

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-246550  
(P2010-246550A)

(43) 公開日 平成22年11月4日(2010.11.4)

(51) Int.Cl.	F I	テーマコード(参考)
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00 Z N A A	2 G 0 4 5
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 A	4 B 0 2 4
C 1 2 Q 1/02 (2006.01)	C 1 2 Q 1/02	4 B 0 6 3
A 6 1 P 35/00 (2006.01)	A 6 1 P 35/00	4 C 0 8 4
A 6 1 P 17/00 (2006.01)	A 6 1 P 17/00	

審査請求 有 請求項の数 36 O L 外国語出願 (全 185 頁) 最終頁に続く

(21) 出願番号 特願2010-119683 (P2010-119683)  
 (22) 出願日 平成22年5月25日 (2010. 5. 25)  
 (62) 分割の表示 特願2002-539507 (P2002-539507) の分割  
 原出願日 平成13年11月2日 (2001. 11. 2)  
 (31) 優先権主張番号 60/245, 756  
 (32) 優先日 平成12年11月3日 (2000. 11. 3)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 592090692  
 ダナ ファーバー キャンサー インSTITUTE, インコーポレイテッド  
 アメリカ合衆国 マサチューセッツ州 O 2 1 1 5 ポストン ビニー ストリート 4 4  
 (71) 出願人 503061809  
 オレゴン ヘルス アンド サイエンス ユニバーシティ  
 アメリカ合衆国, オレゴン 9 7 2 3 9 - 3 0 9 8 ポートランド サウスウェスト サム ジャクソン パーク ロード 3 1 8 1  
 (74) 代理人 100083806  
 弁理士 三好 秀和

最終頁に続く

(54) 【発明の名称】 癌感受性の診断のための方法及び組成物並びに欠損DNA修復メカニズム及びその処置

(57) 【要約】 (修正有)

【課題】 癌感受性の診断のための方法及び組成物、欠損DNA修復メカニズム及びその処置の提供。

【解決手段】 ファンコーニ貧血及び癌の診断、新規治療薬をスクリーニングのためのプローブ及びプライマーとしてのFANCS2遺伝子配列の使用。欠損DNA修復に関する症状を処置するための新規治療薬をスクリーニングに使用する実験マウスモデル、遺伝子治療陽ベクターの調製。FANCD2アイソフォームの同定と、2つのアイソフォームを区別するポリクローナル及びモノクローナル抗体の調製とその診断試験での使用。

【選択図】 図8

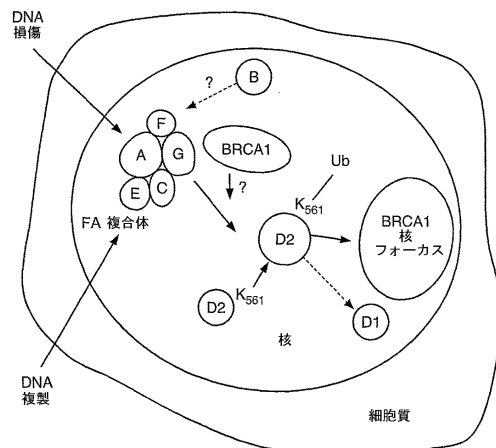


FIG. 8

## 【特許請求の範囲】

## 【請求項 1】

(a) 配列番号：4 に示すアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列；

(b) (a) のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列；

(c) (b) のヌクレオチド配列に相補的なヌクレオチド配列；

(d) 配列番号：5 から 8、187 から 188 に示すヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列；及び

(e) (d) のヌクレオチド配列に相補的なヌクレオチド配列；

から選択されるポリヌクレオチドを含んでなる単離された核酸分子。

10

## 【請求項 2】

ポリヌクレオチドがDNA分子である請求項1に記載の単離された核酸分子。

## 【請求項 3】

ポリヌクレオチドがcDNA分子である請求項2に記載の単離された核酸分子。

## 【請求項 4】

ポリヌクレオチドがRNA分子である請求項1に記載の単離された核酸分子。

## 【請求項 5】

DNA修復を促進するための細胞の核のFANCD2の短い形態から長い形態への変換の生物学的特性を保持するために配列番号：4のアミノ酸配列に十分類似したアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列から本質的に構成される単離された核酸分子。

