Cell, Vol. 115, 523–535, Novmeber 26, 2003, Copyright ©2003 by Cell Press

EMSY Links the BRCA2 Pathway to Sporadic Breast and Ovarian Cancer

Luke Hughes-Davies,[1,2,13,*] David Huntsman,[3,13] Margarida Ruas,[1,13] Francois Fuks,[1] Jacqueline Bye,[4] Suet-Feung Chin,[2] Jonathon Milner,[1] Lindsay A. Brown,[3] Forrest Hsu,[3] Blake Gilks,[3] Torsten Nielsen,[3] Michael Schulzer,[5] Stephen Chia,[6] Joseph Ragaz,[6] Anthony Cahn,[4] Lori Linger,[2] Hilal Ozdag,[2] Elena Cattaneo,[2] E.S. Jordanova,[7] Edward Schuuring,[7,8] David S. Yu,[2,9] Ashok Venkitaraman,[2,9] Bruce Ponder,[2] Aidan Doherty,[10] Samuel Aparicio,[2] David Bentley,[4] Charles Theillet,[11] Chris P. Ponting,[12] Carlos Caldas,[2,*] and Tony Kouzarides[1,*]

[1]Cancer Research UK/Wellcome Trust Institute and Department of Pathology
Tennis Court Road
Cambridge CB2 1QR
United Kingdom

[2]Department of Oncology
University of Cambridge
Hutchison/MRC Research Centre
Hills Road
Cambridge CB2 2XZ
United Kingdom

[3]Genetic Pathology Evaluation Centre of the Department of Pathology and Prostate Center at the Vancouver General Hospital
Department of Pathology, British Columbia Cancer Agency, and the University of British Columbia
Room 509, JBRC
2660 Oak Street
Vancouver, British Columbia V6H 3Z6
Canada

[4]The Wellcome Trust Sanger Institute
Hinxton, Cambridge CB10 1SA
United Kingdom

[5]Department of Statistics and Medicine
University of British Columbia
221 Wesbrook Mall
Vancouver, British Columbia V6T 2B5

[6]Department of Medical Oncology
British Columbia Cancer
600 West 10th Avenue
Vancouver, British Columbia, V5Z 4E6
Canada

[7]Department of Pathology
Leiden University Medical Center
PO Box 9600
2300 RC Leiden
The Netherlands

[8]Department of Pathology
University Hospital Groningen
PO Box 30001
9700 RB Groningen
The Netherlands

[9]MRC Cancer Cell Unit
Hutchison/MRC Research Centre
Hills Road
Cambridge, CB2 2XZ
United Kingdom

[10]Cambridge Institute for Medical Research
Cambridge, CB2 2XY
United Kingdom

[11]INSERM EMI 229
Centre Val d'Aurelle
Montpellier
France

[12]MRC Functional Genetics Unit
University of Oxford
Oxford OX1 3QX
United Kingdom

*Correspondence: lh205@cus.cam.ac.uk (L.H.-D.), cc234@cam.ac.uk (C.C.)., tk106@mole.bio.cam.ac.uk (T.K.)

[13]These authors contributed equally to this work.

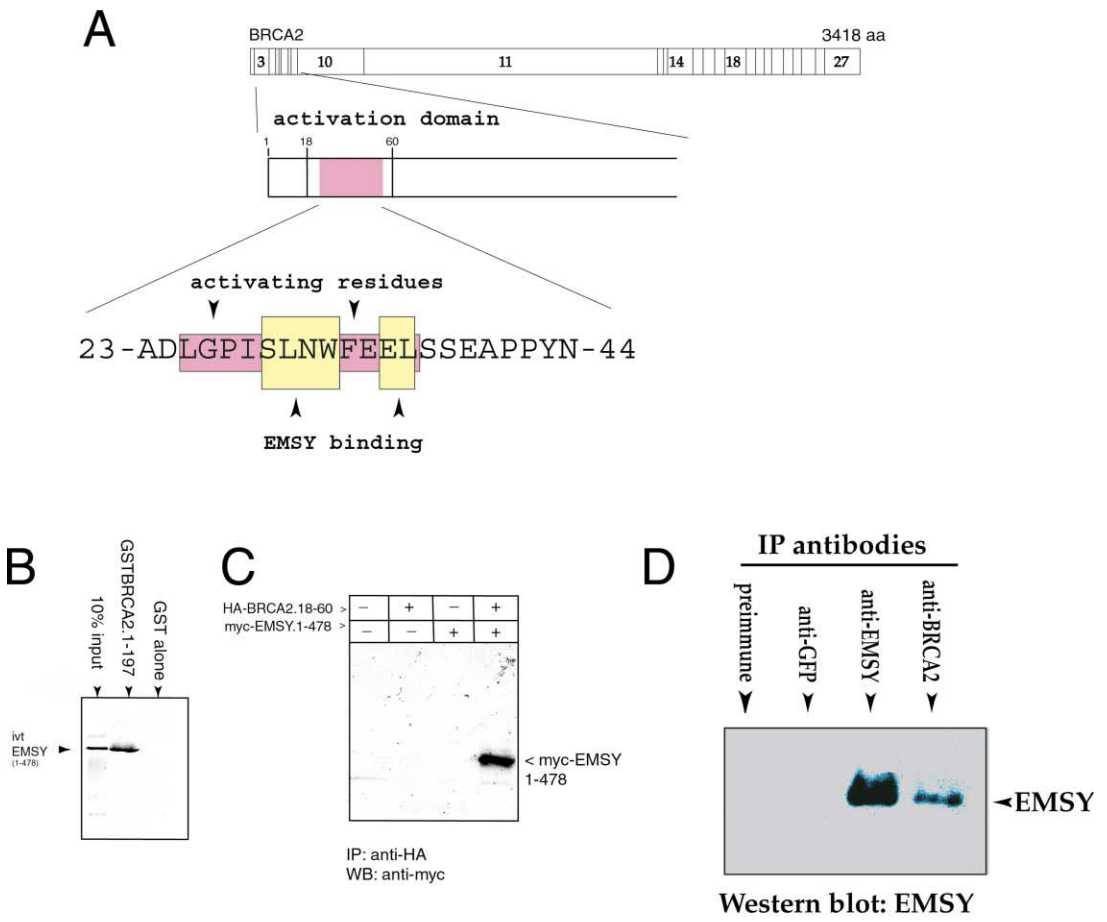
Summary

The *BRCA2* gene is mutated in familial breast and ovarian cancer, and its product is implicated in DNA repair and transcriptional regulation. Here we identify a protein, EMSY, which binds BRCA2 within a region (exon 3) deleted in cancer. EMSY is capable of silencing the activation potential of BRCA2 exon 3, associates with chromatin regulators HP1β and BS69, and localizes to sites of repair following DNA damage. *EMSY* maps to chromosome 11q13.5, a region known to be involved in breast and ovarian cancer. We show that the *EMSY* gene is amplified almost exclusively in sporadic breast cancer (13%) and higher-grade ovarian cancer (17%). In addition, *EMSY* amplification is associated with worse survival, particularly in node-negative breast cancer, suggesting that it may be of prognostic value. The remarkable clinical overlap between sporadic *EMSY* amplification and familial *BRCA2* deletion implicates a BRCA2 pathway in sporadic breast and ovarian cancer.

Introduction

Germline mutations in *BRCA1* and *BRCA2* account for the large majority of autosomal dominant familial breast cancers. However, familial breast cancer is rare, accounting for less than 5% of all cases. The role of the *BRCA1* and *2* genes in sporadic breast cancer remains unclear, since somatic mutations in either gene are uncommon (Gayther et al., 1998). Allele inactivation by promoter hypermethylation occurs at the BRCA1 locus in somatic breast cancer (Dobrovic and Simpfendorfer, 1997), but not at the *BRCA2* locus, even in those tumors with loss of heterozygosity (LOH) at this site (Collins et al., 1997), suggesting that this mechanism of transcriptional repression is unlikely to explain the absence of somatic *BRCA2* mutations in sporadic cancers.

Although there is no sequence similarity between the *BRCA1* and *BRCA2* gene products, they seem to be functionally related, and have been implicated in RAD51-mediated recombinational repair of double-strand DNA breaks, chromatin remodelling, and the regulation of transcription (Venkitaraman, 2002).



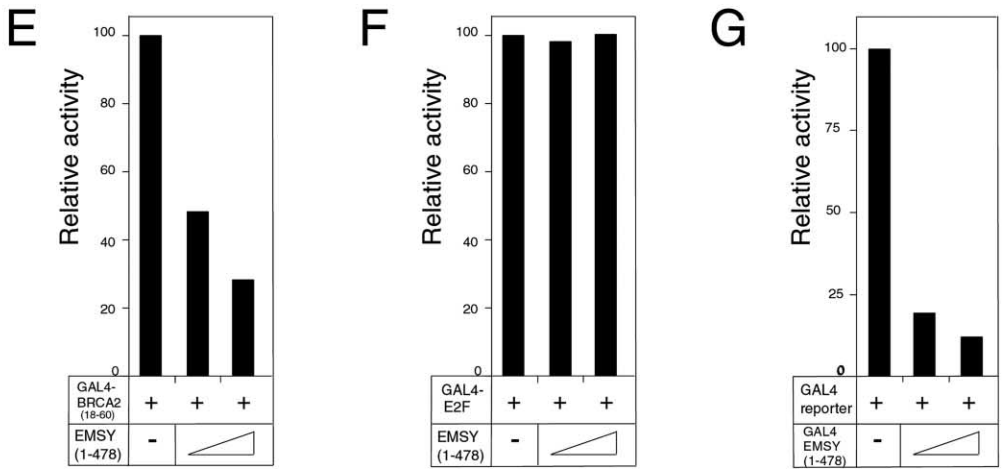


Figure 1. Isolation of *EMSY* and Interaction Studies

(A) The complete exonic structure of the *BRCA2* gene, expanded to show the previously described transcriptionally active region, and then expanded again to show the results of alanine mutagenesis mapping of the BRCA2-EMSY interaction. The yeast two-hybrid bait pBTM116.BRCA2.18–60 was sequentially mutated to establish which residues were important for endogenous transactivation activity in yeast, and which could affect the interaction with EMSY in the yeast two-hybrid experiment. Yellow = mutations which reduce endogenous transactivation capability but do not affect EMSY binding. Pink = mutations which both reduce transactivation and abolish EMSY binding.

