Abstract: Materials and methods for determining a susceptibility to cancer are disclosed, and more particularly mutations found in the PALB2 gene that are linked to an increased risk of Fanconi anemia, childhood cancer, breast cancer, or ovarian cancer.
Materials and Methods for Determining Susceptibility to Cancer

Field of the Invention

The present invention relates to materials and methods for determining a susceptibility to cancer, and more particularly mutations found in the PALB2 gene that are linked to an increased risk of cancer.

Background of the Invention

Fanconi anemia is a genetically heterogeneous condition with 12 complementation groups currently recognised, 11 of which have been attributed to distinct genes, FANCA (FA-A), FANCB (FA-B), FANCC (FA-C), BRCA2 (FA-D1), FANCD2 (FA-D2), FANCE (FA-E), FANCF (FA-F), FANCG (FA-G), PALB2 (FA-J), FANCL (FA-L) and FANCM (FA-M). These genes encode interacting proteins that participate in the recognition and repair of DNA interstrand cross-links.

BRCA2 is a DNA repair protein with a key role in the repair of double-strand DNA breaks by homologous recombination. BRCA2 was originally identified through positional cloning of a familial breast cancer predisposition gene and monoallelic (heterozygous) mutations are associated with high risks of breast and ovarian cancer. Subsequently, biallelic BRCA2 mutations were found to cause a rare subtype of Fanconi anemia, FA-D1.

The phenotype of biallelic BRCA2 mutations differs from other Fanconi anemia subtypes, most strikingly in the risk and spectrum of associated cancers. Collectively, all Fanconi anemia subtypes are associated with greatly increased risks of cancer, primarily acute myelogenous leukaemia and adult squamous cell carcinomas, particularly in the oral cavity and the pharyngo-esophageal and female
genital regions. Childhood solid tumors occur very rarely in Fanconi anemia and, to date, have been almost exclusively associated with FA-D1. The risk of childhood embryonal tumors, particularly Wilms tumor and medulloblastoma, is very high in biallelic BRCA2 mutation carriers.

Approximately one in nine women develops breast cancer. The disease is twice as common in sisters and mothers of affected individuals as in the general population and evidence from twin studies indicates that most of this familial excess is due to inherited susceptibility. Studies over the last two decades have provided strong evidence for a limited number of breast cancer susceptibility genes.

The major known breast cancer susceptibility genes are BRCA1 and BRCA2. Germline mutations in BRCA1 and BRCA2 confer a high relative risk of developing breast cancer, of the order of 10-20 fold by age 60. The prevalence of disease-associated mutations in these genes varies in different populations, but is generally low. In the UK, approximately one in 900 individuals is heterozygous for a BRCA1 mutation and one in 800 heterozygous for a BRCA2 mutation. In Ashkenazi Jews, the prevalence of mutations in both genes is higher, with approximately 1% individuals being heterozygous for mutations in each of these genes. Germline mutations in the TP53 gene also confer a high risk of breast cancer. However, TP53 cancer predisposing mutations are much rarer than those in BRCA1 or BRCA2.

In recent years it has been shown that other genes involved in DNA repair and that interact biologically with BRCA1 or BRCA2 are breast cancer susceptibility genes. The breast cancer risks conferred by mutations in these genes are, however, smaller than those associated with
mutations in BRCA1 and BRCA2. For example, CHEK2 encodes a protein kinase that modulates the activities of TP53, BRCA1 and BRCA2. CHEK2 mutations confer a two-fold relative risk of breast cancer and in many countries such mutations are present in 0.5-1% of the population. Similarly, ATM is critical in the cellular response to double stranded DNA damage and regulates the activities of BRCA1 and TP53. Epidemiological and molecular studies have demonstrated that inactivating mutations in ATM confer an approximately two-fold relative risk of breast cancer. In many populations ATM breast cancer predisposing alleles have a similar prevalence to those in CHEK2, with 0.5-1% of the population being heterozygotes. Finally, we recently demonstrated that truncating mutations in BRIPl, which encodes a protein that binds BRCA1 and mediates some of its functions, also confer a two-fold relative risk of breast cancer. BRIPl mutations appear to be rarer than those in CHEK2 and ATM, with 0.1% individuals in the UK population being heterozygotes, a prevalence similar to that of BRCA1 and BRCA2 mutations in the UK.

Despite these advances in the understanding of breast cancer susceptibility, the known susceptibility genes together account for no more than 25% of the familial risk of breast cancer. Hence, most familial aggregation of the disease remains unexplained. Linkage studies suggest that there is unlikely to be a significant contribution from further high risk susceptibility genes like BRCA1 and BRCA2. Therefore, most remaining breast cancer susceptibility is embodied in lower penetrance alleles.

Biallelic (homozygous or compound heterozygous) mutations in some breast cancer susceptibility genes cause recessive childhood syndromes characterised by sensitivity to various types of DNA damage. Biallelic mutations in ATM
have been recognised as the causative defects in Ataxia
Telangiectasia for more than a decade. In addition, it
has been shown that biallelic mutations in BRCA2 and in
BRIP1 cause subgroups of Fanconi anemia (FA), a rare,
recessive, chromosomal instability disorder characterized
by growth retardation, skeletal abnormalities, bone marrow
failure, cancer predisposition and cellular
hypersensitivity to DNA cross linking agents.

Summary of the Invention

Broadly, the present invention is based on research to
identify additional cancer susceptibility genes and in
particular studies that show that mutations in PALB2 are
associated with Fanconi anemia, childhood cancer, breast
cancer and/or ovarian cancer. PALB2 (for "partner and
localizer of BRCA2") encodes a recently discovered protein
that interacts with BRCA2, is implicated in its nuclear
localisation and stability and is required for some
functions of BRCA2 in homologous recombination and double
strand break repair.

Although many cases of Fanconi anemia with childhood solid
tumors are attributable to BRCA2, we identified
individuals with this phenotype in whom BRCA2 mutations
could be excluded. This raised the possibility that
deficiency of other proteins might give rise to this
combination of features and we considered proteins
functionally related to BRCA2 the most credible
candidates.

As we demonstrated that mutations, and more particularly
biallelic mutations, in PALB2 are responsible for a subset
of Fanconi anemia cases characterised by a high incidence
of childhood solid tumors, a phenotype similar to that
caused by biallelic BRCA2 mutations, we also investigated
whether heterozygous (monoallelic) mutations in PALB2
confer susceptibility to breast cancer and/or ovarian cancer.

PALB2 (for 'partner and localizer of BRCA2') was recently identified as a nuclear partner of BRCA2 (Xia et al, 2006). PALB2 colocalises with BRCA2, promoting its localisation and stability in key nuclear structures, which in turn facilitates BRCA2 functions in DNA repair. Furthermore, knockdown of PALB2 sensitizes cells to MMC treatment which results in interstrand cross-links and double-strand breaks. Sensitivity to MMC is a hallmark of Fanconi anemia and we therefore appreciated that PALB2 might be a candidate Fanconi anemia gene.

Thus, the present invention represents the first proof that mutations in the PALB2 gene are linked to a susceptibility (or predisposition as the terms are used interchangeably herein) to Fanconi anemia, childhood cancer, breast cancer or ovarian cancer. The association with Fanconi anemia includes cancers linked to the occurrence of this condition, including childhood embryonal tumors such as Wilms tumors, medulloblastomas, acute myelogenous leukemia (AML) and neuroblastoma.

Accordingly, in a first aspect, the present invention provides a method of determining whether an individual has an increased susceptibility to Fanconi anemia, childhood cancer or breast cancer, the method comprising determining in a sample obtained from the individual the presence of a mutation in the PALB2 gene, or a polypeptide encoded by the PALB2 gene, wherein the presence of a mutation is indicative of the increased risk of Fanconi anemia, childhood cancer, breast cancer or ovarian cancer.

Examples of the mutations disclosed herein that are associated with Fanconi anemia and childhood cancer are
set out in Table 1 and examples of mutations associated with breast cancer are set out in Table 2. In a preferred embodiment, the mutations are truncating mutations. The mutations associated with Fanconi anemia or 'childhood cancer are generally biallelic, while the mutations associated with breast cancer are generally monoallelic.

Further mutations associated with breast and/or ovarian cancer have been found and are G796X, 172delTTGT, 3116delA, Y1183X, 2982insT, W1038X, W906X, 886delA and 695delG.