20

## 【請求項 6】

配列番号：9 から 191 に少なくとも90%同一であるか、または配列番号：9 から 191 に少なくとも90%同一であるヌクレオチド配列に相補的であるヌクレオチド配列を有するポリヌクレオチドから本質的に構成される単離された核酸分子。

## 【請求項 7】

配列が表6に開示される配列番号：9 から 94 より選択されるイントロン/エクソン配列である請求項6に記載の単離された核酸分子。

## 【請求項 8】

配列が表7に開示される配列番号：115 から 186 より選択されるPCRプライマーである請求項6に記載の単離された核酸分子。

30

## 【請求項 9】

請求項1の単離された核酸分子をベクターに挿入することからなる組換えベクターを製作する方法。

## 【請求項 10】

請求項9の方法により製造される組換えベクター。

## 【請求項 11】

請求項10の組換えベクターを宿主細胞に導入することからなる組換え宿主細胞を製作する方法。

## 【請求項 12】

請求項11の方法により製造される組換え宿主細胞。

40

## 【請求項 13】

(a) FA-D2に関連する相補群のピアレリック変異を有する対象から細胞を入手すること；及び

(b) 細胞を形質転換ウイルスで感染させてFA-D2細胞系を作成すること；  
からなるFA-D2細胞系の作製方法。

## 【請求項 14】

細胞が繊維芽細胞及びリンパ球から選択される請求項13に記載の方法。

## 【請求項 15】

形質転換ウイルスがエプスタイン・バー・ウイルス及びレトロウイルスから選択される

50

請求項 13 に記載の方法。

【請求項 16】

さらに欠損 F A N C D 2 の存在を決定することにより F A - D 2 細胞系を特徴付けることからなる請求項 13 に記載の方法。

【請求項 17】

F A - D 2 細胞系の特徴付けがさらに ( i ) F A N C D 2 に特異的な抗体を用いるウェスタンブロットまたは核免疫蛍光；及び ( i i ) D N A ハイブリダイゼーションアッセイから選択される診断アッセイを細胞系で実施することを含む請求項 16 に記載の方法。

【請求項 18】

請求項 1 の単離された核酸分子を含んでなる組換え宿主細胞を培養することからなるポリペプチドを製造するための組換え方法。

10

【請求項 19】

( a ) 配列番号：4；

( b ) ( a ) に少なくとも 90% 同一であるアミノ酸配列；

( c ) 配列番号：5 から 8、187 から 188 の少なくとも 1 つに少なくとも 90% 同一であるヌクレオチド配列を有するポリヌクレオチドによりコードされるアミノ酸配列；及び

( d ) フラグメントが少なくとも 50 アミノ酸の長さである ( a ) から ( d ) のポリペプチドフラグメント；

から選択されるアミノ酸配列を含んでなる単離されたポリペプチド。

20

【請求項 20】

n t 376 の A から G、n t 3707 の G から A、n t 904 の C から T、及び n t 958 の C から T より選択される変異を有する D N A によりコードされる請求項 19 に記載の単離されたポリペプチド。

【請求項 21】

多型が n t 1122 の A から G、n t 1440 の T から C、n t 1509 の C から T、n t 2141 の C から T、n t 2259 の T から C、n t 4098 の T から G、n t 4453 の G から A より選択される、ポリペプチドをコードする D N A の多型によりポリペプチドが特徴付けられる請求項 19 に記載の単離されたポリペプチド。

【請求項 22】

ポリペプチドがアミノ酸 222 またはアミノ酸 561 での変異により特徴付けられる請求項 19 に記載の単離されたポリペプチド。

30

【請求項 23】

F A N C D 2 タンパク質に結合特異性を有する抗体調製物。

【請求項 24】

さらにモノクローナル抗体を含んでなる請求項 23 に記載の抗体調製物。

【請求項 25】

さらにポリクローナル抗体を含んでなる請求項 23 に記載の抗体調製物。

【請求項 26】

F A N C D 2 タンパク質が F A N C D 2 - S である請求項 23 に記載の抗体調製物。

40

【請求項 27】

F A N C D 2 タンパク質が F A N C D 2 - L である請求項 23 に記載の抗体調製物。

【請求項 28】

( a ) サンプルを第 1 抗体に暴露して F A N C D 2 - L と第 1 複合体を形成し、場合によっては第 2 抗体に暴露して F A N C D 2 - S と第 2 複合体を形成すること；