(B) Evidence for the interaction between EMSY and BRCA2. EMSY associates with BRCA2 in GST pull-down. EMSY 1–478 was cloned in the PING14 vector and used to in vitro translate (IVT) ^{35}S-methionine-labeledrecombinant protein. Incubation of the radiolabeled EMSY (1–478) resulted in binding to the fusion protein GST-BRCA2 (1–197), but not to the GST control.

Soon after the cloning of *BRCA1* and *BRCA2*, transcriptional activation domains were described in both proteins. One such domain is present in the highly conserved exon 3 of *BRCA2* (Figure 1A) (Milner et al., 1997). The importance of this domain in the tumor suppressor function of *BRCA2* is suggested by the report of a breast and ovarian cancer family whose germline mutation was shown to be an exact in-frame deletion of *BRCA2* exon 3 alone (Nordling et al., 1998), although such a deletion might also have general effects on protein structure and function. The physiological role for a BRCA2 transactivation domain remains uncertain. However, it has recently been shown that BRCA2 can stimulate transcription of androgen receptor-regulated genes in cooperation with histone acetytransferases (Shin and Verma, 2003).

Here we identify a protein, EMSY, which binds BRCA2 and like BRCA2 has functions associated with DNA repair and transcriptional regulation. EMSY has the ability to contact chromatin-regulating proteins that contain the newly defined Royal Family domain. Clinical studies show that the *EMSY* gene is frequently amplified in sporadic cases of breast and ovarian cancer.

Results

Isolation of *EMSY*

We performed yeast two-hybrid screens with small overlapping sequences from the BRCA2 N terminus (Figure 1A). A screen with the region of BRCA2 spanning amino acids 1–46 identified 10 specifically interacting clones, while an overlapping region (amino acids 18–60) yielded 6 clones. All sixteen interactors proved to be fragments of the same cDNA, encoding a protein which we have called EMSY (since the word "SISTER" appears in the first line of the protein sequence, we named this protein after the first author's sister, who is a breast cancer nurse). These results indicate that a region at the amino terminus of EMSY (corresponding to amino acids 1–80 of the predicted EMSY protein) binds within residues 18–46 of BRCA2 exon 3.

We examined the EMSY binding site within the BRCA2 exon 3 activation domain with an alanine scanning mutagenesis approach (Figure 1A). The interaction site was narrowed down to the BRCA2 amino acid region 23–46 with a deletion series (not shown). Each amino acid residue within this critical region of BRCA2 was then mutated to alanine with site-directed mutagenesis, and each of these mutants was tested in the yeast two-hybrid assay for interaction with EMSY. These experiments showed that a short patch of 12 amino acids at the BRCA2 N terminus is required for transcriptional activation. The EMSY-BRCA2 interaction is affected by mutation of six of these residues.

To confirm the interaction between EMSY and BRCA2, we used a bacterially expressed recombinant glutathione-S-transferase (GST) BRCA2 (1–197) fusion protein to pull down the radiolabeled in vitro translated N-terminal region of EMSY (1–478) (Figure 1B). The interaction was also tested with transfection of the respective epitope-tagged mammalian expression constructs into 293T cells, followed by coprecipitation and Western blotting of whole cell extract. EMSY (1–478) interacted specifically with BRCA2 (18–60), corresponding to the exon 3 activation domain (Figure 1C). We then generated two rabbit polyclonal anti-EMSY antibodies to short fragments of recombinant bacterially expressed human GST-EMSY proteins corresponding to EMSY amino acids 130–180 (Abcam ab123) and EMSY amino acids 178–226 (ab2). Both antibodies strongly detected a single band on Western blot of whole cell U2OS extract running slightly slower than the predicted 160 kDa. Fractionation experiments show that EMSY is exclusively nuclear (data not shown). With these characterized EMSY antibodies, we used coimmunoprecipitation to show that native BRCA2 and EMSY proteins associate in vivo in untransfected cells (Figure 1D).

Since we had shown that EMSY binds to residues that are involved in transcriptional activation, we asked if the binding of EMSY might affect this activity. Figures 1E and 1F show that the expression of EMSY 1–478 suppresses the activity of GAL4-BRCA2 18–60, but does not affect the activity of a control fusion GAL4-E2F. The EMSY 1–478 region possesses independent repressive functions, since a GAL4-EMSY 1–478 fusion is able to silence transcription of a high basal level promoter (Figure 1G).

Collectively, these results demonstrate that EMSY is a repressor protein which has the capacity to bind a small epitope within the BRCA2 exon 3-encoded transcriptional activation domain and specifically silence this activity.

EMSY Relocalization to Sites of DNA Damage

Immunofluorescence studies with the anti-EMSY antibody (Abcam ab123) showed that EMSY is nuclear, with

(C) EMSY (1–478) coimmunoprecipitates with BRCA2 (18–60) from transfected cellextracts. Epitope-tagged fragments of EMSY (pcDNA.EMSY1–478.myc) and BRCA2 (pcDNA.HA.BRCA2.18–60) were transfected into 293T cells, immunoprecipitated with an anti-HA (haemagglutin) antibody (Boehringer-Mannheim), and Western blotted with an anti-myc antibody (Roche).

(D) Endogenous coimmunoprecipitation of EMSY and BRCA2 from unmanipulated asynchronously dividing HeLa cells. Immunoprecipitation with an unrelated antibody (anti-green fluorescent protein, Abcam, Cambridge, UK), preimmune serum (from a different rabbit), or anti-BRCA2 antibody (Ab-2, Oncogene Research Products, San Diego CA) was followed by Western blotting with the affinity-purified anti-EMSY antibody (Abcam ab123). Neither of the unrelated antibodies could coprecipitate EMSY, whereas an EMSY signal was detected in the immunoprecipitate obtained with the anti-BRCA2 antibody. The anti-EMSY antibody was also used to immunoprecipitate endogenous EMSY protein.

(E) EMSY N terminus specifically represses BRCA2 transcriptional activity. Cotransfection of increasing amounts of pcDNA.EMSY (1–478) repressed the ability of the pCMV.GAL4. BRCA2 18–60 region to activate a reporter construct in a CAT assay, and this effect was shown to be specific in (F), since no effect was seen on the E2F activation domain.

(G) EMSY was also shown to have endogenous repressive functions. Amino acids 1–478 of EMSY were cloned as a fusion protein to GAL4 to allow recruitment to a GAL4 coupled high basal TK promoter driving expression of the CAT gene. Transient transfection of increasing amounts (0.5 and 5 μg) of the pcDNA3.GAL4-EMSY1–478 construct caused increasing transcriptional repression. The activity of the reporter in the presence of the empty pcDNA-GAL4 vector was normalized to 100.

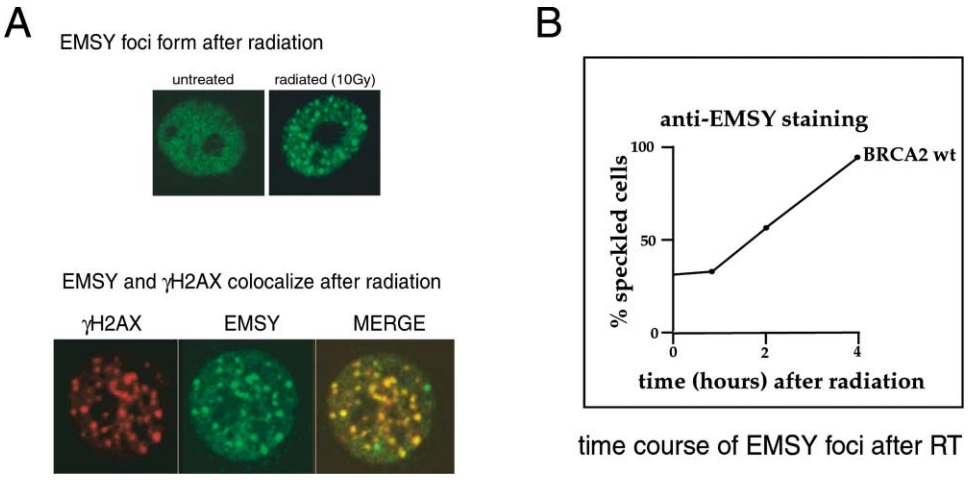


Figure 2. EMSY Relocalizes in Response to DNA Damage

(A) EMSY responds to DNA damage. EMSY antibody immunofluorescence staining of immortalized mouse wild-type embryonic fibroblasts before and after treatment with ionizing radiation (10 Gray at 1.8 Gy/minute, 250 kVp). EMSY assembles into nuclear speckles over several hours (see graph in [B] for time course). Costaining with a monoclonal mouse antibody to γ-H2AX (Roche) reveals that EMSY relocalizes to DNA damage sites.

a granular appearance (Figure 2A). We asked whether EMSY can relocalize to sites of DNA repair after DNA damage, as is the case with BRCA2. Figure 2 shows that treatment with ionizing radiation causes a substantial increase in the number of cells with EMSY-containing nuclear speckles. The time course of DNA damage-induced speckling in the wild-type cells is shown (Figure 2B). To establish whether EMSY relocalizes to DNA repair sites, we performed coimmunofluorescence experiments with an anti-phospho-histone 2AX (γ-H2AX) antibody. H2AX is phosphorylated rapidly in response to DNA damage, forming γ-H2AX speckles that mark the site of DNA repair (Paull et al., 2000). EMSY colocalizes almost completely with γ-H2AX foci after radiation (Figure 2A), suggesting that EMSY has some role in DNA repair. These immunofluorescence experiments were performed in a spontaneously immortalized wild-type murine embryonic fibroblast cell line derived in the Venkitaraman laboratory. To date, demonstration of damage-inducible EMSY speckling in human cell lines has been unsuccessful.