In a further aspect, the present invention provides a method which comprises having determined whether an individual has an increased susceptibility to Fanconi anemia, childhood cancer, breast cancer or ovarian cancer according to the methods disclosed herein, one or more of the further step of:

(a) correlating the presence of said truncating mutations to a susceptibility to Fanconi anemia, childhood cancer or breast cancer; and/or
(b) saving data representing the result of the test on a recordable media, and/or
(c) transmitting the data representing the result of the test to a recipient.

In a further aspect, the present invention provides a kit for detecting mutations in the PALB2 gene associated with a susceptibility to cancer according to any one of the preceding claims, the kit comprising:

(a) one or more sequence specific probes as disclosed herein; and/or
(b) one or more sequence specific primers for amplifying a portion of the PALB2 nucleic acid sequence as disclosed herein; and/or
(c) one or more specific binding partners capable of
specifically binding to normal or mutated PALB2 polypeptide as disclosed herein; and/or

(d) a microarray as disclosed herein.

In a further aspect, the present invention provides novel nucleic acid and polypeptide sequences that includes an isolated nucleic acid molecule encoding the PALB2 gene having at least 90% nucleic acid sequence identity with the sequence as set out in SEQ ID NO: 2, wherein the nucleic acid comprises one of the mutations set out in Table 1 or Table 2.

In further aspects, the present invention further relates to a replicable vector comprising these nucleic acid sequences and to host cells transformed with the vector, e.g. for use in expressing PALB2 nucleic acid by culturing the host cells so that the polypeptide encoded by the PALB2 nucleic acid is produced. The present invention also provides polypeptides encoded by these nucleic acid molecules and antibodies capable of specifically binding to the PALB2 polypeptides.

Aspects of the present invention will now be further described by way of example with reference to the accompanying drawings, by way of example and not limitation. Further aspects of the invention will be apparent to those or ordinary skill in the art.

**Brief Description of the Figure and Sequences**

**Figure 1.** Evidence that PALB2 deficiency causes FA-N. a) Mutation analysis. Schematic representation of PALB2 showing the position of mutations identified in FA-N cases. Further details of the mutations and the individuals that they were identified in are given in Table 1. The 3549OG, Y1183X mutation was identified in two separate families. b) PALB2 and BRCA2 immunoblotting.
PALB2 protein is absent in cells from FA-N cases IFAR-847 (LCL), LNEY (fibroblast) and GESH (fibroblast) and is present in control cells (left hand panel). BRCA2 is present in FA-N and control cells and is absent in FA-Dl cells (right hand panel). Tubulin or RAD50 was used as a loading control, c) Complementation of cellular FA phenotype. Transduction of cells - from IFAR-849 with the vector pOZC-PALB2 restores PALB2 expression similar to that in a normal control, whereas the mock vector pSllEG (expressing GFP) does not. The transduction leads to reduction in MMC-induced G2 phase arrest in cells transduced with pOZC-PALB2 (black) compared to that in the mock-transduced line pSllEG (grey) grown for the same time at identical MMC concentration. This effect has been reported in complemented FA cell lines from other subtypes d) FANCD2 monoubiquitination. There is normal FANCD2 monoubiquitination in control cells and cells from FA-N cases IFAR-847, LNEY and GESH, whereas FA-C cells show absent FANCD2 monoubiquitination. This indicates that PALB2 acts downstream of FANCD2 in the FA-BRCA pathway.

Figure 2. Cellular phenotype of PALB2 deficiency. a) Severe chromosomal damage in a fibroblast metaphase from GESH treated with 50ng/ml MMC. b) Dose-response curves demonstrate diminished survival of lymphocytes from patient LOAO as opposed to his unaffected sibling. The cells were transduced with eGFP or Fanconi anemia cDNAs as indicated, c) No cells with >5 RAD51 nuclear foci were detectable in fibroblasts from GESH after induction of DNA damage, whereas a normal control clearly showed nuclei (blue) with >5 RAD51 signals (red) . d) The proportion of nuclei with >5 RAD51 foci (black columns) was ~0 in fibroblasts from GESH and LNEY compared to 16% in control cells and 24% in FA-C cells (grey columns, proportion of nuclei with <5 foci) .
**Figure 3.** PALB2 mutations identified in individuals with breast cancer. Chromatograms of five PALB2 mutations identified in eight familial breast cancer cases and the wild-type sequence.

**Figure 4.** Abridged pedigrees of eight breast cancer families with PALB2 mutations. Individual screened for PALB2 mutations is indicated with an arrow. Individuals with breast cancer are shown as filled circles, with the age at diagnosis given underneath. Other cancers or medical conditions are not shown. Samples were not available from individuals with breast cancer that are not genotyped. The PALB2 mutation in each family is shown in Figure 1 and Table 1 and is shown under the individual.

BC, breast cancer; PALB2_WT, PALB2 mutation absent.

**Figure 5.** Schematic diagram of the Fanconi anemia-BRCA pathway. The FA-core complex consists of eight FA proteins (A, B, C, E, F, G, L and M) and is essential for the monoubiquitination and activation of FANCD2 following DNA damage. Activated FANCD2 is translocated to DNA repair foci where it colocalises with other DNA damage response proteins including BRCA2 and RAD51 and participates in homology directed repair. Shaded proteins are encoded by genes that cause Fanconi anemia. Proteins outlined in blue are encoded by genes that confer susceptibility to breast cancer. **BRIP1, BRCA2 and PALB2** are both Fanconi anemia genes and breast cancer susceptibility genes and their encoded proteins all function downstream of FANCD2.

SEQ ID NO: 1 shows the amino acid sequence of PALB2.

SEQ ID NO: 2 shows coding sequence of the PALB2 gene.
**Detailed Description**

**PALB2 gene and polypeptide sequences**

The PALB2 gene and polypeptide sequences are disclosed in Xia et al., and are publicly available on GenBank as sequence accession numbers NM024675 and BC044254. The polypeptide sequence is 1186 amino acids in length and is provided a SEQ ID NO: 1. The coding sequence of the PALB2 gene is reproduced herein as SEQ ID NO: 2. PALB2 nucleic acid includes the sequence shown in SEQ ID NO: 2, alleles and sequence variants thereof and a complementary sequences of any of these nucleic acids. The numbering used herein refers to these sequences and in particular in Tables 1 and 2 to the coding sequence of the PALB2 gene shown in SEQ ID NO: 2. However, the present invention is also applicable to the use of alleles and sequence variants of this gene that may include one or more of the mutations as disclosed herein.

PALB2 nucleic acid and amino acid sequences preferably have at least 90% sequence identity, more preferably 98% sequence identity, and most preferably at least 98% sequence identity, to their respective sequences set out in SEQ ID NO: 1 and 2. "Percent (%) amino acid sequence identity" with respect to the PALB2 polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the PALB2 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. The % identity values can be generated by WU-BLAST-2 which was obtained from [Altschul et al., Methods in Enzymology, 266:460-480 (1996); http://blast.wustl.edu/blast/README.html]. WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set
with the following values: overlap span=1, overlap fraction=0.125, word threshold (T)=11. The HSPS and HSPS2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched; however, the values may be adjusted to increase sensitivity. A % amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the "longer" sequence in the aligned region. The "longer" sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-Blast-2 to maximize the alignment score are ignored).

Similarly, "percent (%) nucleic acid sequence identity" with respect to the coding sequence of the PALB2 polypeptides identified herein is defined as the percentage of nucleotide residues in a candidate sequence that are identical with the nucleotide residues in the PALB2 coding sequence as provided in SEQ ID NO: 1. The identity values used herein were generated by the BLASTN module of WU BLAST-2 set to the default parameters, with overlap span and overlap fraction set to 1 and 0.125, respectively.

Particular mutant alleles of the present invention are set out in Tables 1 and 2 and are described using the nomenclature Antonarakis, Recommendations for a nomenclature system for human gene mutations, Nomenclature Working Group, Hum. Mutat., 11(1):1-3, 1998. These mutations are generally associated with the production of truncated forms of PALB2 polypeptide show in the experimental work described herein to be associated with susceptibility to cancer, and especially to Fanconi anemia, childhood cancer, breast cancer or ovarian cancer.
Implications for screening, e.g. for diagnostic or prognostic purposes, are discussed below.

