Australia, and the NIH (via U01 CA69638).

M.C. Southey is an NHMRC Senior Research Fellow. J.L. Hopper is an NHMRC Senior Principal Research Fellow. T. Nguyen-Dumont is a Susan G. Komen for the Cure Postdoctoral Fellow.

Grant Support

This work was supported by the Cancer Council Victoria (Grant ID 628774 to M.C. Southey); the NIH (R01CA155767 to D.E. Goldgar, S.V. Tavtigian, and M.C. Southey, R01CA121245 to S.V. Tavtigian, and P30CA042014); The Australian NHMRC (Grant IDs 466668, 509038, and APP1025145); The Victoria Breast Cancer Research Consortium; The University of Melbourne (infrastructure award to J.L. Hopper); a Victorian Life Sciences Computation Initiative (VLSI) grant (number VR00353) on the Peak Computing Facility at the University of Melbourne, an initiative of the Victorian State Government; and by the Government of Canada through Genome Canada, the Canadian Institutes of Health Research, and the Ministère de l'enseignement supérieur, de la recherche, de la science, et de la technologie du Québec through Génome Québec.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 26, 2014; revised April 19, 2014; accepted April 28, 2014; published OnlineFirst May 2, 2014.

REFERENCES


RINT1 Sequence Variants and Cancer Predisposition

Rare Mutations in RINT1 Predispose Carriers to Breast and Lynch Syndrome–Spectrum Cancers


Cancer Discovery 2014;4:804-815. Published OnlineFirst May 2, 2014.

Updated version
Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-14-0212

Supplementary Material
Access the most recent supplemental material at:
http://cancerdiscovery.aacrjournals.org/content/suppl/2014/05/01/2159-8290.CD-14-0212.DC1

Cited articles
This article cites 47 articles, 8 of which you can access for free at:
http://cancerdiscovery.aacrjournals.org/content/4/7/804.full#ref-list-1

Citing articles
This article has been cited by 6 HighWire-hosted articles. Access the articles at:
http://cancerdiscovery.aacrjournals.org/content/4/7/804.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.