The EMSY ENT Domain and Plant Agenet Domains

The complete *EMSY* cDNA sequence was cloned. The open reading frame encodes a predicted protein of 1322 amino acids. The sequence is entirely novel. There are no other EMSY-like proteins in the human genome, and a single *EMSY* gene can be found in frogs and fish (Figure 3A). However, we identified an 80 amino acid evolutionarily conserved domain at the amino terminus of EMSY, which we have called the *E*MSY *N-T*erminal domain (ENT domain; Pfam number PF03735) (Figure 3A). In *Arabidopsis*, there are 9 different proteins which possess an ENT domain. Examination of these genes shows that they all contain a new Tudor-like domain, which we call the Agenet domain (Pfam number PF05641)

EMSY and the Royal Family Connection

As we have recently reported (Maurer-Stroh et al., 2003), the Agenet domain forms part of a superfamily we term the Royal Family of protein domains. Royal Family members include Chromo, PWWP, MBT, Tudor, and Agenet (Figure 3C). These domains are similar in both primary sequence (Maurer-Stroh et al., 2003) and tertiary structure. Representatives of Tudor, PWWP, Chromo, and MBT have recently been solved at the structural level. Our examination of these structures shows that they exhibit a common fold. Within this common fold, there is conservation of aromatic residues found within the HP1 chromodomain, which are known to recognize methylated lysines within histones (Nielsen et al., 2002; Jacobs and Khorasanizadeh, 2002). These findings raise the possibility that some members of the Tudor, PWWP, and MBT family may also be recognizing modified histones. This would be compatible with the fact that these domains are often found in chromatin regulating proteins.

An HP1 and BS69 Binding Platform within EMSY

We performed yeast two-hybrid screens with the first 140 amino acids of EMSY (which includes the ENT domain) in an attempt to gain some insight into the function of this protein. We noticed that over 80% of the 112 interacting clones encoded proteins that contained Royal Family domains. These include HP1 (which has a chromodomain), BS69 (PWWP domain), and a novel MBT domain-containing protein. This suggests that EMSY may coordinate the binding of proteins containing the Royal Family domain. To test this, we chose two interactors, HP1 and BS69, for further study.

All three isoforms of HP1 (α, β, γ) were detected in the preliminary yeast two-hybrid screens. The interaction was also seen between recombinant bacterial expressed EMSY and each of the HP1 isoforms in a GST pulldown (Figure 4A). However, in an endogenous coimmunoprecipitation, we were only able to confirm that EMSY and HP1β interact in vivo (Figure 4B), whereas we could not demonstrate an interaction between EMSY and HP1α or HP1γ (data not shown).

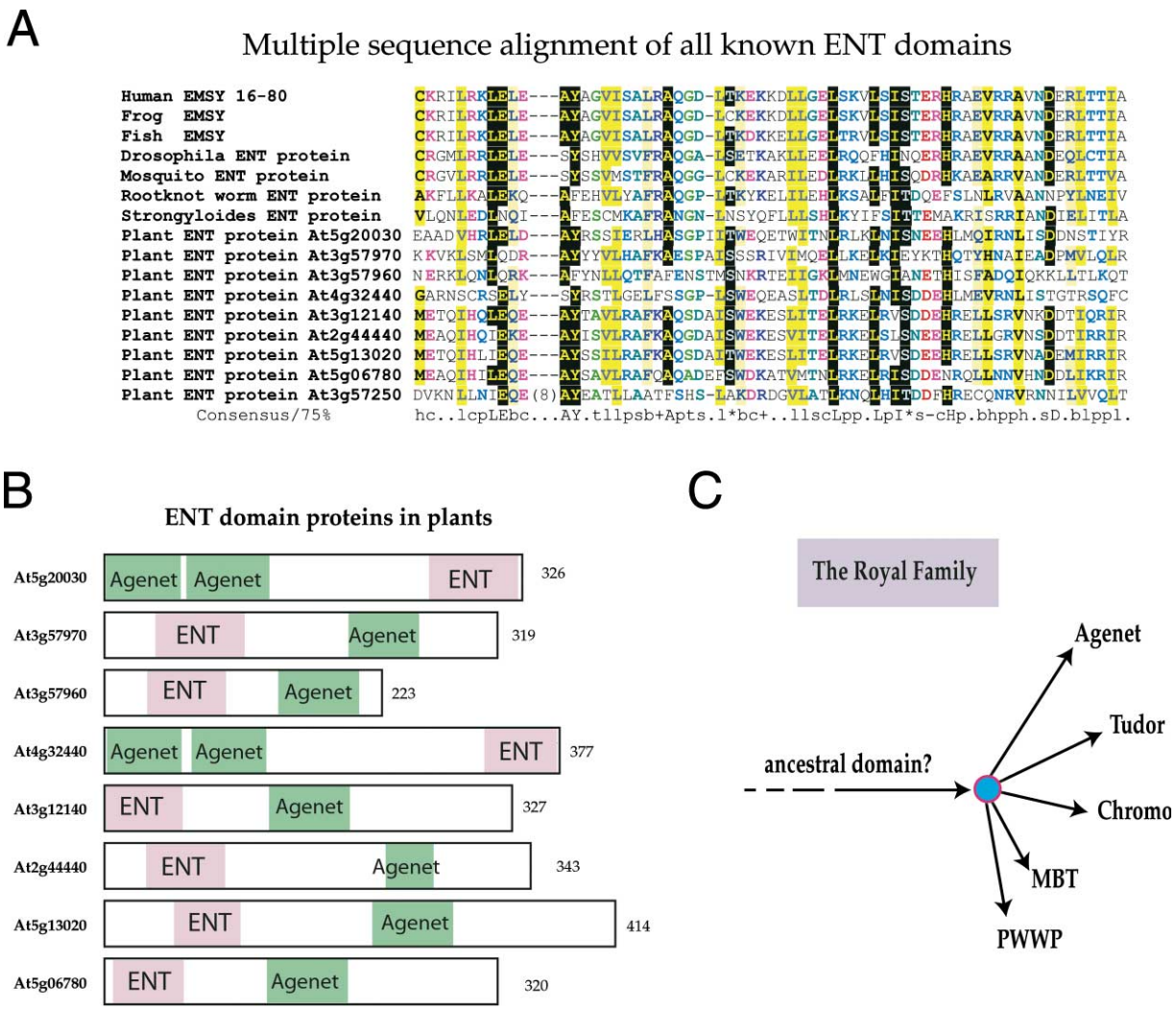


Figure 3. *EMSY* Bioinformatics

(A) Multiple sequence alignment of ENT domains presented using CHROMA and a 75% consensus. Species abbreviations (with accession numbers in brackets): mosquito, *Aedes aegypti* (agCP8691); plant, *Arabidopsis thaliana* (accession numbers given in figure); human, *Homo sapiens* (CAD22881); *Drosophila*, *Drosophila melanogaster* (CG15356); zebrafish, *Danio rerio* (EST 14163341); rootknot worm, *Meloidogyne javanica* (EST 9829797); strongyloides, *Strongyloides ratti* (EST 14288209); and frog, *Silurana tropicalis* (EST 17663399).

(B) Schematic stick figure representation of the eight *Arabidopsis* ENT-Agenet proteins. The ENT domain is shown in pink, the Agenet domain in green.

(C) Diagram of the Royal Family of domains, illustrating the possible evolutionary divergence from a single ancestral domain. Within the Royal Family, there are five domains that contain sequence similarity. These are PWWP (Stec et al., 2000), MBT (Bornemann et al., 1998), Tudor (Yang et al., 2003), Agenet (Maurer-Stroh et al., 2003), and the chromodomain (Jones et al., 2000). Each Royal Family member has been implicated in chromatin remodeling functions. The PWWP domain was named after its central core "Pro-Trp-Trp-Pro," and is usually found with other chromatin-associated domains such as the SET, bromo-, and PHD domains; they may be involved in the heterochromatin localization of the DNA methyltransferase, Dnmt3a (Bachman et al., 2001) The Tudor domain was first identified in the *Drosophila* Tudor protein (the *tudor* phenotype includes male sterility, which is often thought to be characteristic of the Tudor kings of England, hence the name), and is also found in a lysine-specific histone methyltransferase that contributes to HP1-mediated silencing (Schultz et al., 2002). The chromodomain was first identified as a homology between heterochromatin protein 1 (HP1) and polycomb protein (Pc), and is crucial for targeting to methylated histone tails (Bannister et al., 2001; Lachner et al., 2001); the MBT (malignant brain tumor) domain was identified during a screen for tumor suppressor genes in *Drosophila* and is confined to members of the polycomb group of proteins which are involved in the maintenance of heterochromatic states (Bornemann et al., 1998).