The finding of mutations to the wild type PALB2 gene sequence means that, in some aspects, the present invention provides novel PALB2 nucleic acid sequences, in particular the mutations set out in Tables 1 and 2 or described elsewhere in the present application. Generally, nucleic acid according to the present invention is provided as an isolate, in isolated and/or purified form, or free or substantially free of material with which it is naturally associated, such as free or substantially free of nucleic acid flanking the gene in the human genome, except possibly one or more regulatory sequence(s) for expression. Nucleic acid may be wholly or partially synthetic and may include genomic DNA, cDNA or RNA. Where nucleic acid according to the invention includes RNA, reference to the sequence shown should be construed as reference to the RNA equivalent, with U substituted for T.

Nucleic acid sequences encoding all or part of the PALB2 gene and/or its regulatory elements can be readily prepared by the skilled person using the information and references contained herein and techniques known in the art (for example, see Sambrook, Fritsch and Maniatis, "Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, and Ausubel et al, Short Protocols in Molecular Biology, John Wiley and Sons, 1992). These techniques include (i) the use of the polymerase chain reaction (PCR) to amplify samples of such nucleic acid, e.g. from genomic sources, (ii) chemical synthesis, or (iii) preparing cDNA sequences.

In order to obtain expression of the PALB2 nucleic acid sequences, including the novel mutated sequences disclosed herein, the sequences can be incorporated in a vector
having control sequences operably linked to the PALB2 nucleic acid to control its expression. The vectors may include other sequences such as promoters or enhancers to drive the expression of the inserted nucleic acid, nucleic acid sequences so that the PALB2 polypeptide is produced as a fusion and/or nucleic acid encoding secretion signals so that the polypeptide produced in the host cell is secreted from the cell. PALB2 polypeptide can then be obtained by transforming the vectors into host cells in which the vector is functional, culturing the host cells so that the PALB2 polypeptide is produced and recovering the PALB2 polypeptide from the host cells or the surrounding medium. Prokaryotic and eukaryotic cells are used for this purpose in the art, including strains of E. coli, yeast, and eukaryotic cells such as COS or CHO cells. The choice of host cell can be used to control the properties of the PALB2 polypeptide expressed in those cells, e.g. controlling where the polypeptide is deposited in the host cells or affecting properties such as its glycosylation.

Methods of determining the presence of mutations
A wide range of techniques are known in the art for determining the presence of a presence of mutations in a gene such as PALB2, or in the polypeptide encoded by it. These techniques may be employed by the skilled person for use in accordance with the present invention. In general, the purpose of carrying of the methods disclosed herein on a sample from an individual is to determine whether the individual carries a PALB2 allele predisposing them to cancer. The purpose of such analysis may be used for diagnosis or prognosis, e.g. to serve to detect the presence of an existing cancer, to help identify the type of cancer, to assist a physician in determining the severity or likely course of the cancer and/or to optimise treatment of it. More preferably, the methods can be used
to detect PALB2 alleles that are statistically associated with a susceptibility to cancer in the future, e.g. early onset breast cancer, identifying individuals who would benefit from regular screening to provide early diagnosis of cancer or for whom changes in lifestyle or diet may help to ameliorate the increased susceptibility to a particular form of cancer.

Broadly, the methods divide into those screening for the presence of PALB2 nucleic acid sequences and those that rely on detecting the presence of PALB2 polypeptide. Exemplary techniques and their advantages and disadvantages are reviewed in Nature Biotechnology, 15:422-426, 1997. The methods make use of biological samples from individuals that may contain the nucleic acid or polypeptides. Examples of biological samples include blood, plasma, serum, tissue samples and saliva.

Nucleic acid based testing may be carried out using preparations containing genomic DNA, cDNA and/or mRNA. Testing cDNA or mRNA has the advantage of the complexity of the nucleic acid being reduced by the absence of intron sequences, but the possible disadvantage of extra time and effort being required in making the preparations. RNA is more difficult to manipulate than DNA because of the widespread occurrence of RN'ases.

Techniques that involve looking for mutations in PALB2 nucleic acid sequence include direct sequencing, restriction fragment length polymorphism (RFLP) analysis, single-stranded conformation polymorphism (SSCP), PCR amplification of specific alleles, amplification of DNA target by PCR followed by a mini-sequencing assay, allelic discrimination during PCR, Genetic Bit Analysis, pyrosequencing, oligonucleotide ligation assay, or analysis of melting curves.
Techniques that involve looking for mutations in PALB2 polypeptides include the use of specific binding members such as antibodies to detect mutated and/or normal PALB2 polypeptides.

Sequencing
Nucleic acid in a test sample may be sequenced and the sequence compared with the sequence shown in SEQ ID NO: 1, for example to determine whether the sequence contains a truncating mutation, such as one of the mutations shown in Table 1 or Table 2, and hence is associated with a susceptibility to cancer. Since it will not generally be time or labour efficient to sequence all nucleic acid in a test sample, or even the whole PALB2 gene, a specific amplification reaction such as PCR using one or more pairs of primers may be employed to amplify the region of interest in the nucleic acid, for instance the PALB2 gene or a particular region in which mutations associated with cancer susceptibility occur. Exemplary primers for this purpose can be designed by the skilled person based on the information provided herein. The amplified nucleic acid may then be sequenced as above and/or tested in any other way to determine the presence or absence of a particular feature. Nucleic acid for testing may be prepared from nucleic acid removed from cells or in a library using a variety of other techniques such as restriction enzyme digest and electrophoresis.

Probes
Mutations in nucleic acid may also be screened using a mutant- or allele-specific probe. Such a probe corresponds in sequence to a region of the PALB2 gene, or its complement, containing a sequence mutation known to be associated with cancer susceptibility, for example as set out in Table 1. Under suitably stringent conditions,
specific hybridisation of such a probe to test nucleic acid is indicative of the presence of the sequence alteration in the test nucleic acid. For efficient screening purposes, more than one probe may be used on the same test sample. This approach may be adapted to use a microarray as discussed in more detail below.

The binding of the probe to target nucleic acid (e.g. DNA) may be measured using any of a variety of techniques at the disposal of those skilled in the art. For instance, probes may be radioactively, fluorescently or enzymatically labelled. Other methods not employing labelling of probe include examination of restriction fragment length polymorphisms, amplification using PCR, RNAase cleavage and allele specific oligonucleotide probing.

Probing may employ the standard Southern blotting technique. For instance DNA may be extracted from cells and digested with different restriction enzymes. Restriction fragments may then be separated by electrophoresis on an agarose gel, before denaturation and transfer to a nitrocellulose filter. Labelled probe may be hybridised to the DNA fragments on the filter and binding determined. DNA for probing may be prepared from RNA preparations from cells.

Those skilled in the art are well able to employ suitable conditions of the desired stringency for selective hybridisation, taking into account factors such as the length of the probe and base composition, temperature and so on. By way of example, stringent conditions include those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridisation a
denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 760 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

Approaches which rely on hybridisation between a probe and test nucleic acid and subsequent detection of a mismatch may be employed. Under appropriate conditions (temperature, pH etc.), an oligonucleotide probe will hybridise with a sequence which is not entirely complementary. The conditions of the hybridisation can be controlled to minimise non-specific binding, and preferably stringent to moderately stringent hybridisation conditions are preferred. The skilled person is readily able to design such probes, label them and devise suitable conditions for the hybridisation reactions, assisted by textbooks such as Sambrook et al (1989) and Ausubel et al (1992). The degree of base-pairing between the two molecules will be sufficient for them to anneal despite a mismatch. Various approaches are well known in the art for detecting the presence of a mismatch between two annealing nucleic acid molecules.

For instance, RN'ase A cleaves at the site of a mis-match. Cleavage can be detected by electrophoresing test nucleic acid to which the relevant probe or probe has annealed and looking for smaller molecules (i.e. molecules with higher
electrophoretic mobility) than the full length probe/test hybrid. Other approaches rely on the use of enzymes such as resolvases or endonucleases.

Thus, an oligonucleotide probe that has the sequence of a region of the normal PALB2 gene (either sense or anti-sense strand) in which mutations associated with cancer susceptibility are known to occur (e.g. see Tables 1 or 2) may be annealed to test nucleic acid and the presence or absence of a mismatch determined. Detection of the presence of a mismatch may indicate the presence in the test nucleic acid of a mutation associated with cancer susceptibility. On the other hand, an oligonucleotide probe that has the sequence of a region of the PALB2 gene including a mutation associated with cancer susceptibility may be annealed to test nucleic acid and the presence or absence of a mismatch determined. The absence of a mismatch may indicate that the nucleic acid in the test sample has the normal sequence. In either case, a plurality of probes to different regions of the gene may be employed.