( b ) マーカーでサンプル中の第 1 複合体及び第 2 複合体の量を検出すること；からなる生物学的サンプル中の F A N C D 2 アイソフォームを測定するための診断方法。

【請求項 29】

サンプルが無傷細胞を含んでなる請求項 28 に記載の診断方法。

【請求項 30】

50

サンプルがライゼート中に溶解された細胞を含んでなる請求項 28 に記載の診断方法。

【請求項 31】

生物学的サンプルが癌に対する感受性を有するかまたは初期段階の癌を有するヒト対象からのものである請求項 28 に記載の診断方法。

【請求項 32】

生物学的サンプルがヒト対象の癌からのものであり、ここで癌が黒色腫、白血病、アストチトーム、神経膠芽細胞腫、リンパ腫、神経膠腫、ホジキンリンパ腫、慢性リンパ球性白血病及び膵臓、乳房、卵巣、子宮、精巣、脳下垂体、腎臓、胃、食道及び直腸の癌から選択される請求項 31 に記載の診断方法。

【請求項 33】

生物学的サンプルがヒト胎児からのものである請求項 28 に記載の診断方法。

【請求項 34】

生物学的サンプルが成人からのものである請求項 28 に記載の診断方法。

【請求項 35】

生物学的サンプルが血液サンプル、対象及び細胞系からの組織の生検サンプルから選択される請求項 28 に記載の診断方法。

【請求項 36】

生物学的サンプルが心臓、脳、胎盤、肝臓、骨格筋、腎臓、膵臓、脾臓、胸腺、前立腺、精巣、子宮、小腸、結腸、末梢血またはリンパ球に由来する請求項 28 に記載の診断方法。

【請求項 37】

マーカーが蛍光マーカー、場合によっては FANCD2 - L 抗体に抱合されていてよい蛍光マーカーである請求項 28 に記載の診断方法。

【請求項 38】

マーカーが化学ルミネセンスマーカー、場合によっては FANCD2 - L 抗体に抱合されていてよい化学ルミネセンスマーカーである請求項 28 に記載の診断方法。

【請求項 39】

さらに第 1 及び第 2 複合体が基質に抱合された第 3 抗体に結合することからなる請求項 28 に記載の診断方法。

【請求項 40】

ライゼートを分離手順に供して FANCD2 アイソフォームを分離し、第 1 または第 2 FANCD2 抗体への結合を決定することにより分離されたアイソフォームを同定する請求項 30 に記載の診断方法。

【請求項 41】

FANCD2 タンパク質に対する抗体を選択し、細胞集団における FANCD2 - L アイソフォームの量が野生型集団における量と比較して低減されているかどうかを決定し； FANCD2 - L タンパク質の量が低減されている場合、次いで細胞集団における FANCA、FANCB、FANCC、FANCD1、FANCE、FANCF または FANCG タンパク質のいずれかの量が野生型集団における量と比較して変化しているかどうかを決定して、細胞集団におけるファンコーニ貧血経路における欠損を同定すること；  
からなる対象からの細胞集団におけるファンコーニ貧血経路における欠損を同定するための診断試験。

【請求項 42】

アイソフォームの量の決定が FANCD2 - L 及び FANCD2 - S アイソフォームの分離に依存する請求項 41 に記載の診断試験。

【請求項 43】

分離がゲル電気泳動により達成される請求項 41 に記載の診断試験。

【請求項 44】

分離が遊走結合バンド試験ストリップにより達成される請求項 41 に記載の診断試験。

【請求項 45】

10

20

30

40

50

F A N C D 2 - L が低減された量で作製される細胞集団を選択すること；  
細胞集団を候補治療分子のライブラリーの個々のメンバーに暴露すること；及び  
細胞集団において F A N C D 2 - L の量を増加させるこれらの個々のメンバー分子を同  
定すること；  
からなる治療薬を同定するためのスクリーニングアッセイ。

【請求項 4 6】

細胞集団がインビトロ細胞集団である請求項 4 5 に記載のスクリーニングアッセイ。

【請求項 4 7】

細胞集団がインビボ細胞集団であり、インビボ細胞集団が実験動物内にあり、実験動物  
が変異 F A N C D 2 遺伝子を有している請求項 4 5 に記載のスクリーニングアッセイ。

10

【請求項 4 8】

実験動物がノックアウトマウスであり、ここでマウスの F A N D 2 遺伝子がヒト変異 F  
A N C D 2 遺伝子により置換されている請求項 4 5 に記載のスクリーニングアッセイ。