Using deletions of both HP1 and EMSY, we found that the HP1 chromoshadow domain interacts with EMSY (Figure 4C). Deletion mapping narrowed down the interaction site within EMSY to a region between amino acids 80 and 140 (data not shown), and examination of this region revealed a candidate HP1 binding consensus site adjacent to the ENT domain (Figure 4D). Combined mutagenesis of two residues within this region, V100A and L102A, abolished the binding of GST-EMSY to his-HP1, confirming that this region is central to the EMSY-HP1 interaction (Figures 4E and 4F).

To analyze the EMSY-BS69 interaction in vivo, we had to transfect epitope-tagged BS69 into cells, since antibodies against BS69 are not available. Figure 5A

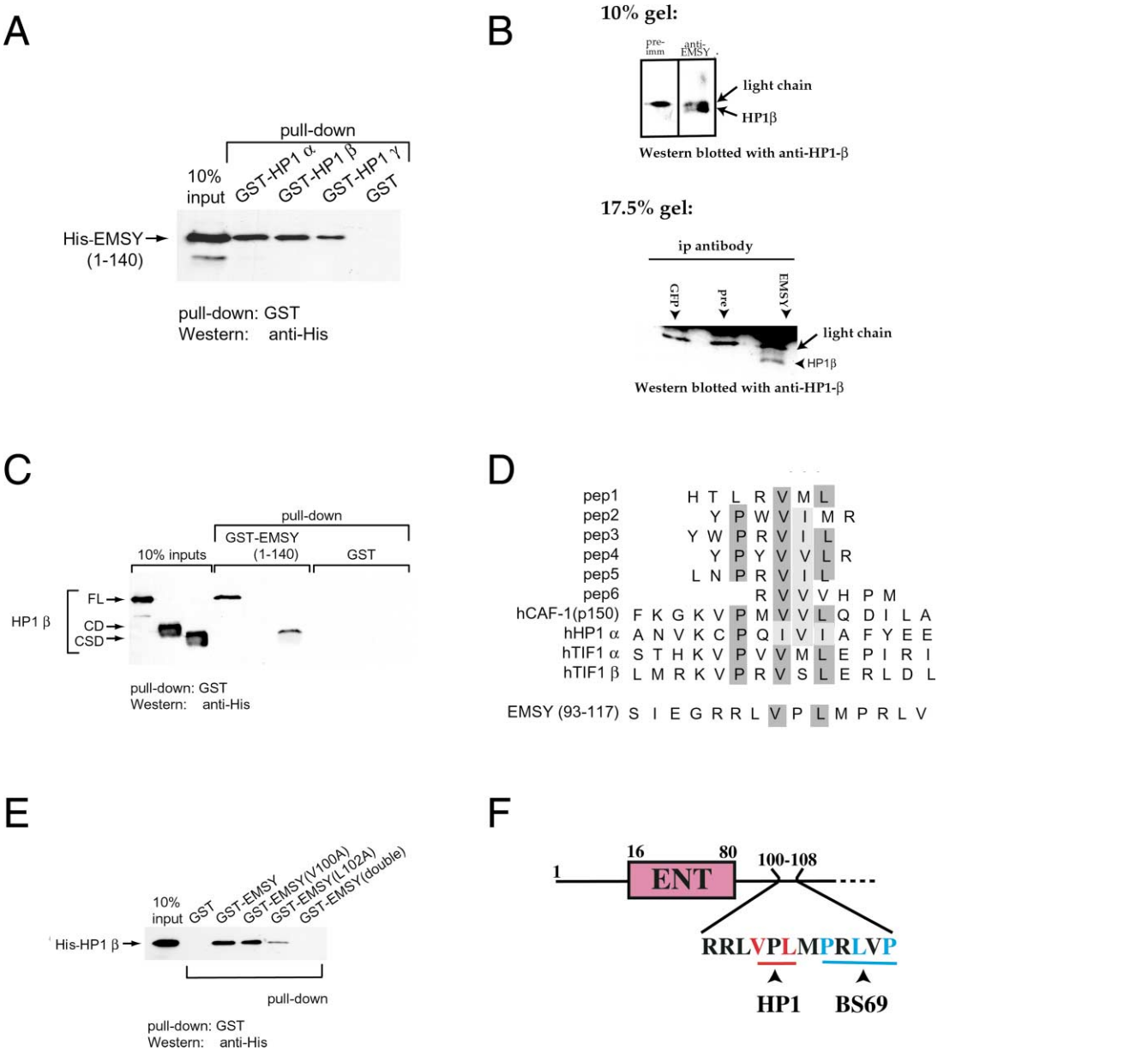


Figure 4. Evidence for HP1β Binding adjacent to the EMSY ENT Domain

(A) GST pull-downs of bacterially expressed recombinant His-EMSY (1–140) by each of the three isoforms of GST-HP1.

(B) Coimmunoprecipitation of EMSY and HP1β. HP1β (molecular weight 28 kDa) is predicted to run just below the light chain from the immunoprecipitating rabbit antibody (31 kDa), and a band doublet corresponding to HP1β and the light chain can be seen at this position after IP with the EMSY antibody in a 10% SDS-PAGE gel, but not with the preimmune sera (where a single band corresponding to the light chain can be seen). This experiment was repeated with a 17.5% gel (lower panel) to achieve better separation of the immunoglobulin light chain and HP1β; neither the preimmune sera or an unrelated antibody (anti-GFP) was able to immunopurify HP1β, but the anti-EMSY antibody does, confirming the interaction in vivo. The HP1β antibody was purchased from Upstate.

(C) GST fusion protein pull-down of his-tagged HP1β deletions reveals that the EMSY interaction is mediated through the HP1β chromoshadow domain. An anti-his tag antibody was used to detect the pull-down proteins (Quiagen). CSD, chromoshadow domain; CD, chromodomain; FL, full-length.

(D) Deletion mapping of the EMSY amino terminus narrowed down the interaction site to a region between amino acids 80 and 140 (data not shown). Sequence comparison of HP1 binding sites suggests a possible HP1 binding site in EMSY. Pep = peptides isolated from a random peptide display phage library; hCAF1 = human chromatin assembly factor 1; hTIF1 = human transcription intermediate factor 1 (Smothers and Henikoff, 2000).

(E) Alanine mutagenesis of the GST-EMSY fusion protein demonstrates that EMSY residues 100 and 102 are indeed critical for the HP1β interaction.

(F) Summary diagram illustrating a binding platform for two Royal Family domain-containing proteins, HP1 and BS69, lying adjacent to the ENT domain in EMSY.

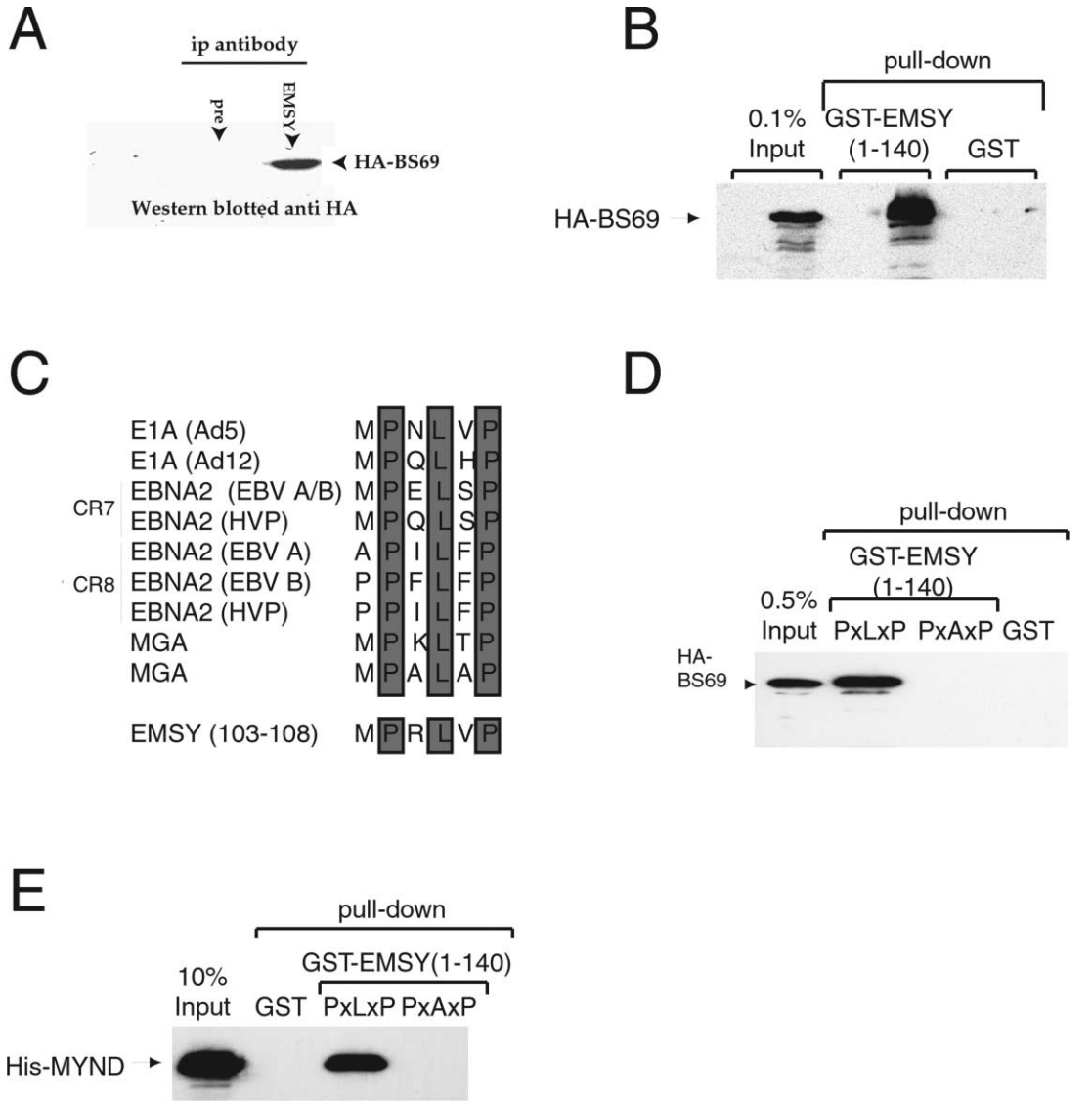


Figure 5. Evidence for BS69 Binding adjacent to the EMSY ENT Domain

(A) Coimmunoprecipitation of endogenous EMSY and HA-BS69 from HA-BS69 transfected HeLa cells, using the previously characterized anti-EMSY antibody (Abcam ab123) for IP and a mouse monoclonal anti-HA antibody (Boehringer-Mannheim) for Western detection.