**PCR methods**

Allele or variant-specific oligonucleotides may similarly be used in PCR to specifically amplify particular sequences if present in a test sample. Assessment of whether a PCR band contains a gene variant may be carried out in a number of ways familiar to those skilled in the art. The PCR product may for instance be treated in a way that enables one to display the mutation or polymorphism on a denaturing polyacrylamide DNA sequencing gel, with specific bands that are linked to the gene variants being selected. PCR techniques for the amplification of nucleic acid are described in US Patent No. 4683195, Mullis et al, Cold Spring Harbor Symp. Quant. Biol., 51:263, (1987), Ehrlich (ed), PCR technology, Stockton Press, NY, 1989,
Restriction digest

The presence of differences in sequence of nucleic acid molecules may be detected by means of restriction enzyme digestion, such as in a method of DNA fingerprinting where the restriction pattern produced when one or more restriction enzymes are used to cut a sample of nucleic acid is compared with the pattern obtained when a sample containing the normal gene or a variant or allele is digested with the same enzyme or enzymes.

Antibodies

There are various methods for determining the presence or absence in a test sample of a mutated form of the PALB2 polypeptide. For example, a sample may be tested for the presence of a binding partner for a specific binding member such as an antibody (or mixture of antibodies), specific for one or more particular variants of the polypeptide, for example the normal PALB2 polypeptide and mutated forms thereof.

In such cases, the sample may be tested by being contacted with a specific binding member such as an antibody under appropriate conditions for specific binding, before binding is determined, for instance using a reporter system as discussed. Where a panel of antibodies is used, different reporting labels may be employed for each antibody so that binding of each can be determined.

A specific binding member such as an antibody may be used to isolate and/or purify its binding partner polypeptide from a test sample in preference to other components that may be present in the sample. This may be used to
determine whether the polypeptide has the sequence shown in SEQ ID NO: 1, or if it is a mutant form. Amino acid sequence is routine in the art using automated sequencing machines. A "specific binding pair" comprises a specific binding member (sbm) and a binding partner (bp) which have a particular specificity for each other and which in normal conditions bind to each other in preference to other molecules. The skilled person will be able to think of many other examples and they do not need to be listed here. It has become a matter of routine in the art for the skilled person to make antibodies that are capable of specifically binding to different polypeptides.

The reactivities of antibodies on a sample may be determined by any appropriate means. Tagging with individual reporter molecules is one possibility. The reporter molecules may directly or indirectly generate detectable, and preferably measurable, signals. The linkage of reporter molecules may be directly or indirectly, covalently, e.g. via a peptide bond or non-covalently. Linkage via a peptide bond may be as a result of recombinant expression of a gene fusion encoding antibody and reporter molecule.

One favoured mode is by covalent linkage of each antibody with an individual fluorochrome, phosphor or laser dye with spectrally isolated absorption or emission characteristics. Suitable fluorochromes include fluorescein, rhodamine, phycoerythrin and Texas Red. Suitable chromogenic dyes include diaminobenzidine.

Other reporters include macromolecular colloidal particles or particulate material such as latex beads that are coloured, magnetic or paramagnetic, and biologically or chemically active agents that can directly or indirectly cause detectable signals to be visually observed,
electronically detected or otherwise recorded. These molecules may be enzymes which catalyse reactions that develop or change colours or cause changes in electrical properties, for example. They may be molecularly excitable, such that electronic transitions between energy states result in characteristic spectral absorptions or emissions. They may include chemical entities used in conjunction with biosensors. Biotin/avidin or biotin/streptavidin and alkaline phosphatase detection systems may be employed.

Microarrays
There is also an increasing tendency in the diagnostic field towards miniaturisation of such assays, e.g. making use of binding agents (such as antibodies or nucleic acid sequences) immobilised in small, discrete locations as arrays on solid supports or on diagnostic chips. The use of microarrays can be particularly valuable as they can provide great sensitivity, particularly through the use of fluorescent labelled reagents, require only very small amounts of biological sample from individuals being tested and allow a variety of separate assays can be carried out simultaneously. This latter advantage can be useful as it provides an assay for different mutations in the PALB2 gene or mutations in other genes to be carried out using a single sample, e.g. in forms of genetic profiling.

Microarrays are libraries of biological or chemical entities immobilised in a grid/array on a solid surface and methods for making and using microarrays are well known in the art. A variation on this theme is immobilisation of these entities onto beads, which are then formed into a grid/array. The entities immobilised in the array can be referred to as probes. These probes interact with targets (a gene, mRNA, cDNA, protein, etc) and the extent of interaction is assessed using
fluorescent labels, colorimetric/chromogenic labels, radioisotope labels or label-free methods (e.g. scanning Kelvin microscopy, mass spectrometry, surface plasmon resonance, etc). The interaction may include binding, hybridization, absorption or adsorption. The microarray process provides a combinatorial approach to assessing interactions between probes and targets. The basic nucleic acid microarray concept is described in US Patent No: 5700637 and 6054270.

One type of array uses nucleic acid molecules as the probes. A DNA microarray is a collection of microscopic DNA spots attached to a solid substrate, e.g. glass, plastic or silicon chip, forming an array. DNA microarrays are now commercially available. There are three basic forms: spotted microarrays, lithographic microarrays and a bead-based systems. Each involves analysing DNA sequences by the immobilisation of cDNA probes or in situ creation of oligonucleotide sequences and subsequent hybridisation with target mRNA/cDNA complementary to the probes. Often the target cDNA are fluorescently labelled. Sequencing by hybridization approaches are described, for example, in US Patent No: 6913879, 6025136, 6018041, 5525464 and 5202231.

Two approaches exist to the creation and immobilisation of DNA probes. In the first approach oligonucleotide sequences are built in situ base by base on the chip. In the second, cDNA or oligonucleotide probes are deposited on the array using contact or non-contact printing methods.

In the spotted microarray approach, oligonucleotides, cDNA or small fragments of PCR products corresponding to mRNAs are printed in an array pattern on a solid substrate by either a spotting robot using pins or variations on ink-
jet printing methods. The spots are typically in the 30-
500 µm size range with separations of the order of 100 µm
or more. A lack of uniformity of spot size, variations of
spot shape and donut or ring-stain patterns caused during
the drying of spots can result in non-uniform
immobilisation of the DNA and hence non-uniform
fluorescence following the hybridisation.

In lithographic microarrays, sequences of oligonucleotides
(A, C, T, G) are built up by selective protection and
deprotection of localised areas of the substrate. This
approach has been employed, inter alia, by Affymetrix.
Affymetrix chips generally provide higher probe densities
(spot sizes of the order of 10 µm or greater), but have
shorter sequence lengths than in spotted or bead
microarrays. The fluorescent labelling of target cDNA
remains a key part of the detection strategy. The
photolithographic approach is described in US Patent No:
6045996 and 5143854.

An alternative method for making arrays employs bead based
microarrays. An example of this approach is the system
used by Illumina (http://www.illumina.com/) in which
probes are immobilised on small (3-5 µm diameter) beads.
After hybridisation the beads are cast onto a surface and
drawn into wells by surface tension. In the Illumina
system, the wells are etched into the ends of optical
fibres in fibre bundles. The fluorescence signal is then
read for each bead. The method includes a tagging of each
bead so that the bioactive agent on each bead can be
decoded from the probe position and a decoding system is
needed to distinguish the different probes used. The bead
based system is described in US Patent No: 6023540,
6327410, 6266459, 6620584 and 7033754.

Whilst the above descriptions relate to DNA microarrays,
the same principles have been extended to protein and chemical microarrays. In these cases the probes immobilised on the surface are specific proteins, antibodies, small molecule compounds, peptides, carbohydrates, etc rather than DNA sequences. The targets are complex analytes, such as serum, total cell extracts, and whole blood. The key concepts of an array of probes, which undergo selective binding/interaction with a target and which are then interrogated via, for example, a fluorescent signal remain central to the method.


Accordingly, in a further aspect, the present invention provides a microarray, or the components for forming a microarray (e.g. a bead array), wherein the microarray comprises one or more binding agents present or locatable on a substrate at a plurality of locations, wherein the one or more binding agents are capable of specifically binding to PALB2 nucleic acid containing a truncating mutation or to a truncated PALB2 polypeptide encoded by the nucleic acid. The microarray will preferably also comprise a plurality of further binding agents for carrying out other tests on the sample, for example to determine the presence of other mutations that are associated with a susceptibility to a disease or condition, such as cancer.
Kits

In a further aspect, the present invention provides kits for carrying out the methods disclosed herein. The components of the kit will be dependent on whether the method is for determining the presence of a mutation in the *PALB2* gene, or a polypeptide encoded by the *PALB2* gene, for example the presence of a truncating mutation, or truncated polypeptide.