【請求項 4 9】

F A N C D 2 が低減された量で作製される細胞集団に化学発癌物質を添加し、分子のい  
ずれかのメンバーが F A N C D 2 - L の量を増加させて化学発癌物質の有害な影響から細  
胞を保護するかどうかを決定する請求項 4 5 に記載のスクリーニングアッセイ。

【請求項 5 0】

動物の F A N C D 2 遺伝子が除去されており、場合によっては請求項 1 の核酸分子によ  
り置換されていてもよい実験動物モデル。

20

【請求項 5 1】

対象からの細胞サンプルにおいて、変異が疑われる F A N C D 2 アレルの変異 F A N C  
D 2 ヌクレオチド配列を同定する方法であって、変異が疑われる F A N C D 2 アレルのヌ  
クレオチド配列を野生型 F A N C D 2 ヌクレオチド配列と比較し、ここで疑わしい変異体  
及び野生型配列間の差異により細胞サンプル中の変異 F A N C D 2 ヌクレオチド配列が同  
定される方法。

【請求項 5 2】

変異が疑われるアレルが生殖細胞系アレルである請求項 5 1 に記載の方法。

【請求項 5 3】

変異 F A N C D 2 ヌクレオチド配列の同定により対象における癌の素因が診断される請  
求項 5 1 に記載の方法。

30

【請求項 5 4】

変異 F A N C D 2 ヌクレオチド配列の同定によりファンコーニ貧血を有する子孫を産む  
対象の危険性の増加が診断される請求項 5 1 に記載の方法。

【請求項 5 5】

変異が疑われるアレルが腫瘍型の体細胞アレルであり、変異 F A N C D 2 ヌクレオチド  
配列を同定することにより腫瘍型が診断される請求項 5 1 に記載の方法。

【請求項 5 6】

野生型のヌクレオチド配列及び変異が疑われる F A N C D 2 ヌクレオチド配列が遺伝子  
、 m R N A 及び m R N A から作製される c D N A から選択される請求項 5 1 に記載の方法  
。

40

【請求項 5 7】

変異が疑われる F A N C D 2 アレルのポリヌクレオチド配列を野生型 F A N C D 2 ポリ  
ヌクレオチド配列と比較することがさらに：変異 F A N C D 2 ヌクレオチド配列に特異的  
にハイブリダイズする F A N C D 2 プロブを選択すること、及びプロブでのハイブリ  
ダイゼーションにより変異配列の存在を検出することを含む請求項 5 1 に記載の方法。

【請求項 5 8】

変異が疑われる F A N C D 2 アレルのポリヌクレオチド配列を野生型 F A N C D 2 ポリ  
ヌクレオチド配列と比較することがさらに：野生型 F A N C D 2 D N A に特異的なブラ  
イマーのセットを用いて F A N C D 2 遺伝子の全てまたは一部を増幅して増幅された F A

50

N C D 2 D N A を製造すること、及び F A N C D 2 D N A を解読して変異配列を同定することを含む請求項 5 1 に記載の方法。

【請求項 5 9】

変異 F A N C D 2 ヌクレオチド配列がヒト対象の F A N C D 2 アレルにおける生殖細胞系の変化であり、変化が表 3 に示す変化から選択される請求項 5 1 に記載の方法。

【請求項 6 0】

変異 F A N C D 2 ヌクレオチド配列がヒト対象の F A N C D 2 アレルにおける体細胞の変化であり、変化が表 3 に示す変化から選択される請求項 5 1 に記載の方法。

【請求項 6 1】

対象からの組織サンプルにおける F A N C D 2 遺伝子の生殖細胞系の配列またはその m R N A の配列を F A N C D 2 遺伝子の生殖細胞系の配列またはその m R N A の配列と比較することを含む対象における癌に対する感受性を診断する方法であって、ここで対象の F A N C D 2 遺伝子の生殖細胞系の配列またはその m R N A の配列における変化が癌に対する感受性を示す方法。