(B) Pull-down of HA-BS69 from transfected 293T whole cell extracts with EMSY (1–140)-GST fusion protein.

(C) Evidence for a BS69 core binding motif in EMSY; the sequence is compared to core motifs in known BS69 interactors (E1A = adenoviral E1A oncoprotein; EBNA = Epstein Barr nuclear antigen; CR7 and CR8 refer to conserved regions of the EBNA protein; MGA = myc related cellular transcription factor) (Ansieau and Leutz, 2002).

(D) Mutation analysis confirms that the PxLxP motif at position 103–108 of EMSY is critical for the BS69-EMSY interaction.

(E) Deletion analysis shows that the EMSY interaction is mediated through the MYND domain of BS69, and that mutation of the BS69 core binding site abolishes the BS69 interaction.

shows that the immunoprecipitation of endogenous EMSY coprecipitates an HA-tagged BS69 protein. The N terminus of EMSY (residues 1–140) is sufficient to bind BS69, since a GST-EMSY (1–140) fusion protein can purify transiently transfected HA-tagged BS69 (Figure 5B).

A binding motif for BS69 (PXLXP) has been described for a number of BS69 partners (Figure 5C). This motif was found at amino acids 104–108 of EMSY, which falls within the EMSY binding domain for BS69 (EMSY1–140) (Figure 4F). Mutation of the PXLXP motif to PXAXP completely abolishes the binding of EMSY to HA-tagged full length BS69 (Figure 5D), and a deletion analysis revealed that the MYND domain of BS69 is sufficient for this interaction (Figure 5E).

Together these results indicate that the HP1β chromoshadow domain interacts with a VPL motif in EMSY, whereas the MYND domain of BS69 interacts with a PXLXP motif in EMSY. The binding sites for HP1β and BS69 with EMSY abut each other, and are found directly adjacent to the ENT domain of EMSY. This demonstrates that EMSY has the capacity to contact directly at least

two proteins which contain a Royal Family domain. Since this domain is found in proteins with a chromatin connection, we assume that EMSY functions, at least partly, in the regulation of chromatin.

EMSY's Status in Breast Cancers

The *EMSY* gene locus maps to chromosome 11q13.4-5 by FISH and radiation hybrid mapping. This gene-dense region has been extensively implicated as a site of amplification in sporadic breast cancer (Bekri et al., 1997; Ormandy et al., 2003). A single gene cannot be responsible for all the amplifications at 11q13, because at least four distinct cores of amplification have been recognized (Figure 6; Bekri et al., 1997; Ormandy et al., 2003). Recently, the physical map of this region has been completed (Chromosome 11, build 33, www.ncbi.nlm.nih.gov/genome/guide/human), and it reveals that three cores are clustered within a 1.4 Mb interval. *Cyclin D1* and *EMS1* are strong candidates for driving amplification at two cores within this cluster, whereas the candidate gene for the third (D11S97) is unknown.

A further core is defined by the *GARP* gene (Bekri et al., 1997). The *GARP* amplicon lies 6 Mb telomeric to the other cores, a physical separation which may explain why this locus has the highest independent amplification frequency of all the amplicons on 11q13 (Ormandy et al., 2003). Strikingly, *EMSY* lies immediately adjacent to the *GARP* gene, within 0.1 Mb. Therefore, it must also be contained in this well-described breast cancer amplicon (Figure 6A).

To investigate whether *EMSY* is amplified in breast cancer, we used FISH (fluorescence in situ hybridization) to examine metaphase spreads from 28 breast cancer cell lines. We also did some intraoperative touch preparations from five newly diagnosed sporadic breast primary cancers, and found one with gross *EMSY* amplification (Figure 6B). In the breast cancer cell lines, we found that five of the 28 (18%) had *EMSY* amplification (Table 1, middle column). In one of these cell lines, massive amplification of *EMSY* was independent of *Cyclin D1* (Figure 6B, left panel).

Next, we asked whether this genomic amplification of the *EMSY* gene results in increased expression of *EMSY*. Quantitative RT-PCR confirmed that cell lines with *EMSY* genomic amplification also had the highest levels of *EMSY* expression (Table 1, right column).

In order to confirm that the *EMSY* amplicon could be independently amplified, we performed array comparative genomic hybridization (CGH) on key genes across the locus: *CCND1, EMS1, GARP, and PAK1*, and FISH for *EMSY* (Supplemental Table S1 at http://www.cell.com/cgi/content/full/115/5/523/DC1; see Figure 6A for a description of gene order relative to amplicon cores). We found that in the SUM52 cell line, the amplification interval involved only *EMSY* and *GARP*, and not genes from the more centromeric amplification cores, or *PAK1* which lies immediately telomeric. This confirms previous reports of independent amplification at this locus (Bekri et al., 1997).

When the expression of *EMSY* and *GARP* was compared in the SUM52 cell line, *EMSY* was found to be overexpressed by 6-fold, whereas *GARP* was barely detectable. These results confirm that *EMSY* can be independently overexpressed in breast cancer. Furthermore, the very low levels of *GARP* expression in the five cell lines with *EMSY/GARP* genomic amplification suggest that *EMSY*, not *GARP*, is the critical gene in this amplicon.

Next, a previously described collection of sporadic primary breast cancers was examined with Southern blotting (Courjal et al., 1997). We found high-level *EMSY* amplification in 65 (7.4%) of 875 of these sporadic breast cancers. Again, in order to establish whether *EMSY* amplification could occur independently of other genes in this region, probes to the *GARP*, *Cyclin D1*, and *EMS1* genes were also used in Southern blotting. *EMSY* amplification occurs independently of *Cyclin D1* in 40% of cases. These results show that the *EMSY/GARP* gene region is amplified in breast cancer, and that it is not simply being amplified as a passenger to *Cyclin D1*, a well-characterized breast cancer oncogene which lies 6 Mb centromeric to *EMSY*.

EMSY Amplification Is Associated with a Poor Prognosis in Human Breast Cancer

We then asked whether *EMSY* amplification is associated with clinical outcome by using FISH in tissue microarrays (Kononen et al., 1998) of primary sporadic tumors sent to Vancouver Hospital during 1976 to 1990. Clinical data and outcome, including all breast cancer recurrences and deaths, were available for all patients. Amplification ratios for *EMSY* and *Cyclin D1* were obtained by normalizing gene copy number against centromere 11 copy number. FISH for *EMSY* gene copy number was successful in 551 of 928 cases. There was no difference in survival between cases with and without *EMSY* data ($p = 0.79$).

This more sensitive FISH-based analysis showed *EMSY* amplification (copy number ratio greater than 1.5) in 70 of the 551 breast cancers (13%). In 49 of these there were more than 5 copies of *EMSY* per cell. *Cyclin D1* amplification was seen in 78 cancers (14%), and was present in only a subset of cancers with *EMSY* amplification (42/70 [60%]), which again suggests that both genes are independent oncogenic targets; this proportion of *Cyclin D1*-independent *EMSY* amplification agrees exactly with the data obtained by Southern blotting on a different set of sporadic breast cancers described above.

Examination of records from the cases described above indicates that the median disease-specific survival was 6.4 years for node-negative cases with *EMSY* amplification (copy number ratio > 1.5), versus 14 years for node-negative cases without *EMSY* amplification. Overall, patients with *EMSY* amplification had a worse disease-specific survival compared to patients without *EMSY* amplification ($p = 0.002$) (Figure 7A). Remarkably, this effect was confined to node-negative breast cancer patients ($p = 0.0004$) compared to node-positive cases ($p = 0.14$) (Figures 7B and 7C). A similar effect is seen on overall survival (data not shown). The effect of *EMSY* amplification on survival in patients with node-negative disease remained significant in a multivariate analysis presented below the graph in Figure 7B. We were also able to show that the survival of patients with *EMSY* amplification was not affected by the presence of con-