Generally, the components of the kit will be provided in a suitable form or package to protect the contents from the external environment. The kit may also include instructions for its use and to assist in the interpretation of the results of the test. The kit may also comprise sampling means for use in obtaining a test sample from an individual, e.g. a swab for removing cells from the buccal cavity or a syringe for removing a blood sample (such components generally being sterile).

In one embodiment, the kit may comprise a microarray as described above, optionally in combination with other reagents, such as labelled developing reagents, useful for carrying out testing with the assay. The microarray is preferably a nucleic acid array.

In other embodiments, the kit may be for use in PCR based testing according to the methods disclosed herein and accordingly may comprise one or more primers suitable for amplifying a portion of the *PALB2* nucleic acid sequence where one of the mutations associated with a susceptibility to cancer are located. The kit may include instructions for use of the nucleic acid, e.g. in PCR and/or a method for determining the presence of nucleic acid of interest in a test sample. In addition to one or more primers (or pairs of primers), the kit may also one or more further reagents required for the reaction, such
as polymerase, nucleosides, buffer solution etc. The nucleic acid primer may also be labelled, for example to facilitate detection and/or quantification of the amplified product.

In a further aspect, the present invention provides a computer program for carrying the method for evaluating a property of a clinical treatment in a group of test subjects.

In a further aspect, the present invention provides a data carrier having a program saved thereon for carrying out the method for evaluating a property of a clinical treatment in a group of test subjects.

In a further aspect, the present invention provides a computer programed to carry out the method for evaluating a property of a clinical treatment in a group of test subjects.

Materials and Methods
Fanconi anemia study
Samples and cell lines
Genomic DNA was obtained from 82 individuals with Fanconi anemia and/or their parents with informed consent. These individuals came from several different clinicians and repositories. All had a clinical diagnosis of Fanconi anemia. This diagnosis was confirmed by study of baseline and DEB- or mitomycin C-induced chromosomal breakage or cell cycle arrest in PHA-stimulated cultured peripheral blood lymphocytes as previously described, except in cases that died before suitable samples could be obtained. In some cases, diagnosis of Fanconi anemia was based on fibroblast analysis. None of the cases included in this study were known to be due to an existing Fanconi anemia gene. However, the extent to which known genes were
formally excluded varied, depending on the available samples. Methods used for exclusion were sequencing, retroviral complementation group analysis, and western blot analysis for FANCD1, FANCD2 and FANCJ. BRCA2 mutations had been excluded in all cases by sequencing of all exons and intron/exon boundaries in genomic DNA from either the case and/or parents. The 176 samples from normal individuals were from the 1958 Birth Cohort Collection (see http://www.els.ioe.ac.uk/Cohort/Ncds/mainncds.htm). This research was approved by the London Multicentre Research Ethics Committee (05/MRE02/17).

**PALB2 sequencing**

Primers were designed to amplify the 13 exons and intron-exon boundaries of PALB2. We used a touchdown 68-50°C protocol to amplify all products which were sequenced using the BigDye Terminator Cycle sequencing kit and a 3730 automated sequencer (ABI Perkin Elmer). Sequencing traces were analysed using Mutation Surveyor software (www.softgenetics.com) and by visual inspection.

**Immunoblotting**

PALB2 immunoblots were performed with samples containing 50 µg total protein each on 7% NuPage Tris-Acetate polyacrylamide gels (Invitrogen). Membranes were probed with polyclonal rabbit anti-PALB2 antiserum raised against the first 120 amino acids of PALB2 at a concentration of 1:1000 (gift of B Xia). Secondary antibody was ECL donkey anti-rabbit IgG horseradish peroxidase-linked whole antibody (GE Healthcare) at 1:2000. Detection was by the chemiluminescence technique using the ECL system (Amersham). BRCA2 immunoblots were performed with samples containing 80 µg total protein each on 3-8% NuPage Tris-Acetate polyacrylamide gels (Invitrogen). Membranes were probed with polyclonal rabbit anti-BRCA2 antiserum (Ab-2,
Calbiochem), raised against amino acids 3245-3418 of BRCA2, at a concentration of 1:200. Secondary antibody (1:5000) and detection were the same as for the PALB2 immunoblots. Immunoblotting for detection of monoubiquitinated FANCD2 was performed using previously described methods.

Transduction and complementation
For complementation studies, the γ-retroviral vectors pOZC-PALB2 and pS11EG (expressing GFP) were packaged in PG cells and used for transduction of patient and control cell lines. Gene transfer was monitored by CD25 (pOZC-PALB2) or GFP (pS11EG) expression. Transduced cells were grown for 48 h in the presence of MMC at concentrations of 12 (fibroblasts) or 15 ng/ml (LCL). The cells were vitally stained with Hoechst 33342 fluorescent dye (Molecular Probes) at 16 µg/ml. DNA histograms were recorded by flow cytometry.

Lymphocyte survival assay
Lymphocyte survival was determined using CD3/CD28/IL2-stimulated lymphocytes. Transduction was with FANC cDNAs (as indicated in Fig 2B) and GFP cDNA as a control, separately cloned into a S11-type γ-retroviral vector. The cells were exposed to various concentrations of MMC for 5 days. Live/dead cell ratios were determined by propidium iodide exclusion/uptake on flow cytometry.

Immunofluorescence
To examine RAD51 foci formation, nuclear foci were induced by ionizing irradiation with 8 Gy of fibroblast cultures grown on glass slides. Cells were fixed 8 h later using 4% paraformaldehyde in PBS (pH 6.8) and permeabilized with 0.1% Triton X-100. Primary antibody was mouse monoclonal anti-RAD51 (GeneTex) at a dilution of 1:100. Secondary antibody was Alexa594-conjugated goat anti-mouse IgG.
(Molecular Probes) at 1:200. DAPI was used as DNA counterstain. Nuclear foci were counted on a fluorescence microscope.

Breast cancer study
To investigate the role of PALB2 in breast cancer predisposition we sequenced the complete coding sequence of the gene in series of familial breast cases and controls. This design enhances the power of the study by exploiting the expected enrichment of predisposition alleles in familial cases and has been employed in our previous studies of CHEK2, ATM and BRIPl mutations in breast cancer susceptibility. The 13 exons and intron-exon boundaries of PALB2 were PCR-amplified and sequenced in 747 breast cancer cases, each from an independently ascertained family with multiple cases of breast cancer. Each case was affected with breast cancer at any age and had at least two relatives with breast cancer. All samples were from breast cancer families ascertained through Cancer Genetics clinics in the UK. A family history of Fanconi anemia was not part of the ascertainment criteria for any of the families. Families from non-UK ethnic groups were excluded. All cases had previously been screened and found negative for mutations and large deletions/duplications in BRCA1 and BRCA2. We quantified the extent of the family history of breast cancer in each family using a Family History Score. This was defined as the number of relatives of the index case with breast cancer, weighted by their degree of relatedness to the index case to adjust for the expected allele sharing (score=1 for each affected 1st degree relatives, 0.5 for 2nd degree relatives and 0.25 for 3rd degree relatives; bilateral breast cancers score double). Where possible, additional breast cancer cases from the families of index cases in whom a PALB2 mutation was found were also genotyped for that mutation. We obtained
informed consent from all families and the research was approved by the London Multicentre Research Ethics Committee (MREC/01/2/18).

906 controls were screened for mutations in the 13 exons and intron-exon boundaries of PALB2 in exactly the same way as cases. Control samples were from the 1958 Birth Cohort Collection which is an ongoing follow-up of all persons born in Great Britain during one week in 1958, including a recent biomedical assessment during 2002-2004 at which blood samples and informed consent were obtained for creation of a genetic resource. At least 97% of the controls were of white ethnicity.

We compared numbers and pathogenicity of mutations or variants between cases and controls using a two-sided Fisher exact test. We used the Wilcoxon test to assess the difference between cases and controls of the variables, age and Family History Score and a two-sided Fisher exact test to compare the numbers of bilateral breast cancers in index cases with and without PALB2 mutations. The analyses were performed using STATA 8.2 (StataCorp, 4905 Lakeway Drive, College Station, Texas 77845 USA).