【請求項 6 2】

F A N C D 2 遺伝子の制御領域において変化を検出する請求項 6 1 に記載の方法。

【請求項 6 3】

生殖細胞系の配列における変化の検出が ( a ) 非変性ポリアクリルアミドゲルにおける一本鎖 D N A の電気泳動可動性のシフトを観察すること、( b ) F A N C D 2 遺伝子プローブを組織サンプルから単離されたゲノム D N A にハイブリダイズすること、( c ) アレル特異性プローブを組織サンプルのゲノム D N A にハイブリダイズすること、( d ) 組織サンプルからの F A N C D 2 遺伝子の全てまたは一部を増幅して増幅された配列を製造し、増幅された配列をシーケンシングすること、( e ) 特異的 F A N C D 2 変異アレルのプライマーを用いて組織サンプルから F A N C D 2 遺伝子の全てまたは一部を増幅すること、( f ) 組織サンプルからの F A N C D 2 遺伝子の全てまたは一部を分子クローニングしてクローン化された配列を製造し、クローン化された配列をシーケンシングすること、( g ) ( i ) 組織サンプルから単離した F A N C D 2 遺伝子または F A N C D 2 m R N A 、及び ( i i ) ヒト野生型 F A N C D 2 遺伝子配列に相補的な核酸プローブ間の誤対合を同定することであって、この場合、分子 ( i ) 及び ( i i ) は互いにハイブリダイズして二重鎖を形成する、( h ) 組織サンプル中の F A N C D 2 遺伝子配列の増幅及び増幅された配列の野生型 F A N C D 2 遺伝子配列を含んでなる核酸プローブへのハイブリダイゼーション、( I ) 組織サンプル中の F A N C D 2 遺伝子配列の増幅及び増幅された配列の、変異 F A N C D 2 遺伝子配列を含んでなる核酸プローブへのハイブリダイゼーション、( j ) 組織サンプル中の欠失変異に関するスクリーニング、( k ) 組織サンプル中の点変異に関するスクリーニング、( l ) 組織サンプル中の挿入変異に関するスクリーニング、( m ) 組織サンプル中の F A N C D 2 遺伝子との、F A N C D 2 遺伝子を含んでなる核酸プローブとのインサイチュウ・ハイブリダイゼーションからなる群から選択されるアッセイにより決定される請求項 6 1 に記載の方法。

【請求項 6 4】

( a ) 対象からの遺伝物質を評価し、欠損 D N A 修復を決定すること；

( b ) F A N D 2 及び少なくとも 1 つの F A N C A 、 F A N C B 、 F A N C C 、 F A N C D 1 、 F A N C D E 、 F A N D F 、 F A N D G 、 B R A C A 1 及び A T M を含んでなる遺伝子のセットにおける変異の存在を決定すること；並びに

( c ) 遺伝子のセットにおける変異の存在から癌の感受性を診断すること；  
からなる対象における癌の感受性を診断するための方法。

【請求項 6 5】

対象の損傷からの組織サンプルにおける F A N C D 2 遺伝子の配列またはその m R N A の配列を野生型 F A N C D 2 遺伝子の配列またはその m R N A の配列と比較し、ここで対象の F A N C D 2 遺伝子の配列またはその m R N A の配列における変化が腫瘍性損傷の F A N C D 2 遺伝子での変異を示すことを含む：ヒト対象における F A N C D 2 遺伝子での

10

20

30

40

50

腫瘍性損傷における変異を検出する方法。

【請求項 66】

腫瘍性損傷の F A N C D 2 遺伝子での変異に従って腫瘍性損傷を処置するための治療プロトコルを決定することをさらに含む請求項 65 に記載の方法。

【請求項 67】

ヒト対象の損傷からの組織サンプルにおける F A N C D 2 遺伝子の配列またはその m R N A の配列を野生型 F A N C D 2 遺伝子の配列またはその R N A の配列と比較し、ここで組織サンプルにおける野生型配列の存在が F A N C D 2 遺伝子での変異の欠如を示すことを含むヒト対象からの腫瘍性損傷における F A N C D 2 変異の欠如を確認する方法。

【請求項 68】

( a ) 請求項 25 に記載の F A N C D 2 アイソフォームを測定することにより対象からの細胞サンプルにおいて F A N C D 2 - L の欠乏を生じるかどうかを決定すること；

( b ) ( a ) で欠乏が検出された場合、次いで欠乏が非癌細胞における遺伝的欠損の結果であるかどうかを決定すること；及び

( c ) ( b ) が陽性である場合、D N A 損傷を増加させて対象における正常な組織を保護する治療プロトコルの使用を減らし、( b ) が陰性であり、欠乏は癌細胞における遺伝的欠損にのみ含まれる場合、次いで D N A 損傷を増加させ癌細胞に悪影響を与える治療プロトコルの使用を増やす；