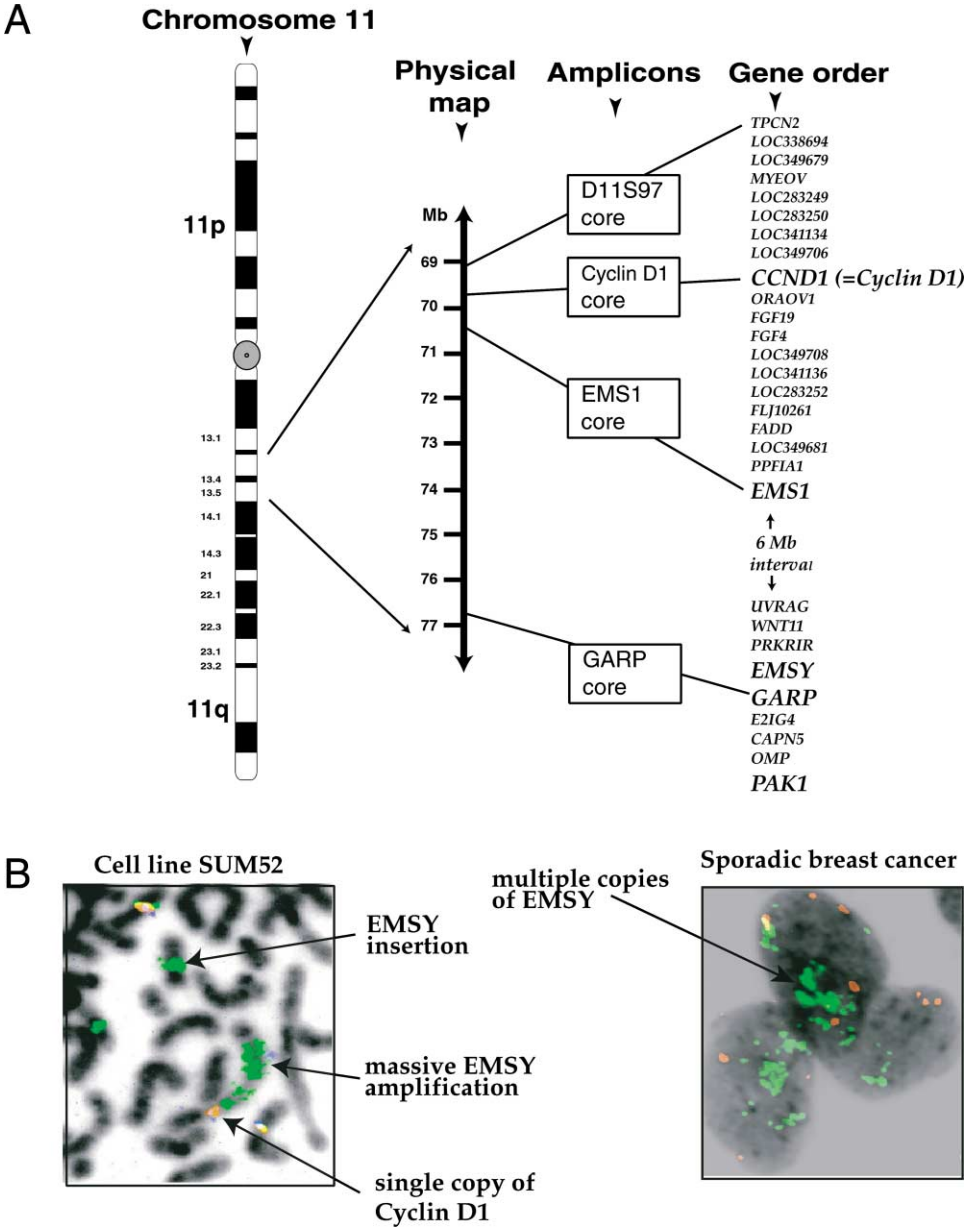


Figure 6. *EMSY* Amplification in Breast Cancer

(A) Chromosomal mapping of *EMSY*. The recently available physical map of the chromosome 11q13.4-5 breast cancer amplified region is displayed next to the cytogenetic map of chromosome 11. The gene order within this region is also listed. Three of the four known breast amplicon cores (*D11S97*, *CCND1/CyclinD1*, and *EMS1*) are clustered within a 1.4 Mb interval, which contains only 20 genes. The fourth core (*GARP*) lies 6 MB telomeric to this three-core cluster. Within the *GARP* core, *EMSY* can be seen lying immediately centromeric to *GARP* in the gene order (and within 0.1 Mb on physical mapping). The genes discussed in the text are shown in larger type within the amplicon.

(B) Left panel shows high-level genomic amplification of *EMSY* by FISH in the SUM52 breast cancer line. The *EMSY* probe is labeled green, and a large homogenously staining region (HSR) is seen together with *EMSY* genomic insertions in other chromosomal regions, a feature of high-level genomic amplification. The simultaneous hybridization with the *Cyclin D1* (red) shows that *EMSY* amplification is independent of *Cyclin D1* in this cell line. Right panel shows high level *EMSY* amplification in a primary breast cancer (touch prep of surgical specimen) with *EMSY* (green) and centromere 11 probe (red).

current *Cyclin D1* amplification (Figure 7B). We conclude that *EMSY* amplification identifies a subset of breast cancers associated with a poor prognosis.

EMSY's Status in Ovarian Cancer

EMSY was assessed in ovarian cancers with tissue arrays. Amplification was seen in 17% of high-grade carcinomas (62/360) and 0% of low-grade or borderline ovarian cancers (0/110) ($p = 0.001$). In contrast to breast cancer, we found that *Cyclin D1* amplification (11%) was less common than *EMSY* amplification (17%), and *EMSY* was less likely to be coamplified with *Cyclin D1* (40% of *EMSY* amplification cases had *Cyclin D1* coamplification). Thus, *EMSY* rather than *Cyclin D1* appears to be

Table 1. Genomic Amplification Is Accompanied by *EMSY* Overexpression

Cell Line	Copy Number Ratio	Ranked Expression
SUM44	>5	1
SUM52	>5	2
MDA-MB-175	>5	3/4 (tied)
SUM190	>5	
MATU	1.7	5
VP229	1.0	6
MDA-MB-453	1.4	7
OCUB-M	0.9	8
VP267	1.5	9
MDA-MB-361	0.9	10
MDA-MB-436	0.9	11
MDA-MB-468	1.5	12
Kpl-1	1.2	13/14 (tied)
SUM102	1.3	
SKBR3	1.0	15
MT-3	1.0	16/17 (tied)
MDA-MB-231	1.2	
ZR-75-1	0.6	18
MDA-MB-157	1.0	19
BT549	1.0	20
SUM185	1.5	21
MCF-7	1.5	22
CAL51	1.0	23
SUM149	1.1	24
MDA-MB-134	>5	not performed
MDA-MB-435S	1.1	not performed
DU4475	1.0	not performed
MCF-7	1.5	not performed

The name of each of the 28 breast cancer cell lines analyzed in this study is listed in the left hand column. *EMSY* genomic copy number is recorded in the middle column. In the right hand column, *EMSY* expression levels are ranked, with the highest expressing cells at the top (real time PCR estimation of relative mRNA levels). *EMSY* expression was assessed using ABI TAQMAN real-time PCR in a subset of 24 of the cell lines in which *EMSY* genomic copy number had been determined. Two independent probes for *EMSY* were designed and gave similar results. The ranking shows that cell lines previously found to have genomic amplification of *EMSY* by FISH also ranked highest in *EMSY* mRNA levels. The expression study was performed blind on numbered samples and *EMSY* expression levels were ranked before the number codes were broken to identify the cell lines.

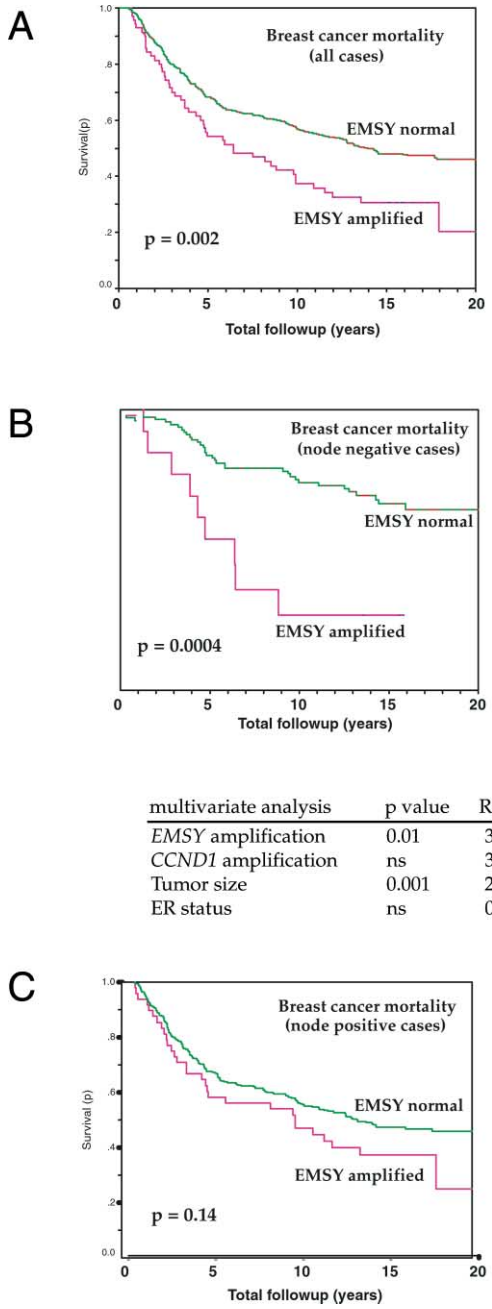
a more frequent gene amplification event involving 11q13 in ovarian cancers.

EMSY's Status in Non-Breast/-Ovarian Cancers

We also examined *EMSY* amplification in non-breast/-ovarian cancers using multitumor arrays. *EMSY* amplification was seen in 1/5 hepatocellular carcinomas, 1/7 melanomas, and 1/57 endometrial carcinomas, but was not identified in more than 200 other cancers from a wide range of primary sites, including 21 sarcomas and 34 colorectal carcinomas.