We estimated the PALB2 combined mutation frequency and the breast cancer risk ratio relative to non-carriers of a PALB2 mutation simultaneously in a segregation analysis using the program MENDEL. We conditioned the likelihood on all phenotypes for each pedigree to correct for ascertainment. Since PALB2 mutation screening had been carried out in all cases and controls we were able to incorporate information from all controls and the full pedigrees of all cases (including those without a PALB2 mutation) in the likelihood together with the segregation information from the families in which a PALB2 mutation
was detected and genotyping was possible in relatives of the index case. We computed risks to non-carriers such that the total risk averaged across carriers and non-carriers agreed with the age-specific breast cancer incidence rates in England and Wales for the period 1992-1997 using segregation analyses. Because this model does not explicitly incorporate the effects of other susceptibility genes, it assumes implicitly that the effects of PALB2 and other potential susceptibility genes can be regarded as independent, as in a multiplicative model.

We estimated the expected number of mutation carriers among the genotyped breast cancer cases (excluding index cases) in the families with identified PALB2 mutations under the null hypothesis according to standard Mendelian inheritance probabilities. The significance of the segregation of mutations with disease was evaluated by computing the observed number of secondary affected individuals that carried the mutation, weighted by degree of relationship. The significance of this statistic was then computed exactly based on the possible inheritance patterns in each family.

Results

Fanconi anemia study

We sequenced the 13 exons and intron-exon boundaries of PALB2 in 82 individuals with Fanconi anemia not due to known genes. We identified pathogenic mutations in seven families (Fig IA, Table 1). In four affected individuals (GESH, IFAR-847, LNEY, IFAR-849) we identified biallelic mutations that result in premature protein truncation. Seven of the mutations were small insertions or deletions that result in translational frameshifts, or base substitutions that generate premature stop codons. RNA analysis in LNEY demonstrated that the second mutation in
this individual, c.3350+4A>G, results in aberrant splicing and generates two abnormal products, r.3202_3350del149, in which exon 12 is skipped, and r.3350insGCAG, which utilises a cryptic splice donor site. Both changes introduce translational frameshifts. Analysis of parental DNA demonstrated that all the mutations had been inherited from different parents, consistent with autosomal recessive inheritance. No sample was available from the affected individuals LOAO, IFAR-007 and ICR-60, but their parents all carried PALB2 mutations. We also sequenced PALB2 in 352 control chromosomes (176 normal individuals). No truncating or splice junction variants were identified providing further evidence that such mutations are pathogenic in the individuals with Fanconi anemia. One mutation, C.35490G, p.Y1183X, was seen in two separate families. A different mutation at the same nucleotide, C.35490A, also results in p.Y1183X and was identified in a third family. C.3549C is in the last exon of PALB2 and there are only three amino acids after codon 1183 before the protein terminates. Truncating mutations close to the end of a protein are generally expected to escape nonsense-mediated RNA decay. However, there was no detectable PALB2 protein in lymphoblastoid cells from IFAR-847 and IFAR-849, who both carry p.Y1183X, indicating that the mutation results in a null allele and confirming its pathogenicity (Fig IB, 1C).

Western blot analysis using antibodies to PALB2 on lymphoblastoid cells or fibroblasts from four individuals with biallelic PALB2 mutations showed absence of PALB2 protein in each case (Fig IB, 1C). Transduction of cells from IFAR-849 with an expression construct carrying wild type PALB2 restored PALB2 expression and reversed MMC-induced G2 phase arrest (Fig 1C). The combined genetic, protein and complementation data provide strong evidence that PALB2 mutations underlie a new Fanconi anemia
complementation group that we have designated subtype FA-N.

Eight of the known Fanconi anemia proteins, (A, B, C, E, F, G, L and M), form a nuclear core complex that mediates monoubiquitination of FANCD2. Activated FANCD2 is translocated to DNA repair foci where it colocalizes with various proteins involved in the DNA damage response, including BRCA2 (refs 6,7). Given the close functional relationship between BRCA2 and PALB2, one would predict that PALB2 acts downstream of FANCD2 in the FA-BRCA pathway. We confirmed this in different cell types from IFAR-847, GESH and LNEY which show normal monoubiquitination of FANCD2 (Fig ID).

The phenotype of FA-N is, in many ways, typical of Fanconi anemia and includes growth retardation and variable congenital malformations (Table 1). The cellular phenotype of PALB2 deficiency is similar to BRCA2 deficiency and more severe than other Fanconi anemia subtypes, with elevated spontaneous chromosome breakage rates and markedly reduced lymphocyte survival and increased chromosome breakage on exposure to MMC (Fig 2A, 2B). There was also no formation of nuclear RAD51 foci in PALB2-deficient fibroblasts after ionising irradiation. Again, this is similar to cells with biallelic BRCA2 mutations and differs from other Fanconi anemia subtypes (Fig 2C, 2D).

FA-N is associated with unusually early mortality due to childhood cancer. All seven FA-N cases developed cancer in early childhood including three Wilms tumors, five medulloblastomas, two cases of AML and one neuroblastoma (Table 1). One individual developed three malignancies within the first year of life and three individuals had two cancers. Cancer treatment was unsuccessful in six
patients, all of whom died before four years of age. LNEY is currently alive at 4.5 years but is in the early stages of treatment for medulloblastoma. The cancer spectrum of biallelic PALB2 mutations is very similar to that associated with biallelic BRCA2 mutations, which also confer high risks of embryonal tumors. The reasons for the association between childhood solid tumors and deficiency of BRCA2 or PALB2 are unclear, but seem likely to be related to functions not shared by other FA proteins.

Monoallelic (heterozygous) BRCA2 mutations are associated with high risks of breast and ovarian cancer and lesser risks of other cancers such as prostate and pancreatic cancer. Given the intimate functional links between PALB2 and BRCA2 and the similar phenotypes associated with biallelic mutations in the genes that encode them, it is plausible that monoallelic PALB2 mutations confer susceptibility to adult cancer. Of interest in this regard are the cancer histories in the seven FA-N families we have identified. The mother of IFAR-007 had early-onset bilateral breast cancer and has a strong family history of breast cancer affecting her sister, mother and other more distant relatives. BRCA1 and BRCA2 mutations have been excluded in this family. No other first-degree relatives of FA-N cases are known to have developed cancer, although most are still less than 50 years of age. However, the maternal grandmother of ICR-60 and the maternal great grandmother of GESH both developed breast cancer at 52 years and 20 years respectively. Mutational analyses of PALB2 in individuals with adult-onset cancer, particularly familial breast cancer, will clarify the role of monoallelic PALB2 mutations in cancer susceptibility. Mutations in BRCA2 and PALB2 together account for almost all our individuals with both Fanconi anemia and childhood solid tumors. However, there are rare Fanconi anemia
cases, not due to known genes, whose cells show normal monoubiquitination of FANCD2. Thus it is possible that other genes encoding proteins that physically and/or functionally interact with BRCA2 cause currently unrecognised subtypes of Fanconi anemia.

Breast cancer study
We identified five different truncating PALB2 mutations, which were present in eight of the 747 (1.1%) familial breast cancer cases compared with none of 906 (0%) controls \( (P = 0.002) \) (Table 2 and Fig 3). Seven mutations were found among the 732 families with female breast cancer only. One mutation was found among the fifteen families that included both male and female breast cancer cases. We were able to evaluate linkage of the mutation with breast cancer in four of the eight PALB2 mutation-positive pedigrees. This revealed limited segregation of the mutation with breast cancer, with only two of five affected relatives of PALB2-mutation carriers also carrying a mutation \( (P = 0.8) \) (Fig 4). This pattern of incomplete segregation in affected relatives is the typical pattern associated with susceptibility alleles that confer modest increased risks and is similar to that previously reported in breast cancer families carrying CHEK2, ATM or BRIP1 mutations.

Segregation analysis incorporating the information from controls and the full pedigrees of the cases estimated the relative risk of PALB2 mutations at 2.3 \( (CI = 1.2-4.1 \ P = 0.006) \). The median age at diagnosis of PALB2-positive cases was 43 years \( [IQR = 40-55] \) compared to a median age at diagnosis of 49 years \( [IQR = 43-56.5] \) in the PALB2-negative breast cancer cases \( (P = 0.31 \text{ for difference}) \). The data suggest that the risks of breast cancer associated with PALB2 mutations may be age-dependent, with a higher RR for women less than 50 years of age \( (RR = 3.1, \)
95% CI = 1.2 - 6.1), although the difference was not significant (P = 0.21). Additional studies are required to address this question. There was no difference in the family history score (P = 0.43) or in the probability of being a bilateral case (P = 0.37) in families with PALB2 mutations compared to families without PALB2 mutations. Assuming a conservative sensitivity of 90% for the PALB2 sequencing we estimate the breast cancer population attributable fraction of PALB2 mutations to be 0.2% (0.05% - 0.6%) and the percentage of the familial relative risk due to PALB2 to be 0.24% (0.013%-1.32%).