ことからなる、癌を有する対象の治療プロトコルを決定する方法。

【請求項 69】

有効量の F A N C D 2 タンパク質または外来性核酸を標的に投与することからなる細胞標的ににおける F A 経路欠損を処置する方法。

【請求項 70】

F A 経路欠損が欠損 F A N C D 2 遺伝子であり、外来性核酸ベクターがさらに請求項 10 に記載のベクターを導入すること含んでなる請求項 69 に記載の方法。

【請求項 71】

ベクターが変異ヘルペスウイルス、E 1 / E 4 欠失組換えアデノウイルス、変異レトロウイルス、感染性新規ウイルス粒子の生成に必須のウイルス遺伝子に関して欠損があるウイルスベクターから選択される請求項 69 に記載の方法。

【請求項 72】

ベクターが脂質ミセルに含まれる請求項 69 に記載の方法。

【請求項 73】

配列番号：4 に記載のポリペプチドを提供し、欠損 F A N C D 2 遺伝子から生じる条件から生じる欠損を機能的に修正すること；

からなる欠損 F A N C D 2 遺伝子で患者を処置する方法。

【請求項 74】

( a ) 対象から細胞サンプルを入手すること；

( b ) 細胞サンプルを D N A 損傷物質に暴露すること；及び

( c ) F A N C D 2 - L が上方制御されているかどうかを検出することであって、上方制御の不在が F A 経路欠損を示している；

からなる F A 経路欠損を検出するための細胞基盤のアッセイ。

【請求項 75】

核フォーカスを検出するためのイムノプロットティング；F A N C D 2 アイソフォームの量を検出するためのウェスタンプロット及び D N A プロブでのハイブリダイゼーションにより m R N A を定量することにより F A N C D 2 の量が測定される請求項 74 に記載の細胞基盤のアッセイ。

【請求項 76】

( a ) 高ストリンジェント条件下で F A N C D 2 遺伝子の配列に結合するプライマー対であって、プライマー対は表 7 に記載される変化した核酸配列を特異的に増幅するために選択される；及び

10

20

30

40

50

(b) 各々のプライマーの容器：

からなる生物学的サンプル中の癌細胞を検出するのに使用するためのキット。

【発明の詳細な説明】

【技術分野】

【0001】

本明細書に記載した研究は国立衛生研究所の援助を受け、NIH グラントvNo.ヘルス・グラントRO1HL52725-04、RO1DK43889-09、1PO1HL48546及びPO1HL54785-04であった。米国政府は請求された発明に対して特定の権利を有する。

【背景技術】

【0002】

本発明はFANCD2遺伝子に欠損を有する対象における癌感受性の診断及び癌が進行したこれらの対象のための適当な処置プロトコルの決定に関する。FANCD2遺伝子に欠損を有する動物モデルを用いて治療薬のスクリーニングをすることができる。

【0003】

ファンコーニ貧血(FA)は先天性欠損、骨髄障害及び癌素因により特徴づけられる常染色体劣性癌感受性症候群である。FA患者からの細胞はDNA鎖間架橋を生み出す作用物質、例えばマイトマイシンCまたはジエポキシブタンに対して特徴的な過敏症を示す。FA患者は急性骨髄性白血病及び皮膚、胃腸；及び婦人科系の癌などのいくつかの型の癌を進行させる。皮膚及び胃腸腫瘍は通常扁平上皮癌である。少なくとも20%のFAの患者で癌を進行させる。癌を進行させる患者の平均年齢は白血病で15歳、肝臓腫瘍で16歳、及びその他の腫瘍で23歳である。(D'Andreaら、Blood 90:1725(1997)、Garcia-Higueraら、Curr. Opin. Hematol. 2:83-88(1999)及びHeijnaら、Am. J. Hum. Genet. 66:1540-1551)