Discussion

We have cloned *EMSY*, a human breast and ovarian cancer gene. The EMSY protein interacts precisely with a highly conserved transactivating region at the N terminus of the breast cancer protein BRCA2, and has endogenous transcriptional repressor activity when recruited to a high basal promoter. EMSY is exclusively nuclear



multivariate analysis	p value	RR
EMSY amplification	0.01	3.1
CCND1 amplification	ns	3.0
Tumor size	0.001	2.3
ER status	ns	0.63

Figure 7. *EMSY* Amplification Is Associated with a Worse Outcome in Breast Cancer

Kaplan-Meier curves comparing breast cancer mortality between women with *EMSY* normal and *EMSY* amplification breast cancer, showing correlation between *EMSY* amplification and breast cancer mortality in all 551 cases (A), in node-negative cases (B) and in node-positive cases (C). The table below (B) summarizes the results of a Cox proportional hazard regression multivariate analysis to determine the effects of tumor size (≤2 cm versus >2 cm), *EMSY* amplification (amplification ratio < 1.5 versus amplification ration ≥ 1.5), *Cyclin D1* amplification (amplification ration < 1.5 versus amplification ration ≥ 1.5), and estrogen receptor (ER) status (<10% versus ≥10% nuclear immunoreactivity) status on survival in node negative patients.

and relocalizes after DNA damage to γ-H2AX speckles. H2AX is essential for the efficient recognition and repair of DNA double-strand breaks (DSBs), and becomes rapidly phosphorylated to γ-H2AX at the site of each nascent DSB, resulting in characteristic nuclear foci (Pilch et al., 2003). BRCA2 is also known to relocalize to DSB repair sites after DNA damage. These results suggest that the functions of *EMSY* and *BRCA2* (two breast cancer-specific genes) are connected.

EMSY Organizes Chromatin Regulators Containing the Royal Family Domain

Although EMSY is a large protein, its primary amino acid sequence is uninformative except for a novel evolutionary conserved domain we have called ENT. In plants, proteins containing the ENT domain invariably contain an Agenet domain. This domain is likely to have a role in chromatin regulation. In contrast, EMSY does not possess a Royal Family domain within its sequence but does associate with proteins containing such domains. This suggests that EMSY's function may be connected to chromatin regulation.

Role of EMSY in Transcription and DNA Repair

The ability of EMSY to assemble chromatin-remodeling proteins adds weight to the emerging role for both the BRCA proteins in chromatin remodeling. A large BRCA2 complex has recently been described which includes proteins strongly reminiscent of the SWI/SNF group and which has the capacity to deacetylate histones (Hakimi et al., 2002); it has also been shown that BRCA1-SWI/SNF is the predominant BRCA1-containing complex in HeLa nuclear extract (Bochar et al., 2000) and that targeting of BRCA1 to chromosomal regions in the mammalian genome results in large-scale chromatin decondensation (Ye et al., 2001). It is not yet clear exactly how chromatin-remodeling functions relate to the known DNA repair activity of the BRCA proteins, although two scenarios have been envisaged: either they are required for the activation of DNA damage response genes or for allowing the DNA repair machinery access to its substrate. Repair genes such as GADD45 are known to be major transcriptional targets for BRCA1, which provides some support for the first scenario (Harkin et al., 1999), while reports of rapid changes in chromatin structure and accessibility after DNA damage provide some support for the second (Downs et al., 2000). Another possibility is that aberrant chromatin remodeling might itself lead to genomic instability (Myung et al., 2003).

We have suggested that the independent activation domain of BRCA2 within exon 3 might have some role in transcription (Milner et al., 1997). The identification of the repressor protein EMSY, which binds and silences this domain, is consistent with such a function. However, the distinction between transcription and repair at the chromatin level now seems outmoded, as it has recently become evident that changes in chromatin occur after DNA damage (Fyodorov and Kadonaga, 2001), and these changes are often indistinguishable from the changes that precede transcription. It is therefore possible that the activation/repression function of BRCA2/EMSY may reflect a role in DNA repair just as much as in transcriptional regulation.

EMSY's Involvement in Breast and Ovarian Cancer

We find that *EMSY* amplification is almost never seen in a wide range of non-breast/-ovarian cancers. Furthermore, in ovarian cancer, *EMSY* amplification is confined to grade 2 or 3 subtypes, not grade 1 or borderline (low malignant potential) tumors. This spectrum of pathology for *EMSY* amplification exactly matches that seen in familial ovarian cancer from *BRCA2* mutation carriers (Boyd et al., 2000; Risch et al., 2001). These findings are important because they show that an abnormality in a BRCA2 interacting protein might incriminate *BRCA2* itself in sporadic cancers. This is an important goal, because despite strenuous worldwide efforts since *BRCA2* was cloned over 8 years ago, it has not yet been possible to show how (or if) *BRCA2* is involved in sporadic breast cancers. *BRCA2* is rarely (if ever) mutated in these cancers. The fact that *EMSY* amplification and *BRCA2* deletion display the same pathology in sporadic and familial cancers suggests that overexpression of *EMSY* and deletion of *BRCA2* have a similar effect on a given pathway. We can only speculate as to which function of these two proteins is impinging on the cancer-predisposing pathway. However, at one level of analysis presented here, regarding their transcriptional activation functions, it is possible to imagine that deletion of the BRCA2 activation domain in exon 3 may have the same effect as overexpressing *EMSY*, a repressor of this *BRCA2* activity. Certainly the physical, functional, and clinical relationships between *EMSY* and *BRCA2* presented here strongly suggest that the pathways they regulate are intimately linked.

Experimental Procedures

Yeast Two-Hybrid Screen

The yeast plasmids pVP16 and pBTM116 and mouse embryo (day 10.5) cDNA library were a gift from Dr. Stanley Hollenberg (Oregon Health Sciences University, Portland, OR), and the screens were performed with 50 mmol 3AT (3-amino-1,2,4-triazole) to suppress the endogenous activity of the cloned fragments from *BRCA2* exon 3.

In Vitro Pull-Downs and Alanine Scanning

The pGEX2TK and PET30 plasmids were used in standard protocols for synthesis of bacterially expressed proteins, and the TNT T7 Coupled Reticulocyte Lysate System (Promega, WI) was used to synthesize radiolabeled protein for pull-downs. Alanine mutations were introduced with the Quickchange method (Stratagene).

Tissue Culture and Transactivation Assays

Cell lines were maintained in DMEM supplemented with 10% fetal calf serum and grown at 37°C in 5% CO_2. The pCDNA3 GAL4 (1–147) vector was used for CAT assays; activity was normalized to 100% for the cells transfected with the control vectors; for transactivation assays, the reporter construct contained five tandem repeats of the Gal4 binding site upstream from the CAT gene; for repression assays, 0.5 microgram of high basal reporter pGal4-TK-CAT was used.

Coimmmuoprecipitations

Subconfluent cells were lysed in NETN buffer (20 mM Tris-Cl [pH 8.0], 100 mM NaCl, 1 mM EDTA, 0.5% NP-40) with protease inhibitors (Boehringer), and supernatant from the lysate was precleared with beads and immunoprecipitated overnight with antibody. The immunoprecipitate was collected on protein A/G sepharose beads,

washed four times in NETN buffer, resuspended in denaturing loading buffer, and run in a 10% SDS-PAGE gel before transferring to nitrocellulose membrane and Western blotting.

Quantitative RT-PCR

Probe 1, CAT CAA CAC AGA CAA CAA ACA CAA CAA CAC AGA AG, Probe 2 CTT TTG CAT TCA TCC CTG CTG AGA TCC. Control probes used were PBGD (porphobilinogen deaminase) TGT CCC TGC CCA GCA TGA AGA TGG and TBP (TATA box binding protein) TAG ACT TCA AGA TCC AGA ACA TGG TGG GGA. Amplicon primers for *EMSY* were *EMSY*-HF1 CCA CCC CAC ATG TCT CCT GTA, *EMSY*-HR1 TGA GCT TGG TGA TGT GGT GAC, *EMSY*-2F TGC CTG TGT GGC CAA, *EMSY*-2R CCT GTG CCC GAA GTG CAC. Optimal conditions for probe 1 were 50°C 2 min, 95°C 10 min, then 45 cycles of 95°C/58°C cycling. Probe 1 was used at 200 nM and primers at 300 nM. For probe 2, the same cycling conditions were used, but probe was used at 100 nM and primers at 900 nM. Normalization to PBGD or TBP gave similar results.

FISH and Tissue Microarray

DNA from the *EMSY* PAC clones (DJ 855A11 824N10, 18d12, 128b4, all PCR verified for exons 2 or 8 of *EMSY*; obtained from HGMP) and from BAC clones containing *EMS1* and *GARP* (gift from Vysis, Downer's Grove, Illinois) were nick translation-labeled with digoxigenin-dUTP or biotin-dUTP (Roche). *Cyclin D1* and centromere 11 probes labeled with Spectrum Orange were purchased from Vysis. Chromosome 11 was paint-labeled by DOP-PCR with rhodamine. Probes were denatured for 10 min at 72°C followed by preannealing for 30 min at 37°C. Target slides were denatured for 1 min at 72°C in 70% formamide/2XSSC and immediately dehydrated in ice-cold ethanol series (70%, 90%, and 100%). Hybridization was done for 72 hr under coverslips at 37°C. The slides were washed and signal detected with antidigoxigenin-FITC (Roche) or avidin-Cy5 (Amersham). Metaphase chromosomes were counterstained with 4,6-diamidino 2-phenylindole (DAPI) in Vectashield solution (Vector Laboratories).