We also identified 62 non-truncating PALB2 variants including 28 non-synonymous variants and 15 synonymous variants. There was no overall evidence that PALB2 missense variants confer susceptibility to breast cancer (P = 0.9). Only four missense variants had an allele frequency greater than 0.5% and there was also no evidence that any of these were breast cancer susceptibility alleles (P > 0.5 for each). This result is consistent with the data from Fanconi anemia cases reported herein in which all identified PALB2 mutations reported to date result in premature protein truncation.

Further study examining mutations associated with breast and/or ovarian cancer
A group of further mutations in the PALB2 gene were identified as follows.
G796X - associated with breast cancer.
172delTTGT - associated with breast and/or ovarian cancer.
3116delA - associated with breast cancer.
Y1183X - associated with breast cancer.
2982insT - associated with breast cancer.
W1038X - associated with breast cancer.
W906X - associated with breast cancer.
886delA - associated with ovarian cancer and/or breast cancer.
695delG - associated with ovarian cancer.

5 Discussion
We have demonstrated that truncating germline PALB2 mutations are present in eight out of 747 BRCA1/2 mutation-negative familial breast cancer cases compared to none out of 906 controls providing strong evidence that PALB2 is a breast cancer susceptibility gene.

These results further elucidate the emerging complex relationship between breast cancer susceptibility and the Fanconi anemia pathway. Fanconi anemia is a genetically heterogeneous recessive condition that currently includes 13 subtypes, 12 of which have been attributed to distinct genes. The known FA genes encode proteins that interact in a complex, and incompletely understood, fashion to facilitate recognition and repair of DNA double-strand breaks. A key process in the pathway involves eight of the known FA proteins forming a nuclear core complex that mediates monoubiquitination and activation of FANCD2. Activated FANCD2 is translocated to DNA repair foci where it colocalizes with BRCA2 and other proteins that effect DNA repair by homologous recombination (Fig 5). A broader role for the FA proteins in monitoring oxidative stress and in the stabilisation of stalled replication forks that arise during recombination and repair has also been proposed. There also appears to be extensive cross-talk between the Fanconi anemia pathway and other DNA repair pathways as FA proteins are found in complexes with many other DNA repair proteins, including BRCA1.

Biallelic mutations of BRCA2 and PALB2 cause Fanconi anemia subtypes FA-D1 and FA-N respectively. The phenotypes associated with biallelic BRCA2 and PALB2
mutations are strikingly similar to each other and differ from the other 10 known Fanconi anemia genes. In particular, FA-D1 and FA-N are associated with high risks of solid childhood malignancies, such as Wilms tumor and medulloblastoma, which occur very rarely in other subtypes. Heterozygous mutations in the BRIP1 gene, which encodes a BRCA1-interacting protein, also confer an elevated risk of breast cancer and biallelic BRIP1 mutations cause Fanconi anemia subtype FA-J. However, FA-J is associated with the classical Fanconi anemia phenotype and no FA-J case with a childhood solid tumor has been reported. The biological explanations for these phenotypic similarities and differences are unknown.

It is plausible that heterozygosity for mutations in other Fanconi anemia genes may also be involved in breast cancer susceptibility. However, epidemiological studies of relatives of FA cases have not demonstrated this, suggesting that breast cancer susceptibility is associated only with a subset of Fanconi anemia genes. This is consistent with the negative results of mutational screens of other FA genes in familial breast cancer cases. The biological features that determine whether an FA gene is also a breast cancer predisposition gene are unknown.

However, it is notable that the three FA genes currently associated with breast cancer susceptibility, BRCA2, PALB2 and BRIP1 are not part of the FA core complex and are the only known FA genes that act downstream of FANCD2 (Fig 5).

We estimate that the female breast cancer risk associated with PALB2 mutations is approximately two-fold. Therefore, despite the fact that PALB2 is functionally associated with BRCA2 and that biallelic mutations in both genes generate a similar phenotype of Fanconi anemia, the increase in breast cancer risk associated with PALB2 monoallelic (heterozygous) mutations is clearly more
modest than that conferred by BRCA2 mutations (10-20 fold). These differences in risk are reminiscent of the differences previously reported between BRCA1 mutations, which also confer a 10-20 fold risk of breast cancer, and mutations in BRIP1, which confer only a two-fold risk. In both instances, the biological explanations for the apparent differences in risk associated with mutations in these genes despite the close physical and functional interactions between their encoded proteins are currently unknown.

Seven of the PALB2 mutations were in the 732 families with female breast cancer only (0.96%). One was among the 15 families (6.7%) with cases of both female and male breast cancer (p = 0.15). While this observation requires further investigation it is interesting because it may suggest that PALB2 mutations confer a higher relative risk of male than female breast cancer and BRCA2 mutations are also known to confer a high relative risk of male breast cancer.

A clearer impression of the structure of breast cancer susceptibility is now beginning to emerge. BRCA1 and BRCA2 are likely to be the only major high penetrance cancer susceptibility genes. Together, mutations in these genes account for <20% of the familial risk of breast cancer. The remaining susceptibility is therefore predominantly due to genes conferring more modest increases in risk. CHEK2, ATM, BRIP1 and now PALB2 have been shown to be breast cancer susceptibility genes. Compared to BRCA1 and BRCA2, inactivating mutations in these four genes confer relatively small increases in risk, approximately 2-2.5 fold. The mutations conferring risk in these four genes are all relatively rare with fewer than 1% of the population being heterozygotes. As a consequence, the contribution of each gene to the overall familial risk of
breast cancer is relatively small; together, they account for approximately 2.3% of the overall familial relative risk. However, since four genes encoding proteins that interact biologically with BRCA1 and BRCA2 are now known to be breast cancer susceptibility genes, it would not be surprising if mutations in other components of these pathways are shown to confer similar risks of breast cancer. Collectively, therefore, this class of susceptibility gene may be making an appreciable contribution to the familial risk of breast cancer.

Diagnostic testing for mutations in BRCA1 and BRCA2 has been routine clinical practice in many countries for several years. It facilitates risk estimation and implementation of cancer prevention strategies and increasingly has potential to influence cancer therapy. It is now evident that there exists a further series of susceptibility genes, characterized by multiple, individually rare, inactivating germline mutations that confer smaller increased breast cancer risks than BRCA1 or BRCA2. The clinical management of this type of susceptibility will require careful and cautious consideration. However, as such genes are increasingly being discovered their clinical utility deserves closer attention, particularly as recent biological studies suggest that cancer cells with mutations in DNA repair genes may have enhanced sensitivity to certain drugs. Further research will be required before genetic testing for these genes can be implemented, in order to generate more precise estimates of the risks, their impact on screening and prevention protocols, and to provide a better understanding of how mutations in these genes interact with non-genetic factors and with other susceptibility genes. The latter may be of particular importance in the context of familial breast cancer where it seems likely that convergence of multiple low
penetrance susceptibility alleles is responsible for much familial clustering. The existence of a large number of different, rare susceptibility alleles in a substantial set of genes may additionally pose significant technical and economic challenges to clinical diagnostic services. However, despite these challenges, the discovery of lower penetrance breast cancer susceptibility genes provides increasing opportunities for improved management of women predisposed to develop breast cancer.
Table 1 PALB2 mutations and clinical features of FA-N cases