【0004】

FAは遺伝的に異種性である。体細胞融合研究は少なくとも7つの別個の相補群で同定された(Joenjeら、Am. J. Hum. Genet. 61:940-944(1997)及びJoenjeら、Am. J. Hum. Genet. 67:759-762(2000))。この観察により、FA遺伝子がDNA架橋に対する細胞応答に關与する多成分性経路を明確化するという仮説に至った。FA遺伝子の5つ(FANCA、FANCC、FANCE、FANCF及びFANCG)がクローン化され、FANCA、FANCC、及びFANCGタンパク質が主に核局在を伴う分子複合体を形成することが示されている。FANCCは細胞質にも局在する。異なるFAタンパク質は、無脊椎動物種のFANCA、FANCC、FANCE、FANCF及びFANCGタンパク質の強相合体を有さない公知の配列モチーフを少ししか、または全く有さない。FANCFは大腸菌(E. coli)RNA結合タンパク質に対して有意性が未知の弱い相同性を有する。2つの最も頻繁な相補群は、一緒ではFA患者の75%から80%を占めるFA-A及びFA-Cである。80kbにわたり、少なくとも43個のエクソンからなるFANCA遺伝子では複数の変異が認識されている。FANCCは14個のエクソンを有し、およそ80kbにわたることが見出されている。FANCC遺伝子における多くの変異が同定されており、これはの重篤度の程度が異なるFAと相関する。FA-Dは明確ではあるが稀な相補群として同定されている。FA-D患者は別のサブタイプの患者とは表現型で区別できるが、FAタンパク質複合体は通常FA-D細胞に集まる(Yamashitaら、P. N. A. S. 95:13085-13090(1998))。

【0005】

クローン化されたFAタンパク質は互いにまたはジェンバンク(GenBank)の別のタンパク質に配列類似性を有さないオーファンタンパク質をコードし、タンパク質配列が明白である機能的ドメインはない。これらのタンパク質の細胞性または生化学的機能に関してはほとんど知られていない。

10

20

30

40

50

## 【 0 0 0 6 】

F A の診断は F A 患者の表現型の広範な多様性により複雑になっている。さらに診断を混乱させることに、およそ 33% の F A 患者が明白な先天性異常を有していない。さらに、既存の診断試験は一般の集団から F A キャリヤを区別しない。診断に伴う問題は D ' A n d r e a r a ( 1 9 9 7 ) により記載されている。F A 細胞で多くの細胞性表現型が報告されているが、最も一貫性があるのは 2 価性アルキル化剤、例えばマイトマイシン C またはジエボキシブタンに対する過敏症である。これらの作用物質は D N A の鎖間架橋 ( D N A 損傷の重要なクラス ) を作製する。

## 【 0 0 0 7 】

癌の形成に関与している非常に多くの制御遺伝子及び生化学的経路のために、癌感受性の診断は複雑になっている。どのように発生したかに依存する異なる癌及び関与する遺伝的損傷は、対象がいずれかの特定の治療的処置にどのように応答するかを決定し得る。欠損修復メカニズムに伴う遺伝的損傷は細胞分割障害及びアポトーシスを生じることができ、これが今度は患者の癌に対する感受性を高め得る。F A は癌感受性に加えて複数の病理学的結果が欠損修復メカニズムに関連付けられる疾病症状である。

10

## 【 0 0 0 8 】

ファンコーニ貧血経路の分子遺伝学及び細胞生物学の理解により、非 F A 患者及び F A 患者に生じる D N A 修復メカニズムにおける欠損に関連する特定のクラスの癌及び症状の予後、診断及び処置に洞察を供し得る。

## 【 発明の概要 】

20

## 【 課題を解決するための手段 】

## 【 0 0 0 9 】

本発明の最初の実施形態では、( a ) 配列番号 : 4 に示すアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列 ; ( b ) ( b ) のポリヌクレオチドに少なくとも 90% 同一であるヌクレオチド配列 ; ( c ) ( b ) のポリヌクレオチドに相補的なヌクレオチド配列 ; ( d ) 配列番号 : 5 から 8、187 から 188 に示すヌクレオチド配列に少なくとも 90% 同一であるヌクレオチド配列 ; 及び ( e ) ( d ) のヌクレオチド配列に相補的であるヌクレオチド配列から選択されるポリヌクレオチドを含む単離された核酸分子が提供される。ポリヌクレオチドは R N A 分子または D N A 分子、例えば c D N A であり。

## 【 0 0 1 0 】

30

本発明の別の実施形態では、D N A 修復を促進するために、細胞の核において F A N C D 2 の短い形態から長い形態に変換する生物学的特性を保持するために配列番号 : 4 のアミノ酸配列に十分類似したアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列から本質的に構成される単離された核酸分子が提供される。また別に、単離された核酸分子は配列番号 : 9 から 191 に少なくとも 90% 同一であるか、または配列番号 : 9 から 191 に少なくとも 90% 同一であるヌクレオチド配列に相補的であるヌクレオチド配列を有するポリヌクレオチドから本質的に構成される。