Tissue microarrays consisted of 0.6 mm cores from representative areas of tumor constructed with a Beecher Instruments Micro Tissue Arrayer. Six-micron-thick array sections were used for FISH analysis. The microarray *EMSY* probes were directly labeled with Spectrum Green by nick translation (Vysis, Downer's Grove, Illinois). The *EMSY* probe was combined with Spectrum Aqua-labeled centromeric probe, CEP11 (Vysis) and Spectrum Orange-labeled *Cyclin D1* (Vysis). All probes label targets on 11q exclusively. Pretreatment of tissue microarray sections and posthybridization washes were based on the supplier's recommendations (Vysis). Slides were counterstained with DAPI. FISH signals were counted in forty intact and nonoverlapping nuclei. The ratio of *EMSY* copy number to centromere copy number is the amplification ratio. The Metasystems Metafer unit (Altlussheim, Germany) was used for scoring.

Copy Number Analysis

EMSY copy number was determined as described above using FISH. Copy number for *GARP*, *PAK1*, *CyclinD1 (=CCND1)*, and *EMS1* was determined using array-comparative genomic hybridization (Vysis).

Expression Analysis

Real-time RT-PCR of *EMSY* (as above), *GARP, PAK1*, *Cyclin D1*, and *EMS1* mRNA (Applied Biosystems, TaqMan Assays on Demand) was carried out on an ABI 7900 Sequence Detection System (Applied Biosystems). The specificity of the PCR products was confirmed by melting curve analysis. Transcripts were semiquantified after normalization to the housekeeping TBP gene by comparison with equally normalized expression levels in the diploid breast cancer cell line CAL51, which was used as a common reference.

Statistical Methods

Multivariate survival analyses were conducted using the proportional hazards method (Cox, 1972). The Kaplan-Meier product limit estimate was used to generate survival curves (Kaplan and Meier, 1958).

Acknowledgments

The work in the T.K. laboratory is supported by grants from Cancer Research UK. L.H.-D. is the recipient of a Cancer Research UK Clinician Scientist award. B.P. is a Gibb Fellow of Cancer Research UK. F.F. is supported by the Belgian Fonds National de la Recherche Scientifique. D.H. is supported by a Michael Smith foundation for Health Research Scholarship. The work was also supported in part by the National Cancer Institute of Canada. We thank Andrew Bannister and Daniel Wolf for advice and John Lester for DNA sequencing services.

Received: June 6, 2003
Revised: October 30, 2003
Accepted: October 30, 2003
Published: November 25, 2003

References

Ansieau, S., and Leutz, A. (2002). The conserved Mynd domain of BS69 binds cellular and oncoviral proteins through a common PXLXP motif. J. Biol. Chem. *277*, 4906–4910.

Bachman, K.E., Rountree, M.R., and Baylin, S.B. (2001). Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. J. Biol. Chem. *276*, 32282–32287.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120–124.

Bekri, S., Adelaide, J., Merscher, S., Grosgeorge, J., Caroli-Bosc, F., Perucca-Lostanlen, D., Kelley, P.M., Pebusque, M.J., Theillet, C., Birnbaum, D., and Gaudray, P. (1997). Detailed map of a region commonly amplified at 11q13→q14 in human breast carcinoma. Cytogenet. Cell Genet. *79*, 125–131.

Bochar, D.A., Wang, L., Beniya, H., Kinev, A., Xue, Y., Lane, W.S., Wang, W., Kashanchi, F., and Shiekhattar, R. (2000). BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. Cell *102*, 257–265.

Bornemann, D., Miller, E., and Simon, J. (1998). Expression and properties of wild-type and mutant forms of the Drosophila sex comb on midleg (SCM) repressor protein. Genetics *150*, 675–686.

Boyd, J., Sonoda, Y., Federici, M.G., Bogomolniy, F., Rhei, E., Maresco, D.L., Saigo, P.E., Almadrones, L.A., Barakat, R.R., Brown, C.L., et al. (2000). Clinicopathologic features of BRCA-linked and sporadic ovarian cancer. JAMA *283*, 2260–2265.

Collins, N., Wooster, R., and Stratton, M.R. (1997). Absence of methylation of CpG dinucleotides within the promoter of the breast cancer susceptibility gene *BRCA2* in normal tissues and in breast and ovarian cancers. Br. J. Cancer *76*, 1150–1156.

Courjal, F., Cuny, M., Simony-Lafontaine, J., Louason, G., Speiser, P., Zeillinger, R., Rodriguez, C., and Theillet, C. (1997). Mapping of DNA amplifications at 15 chromosomal localizations in 1875 breast tumors: Definition of phenotypic groups. Cancer Res. *57*, 4360–4367.

Cox, D.R. (1972). Regression models and life tables. J. Roy. Stat. Soc. *34*, 187–220.

Dobrovic, A., and Simpfendorfer, D. (1997). Methylation of the BRCA1 gene in sporadic breast cancer. Cancer Res. *57*, 3347–3350.

Downs, J.A., Lowndes, N.F., and Jackson, S.P. (2000). A role for Saccharomyces cerevisiae histone H2A in DNA repair. Nature *408*, 1001–1004.

Fyodorov, D.V., and Kadonaga, J.T. (2001). The many faces of chromatin remodeling: SWItching beyond transcription. Cell *106*, 523–525.

Gayther, S.A., Pharoah, P.D., and Ponder, B.A. (1998). The genetics of inherited breast cancer. J. Mammary Gland Biol. Neoplasia *3*, 365–376.

Hakimi, M.A., Bochar, D.A., Chenoweth, J., Lane, W.S., Mandel, G., and Shiekhattar, R. (2002). A core-BRAF35 complex containing

histone deacetylase mediates repression of neuronal-specific genes. Proc. Natl. Acad. Sci. USA *99*, 7420–7425.

Harkin, D.P., Bean, J.M., Miklos, D., Song, Y.H., Truong, V.B., Englert, C., Christians, F.C., Ellisen, L.W., Maheswaran, S., Oliner, J.D., and Haber, D.A. (1999). Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1. Cell *97*, 575–586.

Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science *295*, 2080–2083.

Jones, D.O., Cowell, I.G., and Singh, P.B. (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. Bioessays *22*, 124–137.

Kaplan, E.L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. *53*, 457–481.

Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M.J., Sauter, G., and Kallioniemi, O.P. (1998). Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat. Med. *4*, 844–847.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116–120.

Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C.P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends Biochem. Sci. *28*, 69–74.

Milner, J., Ponder, B., Hughes-Davies, L., Seltmann, M., and Kouzarides, T. (1997). Transcriptional activation functions in BRCA2. Nature *386*, 772–773.

Myung, K., Pennaneach, V., Kats, E.S., and Kolodner, R.D. (2003). Saccharomyces cerevisiae chromatin-assembly factors that act during DNA replication function in the maintenance of genome stability. Proc. Natl. Acad. Sci. USA *100*, 6640–6645.

Nielsen, P.R., Nietlispach, D., Mott, H.R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A.G., Murzina, N.V., and Laue, E.D. (2002). Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature *416*, 103–107.

Nordling, M., Karlsson, P., Wahlstrom, J., Engwall, Y., Wallgren, A., and Martinsson, T. (1998). A large deletion disrupts the exon 3 transcription activation domain of the *BRCA2* gene in a breast/ovarian cancer family. Cancer Res. *58*, 1372–1375.

Ormandy, C.J., Musgrove, E.A., Hui, R., Daly, R.J., and Sutherland, R.L. (2003). Cyclin D1, EMS1 and 11q13 amplification in breast cancer. Breast Cancer Res. Treat. *78*, 323–335.

Paull, T., Rogakou, E.P., Yamazaki, V., Kirchgessner, C.U., Gellert, M., and Bonner, W.M. (2000). A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. Curr. Biol. *10*, 886–895.

Pilch, D.R., Sedelnikova, O.A., Redon, C., Celeste, A., Nussenzweig, A., and Bonner, W.M. (2003). Characteristics of gamma-H2AX foci at DNA double-strand breaks sites. Biochem. Cell Biol. *81*, 123–129.

Risch, H.A., McLaughlin, J.R., Cole, D.E., Rosen, B., Bradley, L., Kwan, E., Jack, E., Vesprini, D.J., Kuperstein, G., Abrahamson, J.L., et al. (2001). Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. Am. J. Hum. Genet. *68*, 700–710.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd. (2002). SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. *16*, 919–932.

Shin, S., and Verma, I.M. (2003). BRCA2 cooperates with histone acetyltransferases in androgen receptor-mediated transcription. Proc. Natl. Acad. Sci. USA *100*, 7201–7206.

Smothers, J.F., and Henikoff, S. (2000). The HP1 chromo shadow domain binds a consensus peptide pentamer. Curr. Biol. *10*, 27–30.

Stec, I., Nagl, S.B., van Ommen, G.J., and den Dunnen, J.T. (2000). The PWWP domain: A potential protein-protein interaction domain in nuclear proteins influencing differentiation? FEBS Lett. *473*, 1–5.

Venkitaraman, A.R. (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell *108*, 171–182.

Yang, L., Mei, Q., Zielinska-Kwiatkowska, A., Matsui, Y., Blackburn, M.L., Benedetti, D., Krumm, A.A., Taborsky, G.J., Jr., and Chansky, H.A. (2003). An ERG (Ets-related gene)-associated histone methyltransferase interacts with histone deacetylases 1/2 and transcription co-repressors mSin3A/B. Biochem. J. *369*, 651–657.

Ye, Q., Hu, Y.F., Zhong, H., Nye, A.C., Belmont, A.S., and Li, R. (2001). BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. J. Cell Biol. *155*, 911–921.

Accession Numbers

The Genbank accession number for *EMSY* sequence is CAD22881; gene symbol *C11ORF30*. The Pfam accession number for the ENT domain protein family is PF03735 and for the Agenet domain protein family is PF05641.