<table>
<thead>
<tr>
<th>ID</th>
<th>Origin</th>
<th>Mutation 1</th>
<th>Mutation 2</th>
<th>Cancer (age at diagnosis)</th>
<th>Other features</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOAOa</td>
<td>Albanian / Moroccan</td>
<td>C 395delT</td>
<td>p V132fs</td>
<td>Presumed splice defect</td>
<td>medulloblastoma (3 yr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3113+5OC</td>
<td></td>
<td>Growth retardation, radial ray hypoplasia, absent right kidney</td>
</tr>
<tr>
<td>GESH</td>
<td>White (German)</td>
<td>c 757_758delCT</td>
<td>p L253fs</td>
<td>p K1098fs</td>
<td>Wilms tumor (1 yr), AML (9 yr), medulloblastoma (1 yr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C 3294_3298</td>
<td></td>
<td>Severe growth retardation, hypoplastic thumbs, left pelvic kidney, anal atresia,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>del1GACGA</td>
<td></td>
<td>microphthalmia, congenital cataract, microphthalmia, skin hypopigmentation</td>
</tr>
<tr>
<td>IFAR-847</td>
<td>Hispanic / white (N Am)</td>
<td>c 2257C&gt;T</td>
<td>p R753X</td>
<td>p Y1183X</td>
<td>Wilms tumor (5 yr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C 3549C&gt;A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNEY</td>
<td>White (German)</td>
<td>c 2394_2395insCT</td>
<td>p T799fs</td>
<td>r exil2del/p G1068fs</td>
<td>medulloblastoma (4 yr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>c 3350,4A</td>
<td>r 3350msGCAG/p Fll1βfs</td>
<td>Growth retardation, microphalma, microphthalmia</td>
</tr>
<tr>
<td>IFML-0G7</td>
<td>White (N Am) / Black</td>
<td>c 2521del1A</td>
<td>p TR41fs</td>
<td>p Y110βfs</td>
<td>Wilms tumor (1 5 yr), medulloblastoma (1 5 yr)</td>
</tr>
<tr>
<td></td>
<td>(N Am)</td>
<td></td>
<td>c 3323del1A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICR-60'</td>
<td>White (UK)</td>
<td>c 2952C&gt;T</td>
<td>p Q988X</td>
<td>p Y1183X</td>
<td>medulloblastoma (2 3 yr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C 3549G&gt;C</td>
<td></td>
<td>Growth retardation, microphalma, skin hyperpigmentation, horseshoe kidney,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
<td>gonadal dysgenesis</td>
</tr>
<tr>
<td>IFAR-849</td>
<td>White (N Am)</td>
<td>c 3116del1A</td>
<td>p N1039fs</td>
<td>p Y1183X</td>
<td>neuroblastoma (0 7 yr) AML (2 yr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C 3549G&gt;C</td>
<td></td>
<td>Growth retardation, microphalma, VSD, ASD, thumb and radial anomalies, skin</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
<td>hyperpigmentation</td>
</tr>
</tbody>
</table>

No samples were available from the affected individuals and the mutations were therefore identified in parental samples.

*Origin relates to mutation 1  †Origin relates to mutation 2
*This mutation is thought to effect splicing as it occurs at a known donor splice site.
†This individual’s clinical details have previously been published.
AML, acute myelogenous leukaemia; IUGR, intrauterine growth retardation; ASD, atrial septal defect; VSD, ventricular septal defect; N Am, North American ancestry.
Table 2. PALB2 mutations identified in familial breast cancer cases.

<table>
<thead>
<tr>
<th>Family</th>
<th>Mutation</th>
<th>Effect</th>
<th>Number of cases (n=747)</th>
<th>Number of controls (n=906)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G2386T</td>
<td>G796X</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2982insT</td>
<td>premature truncation</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3,4</td>
<td>G3113A</td>
<td>W1038X</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>3116delA</td>
<td>premature truncation</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6,7,8</td>
<td>C3549G</td>
<td>Y1183X</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

The pedigrees of Families 1-8 are shown in Fig 3.
References:
All publications, patent and patent applications cited herein or filed with this application, including references filed as part of an Information Disclosure Statement are incorporated by reference in their entirety.


Claims:

1. A method of determining whether an individual has an increased susceptibility to Fanconi anemia, childhood cancer, breast cancer or ovarian cancer, the method comprising determining in a sample obtained from the individual the presence of a mutation in the *PALB2* gene, or a polypeptide encoded by the *PALB2* gene, wherein the presence of a mutation is indicative of the increased risk of Fanconi anemia, childhood cancer, breast cancer or ovarian cancer.

2. The method of claim 1, wherein the mutation is a truncating mutation.

3. The method of claim 1 or claim 2, wherein the mutation is a biallelic mutation.

4. The method of any one of the preceding claims, wherein the mutation is set out in Table 1 or Table 2 or is selected from the group consisting of G796X, 172delTTGT, 3116delA, Y1183X, 2982insT, W1038X, W906X, 886delA and 695delG.

5. The method of any one of the preceding claims, wherein the step of determining the presence of a mutation in the *PALB2* gene uses direct sequencing, hybridisation to a probe, restriction fragment length polymorphism (RFLP) analysis, single-stranded conformation polymorphism (SSCP), PCR amplification of specific alleles, amplification of DNA target by PCR followed by a mini-sequencing assay, allelic discrimination during PCR, Genetic Bit Analysis, pyrosequencing, oligonucleotide ligation assay, or analysis of melting curves.

6. The method of any one of claims 1 to 4, wherein determining the presence of a mutation in the *PALB2* gene
comprises sequencing the PALB2 gene in the sample, or a portion thereof known to contain a mutation, to determine whether the mutation is present in the PALB2 gene in the sample.

7. The method of any one of claims 1 to 4, wherein determining the presence of a mutation comprises contacting nucleic acid in the sample with a sequence specific probe capable of binding to a PALB2 gene sequence comprising one or more mutations under hybridising conditions and the method comprising contacting the probe and the test sample under hybridising conditions and observing whether hybridisation takes place.

8. The method of any one of claims 1 to 4, wherein determining the presence of a mutation in the PALB2 gene comprises digesting a sample comprising the PALB2 gene with one or more restriction enzymes to cut the nucleic acid and produce a restriction pattern for comparison with patterns obtained with a normal PALB2 gene or a mutated form thereof.

9. The method of any one of claims 1 to 4, wherein determining the presence of a mutation comprises contacting a sample containing PALB2 gene, or a portion thereof, with one or more sequence specific primers that are capable of priming the amplification of the nucleic acid if a normal or mutated form of the PALB2 gene is present in the sample.

10. The method of any one of claims 6 to 9 which comprises the initial step of amplifying the PALB2 nucleic acid present in the sample.

11. The method of any one of claims 1 to 4, wherein determining the presence of a mutation comprises
contacting a sample with a specific binding partner capable of specifically binding to normal or mutated PALB2 polypeptide.

12. The method of claim 11, wherein the specific binding member is an antibody.

13. The method of any one of the preceding claims, wherein the step of determining the presence of a mutation uses a microarray.

14. The method of claim 13, wherein the microarray is a spotted microarray, a lithographic microarray or a bead-based microarray.

15. The method of claim 13 or claim 14, wherein the microarray comprises a plurality of nucleic acid probes or a plurality of antibodies.

16. A method which comprises having determined whether an individual has an increased susceptibility to Fanconi anemia, childhood cancer, breast cancer or ovarian cancer according to the method of any one of the preceding claims, one or more of the further step of:

(a) correlating the presence of said truncating mutations to a susceptibility to Fanconi anemia, childhood cancer, breast cancer or ovarian cancer; and/or

(b) saving data representing the result of the test on a recordable media; and/or

(c) transmitting the data representing the result of the test to a recipient.

17. A kit for detecting mutations in the PALB2 gene associated with a susceptibility to cancer according to any one of the preceding claims, the kit comprising:

(a) one or more sequence specific probes as set out
in claim 7; and/or
  (b) one or more sequence specific primers for
amplifying a portion of the PALB2 nucleic acid sequence as
set out in claim 9; and/or
  (c) one or more specific binding partners capable of
specifically binding to normal or mutated PALB2
polypeptide as set out in claim 11; and/or
  (d) a microarray as set out in any one of claims 13
to 15.

18. An isolated nucleic acid molecule encoding the PALB2
gene having at least 90% nucleic acid sequence identity
with the sequence as set out in SEQ ID NO: 2, wherein the
nucleic acid comprises one of the mutations set out in
Tables 1 or 2 or a mutation selected from the group
consisting of G796X, 172delTTGT, 3116del1A, Y1183X,
2982insT, W1038X, W906X, 886del1A and 695delG.

19. A replicable vector comprising nucleic acid of claim
18, operably linked to control sequences to direct its
expression.

20. A host cell transformed with the vector of claim 19.

21. A method of expressing PALB2 nucleic acid comprising
culturing host cells of claim 20 so that the polypeptide
encoded by the PALB2 nucleic acid is produced.

22. The method of claim 21 comprising the further step of
recovering the polypeptide produced.

23. A polypeptide encoded by the nucleic acid molecule of
claim 18.

24. An antibody capable of specifically binding to a
PALB2 polypeptide of claim 23.
Figure 1
Figure 2

Figure 3
Figure 4
Figure 5