## 【 0 0 1 1 】

1 つの実施形態では、前記した単離された核酸分子のいずれかをベクターに挿入することを含む組換えベクターを作製するための方法が提供される。組換えベクター生成物をこの方法により作製し、ベクターを宿主細胞に導入して組換え宿主細胞を形成することができる。

40

## 【 0 0 1 2 】

本発明の 1 つの実施形態では、( a ) F A - D 2 に関連する相補群におけるピアレリック変異を有する対象から細胞を得ること ; 及び ( b ) 形質転換ウイルスで細胞を感染させて、細胞を繊維芽細胞及びリンパ球から選択できる F A - D 2 細胞系を作成すること、並びにエプスタイン・バー・ウイルス及びレトロウイルスから選択されるウイルスを形質転換すること ; を含む F A - D 2 細胞系を作製するための方法が提供される。例えば ( i ) ウェスタンブロットまたは F A N C D 2 に特異的な抗体を用いる核免疫蛍光及び ( i i ) D N A ハイブリダイゼーションアッセイから選択される診断アッセイを実施することによ

50

り細胞系における欠損 F A N D C 2 の存在を決定することにより F A - D 2 細胞系を特徴づけすることができる。

【 0 0 1 3 】

本発明の 1 つの実施形態では、宿主細胞が前記したいずれかの単離された核酸分子を含む組換え宿主細胞を培養することを含む、ポリペプチドを生成するための組換え方法が提供される。

【 0 0 1 4 】

本発明の 1 つの実施形態では、単離されたポリペプチドは ( a ) 配列番号 : 4 ; ( b ) ( a ) に少なくとも 9 0 % 同一であるアミノ酸配列 ; ( c ) 配列番号 : 5 から 8 、 1 8 7 から 1 8 8 の少なくとも 1 つに少なくとも 9 0 % 同一であるヌクレオチド配列を有するポリヌクレオチドによりコードされるアミノ酸配列 ; ( d ) 配列番号 : 5 から 8 、 1 8 7 から 1 8 8 の少なくとも 1 つに相補的な配列に少なくとも 9 0 % 同一であるヌクレオチド配列を有するポリヌクレオチドによりコードされるアミノ酸配列 ; 及び ( e ) フラグメントが少なくとも 5 0 個のアミノ酸の長さである ( a ) から ( d ) のポリペプチドフラグメントから選択されるアミノ酸配列を含む。

10

【 0 0 1 5 】

単離されたポリペプチドは n t 3 7 6 A から G 、 n t 3 7 0 7 G から A 、 n t 9 0 4 C から T 及び n t 9 5 8 C から T より選択される変異を有する DNA によりコードされ得る。また別にポリペプチドをコードする DNA の多型によりポリペプチドを特徴づけることができ、多型は n t 1 1 2 2 A から G 、 n t 1 1 4 0 T から C 、 n t 1 5 0 9 C から T 、 n t 2 1 4 1 C から T 、 n t 2 2 5 9 T から C 、 n t 4 0 9 8 T から G 、 n t 4 4 5 3 G から A より選択される。また別にアミノ酸 2 2 2 またはアミノ酸 5 6 1 での変異によりポリペプチドを特徴づけることができる。

20

【 0 0 1 6 】

本発明の 1 つの実施形態では、 F A N C D 2 タンパク質に関して結合特異性を有する抗体調製物が記載されており、ここで抗体はモノクローナル抗体またはポリクローナル抗体でよく、ここで F A N C D 2 は F A N C D 2 - S または F A N C D 2 - L でよい。

【 0 0 1 7 】

本発明の 1 つの実施形態では、生物学的サンプル中の F A N C D 2 アイソフォームを測定するための診断方法が提供され、ここで方法は ( a ) サンプルを第 1 抗体に暴露して F A N C D 2 - L との第 1 複合体を形成し、場合によっては第 2 抗体に暴露して F A N C D 2 - S との第 2 複合体を形成してもよいこと ; 及び ( b ) サンプル中の第 1 複合体及び第 2 複合体の量をマーカーを用いて検出することを含む。サンプルは無傷の細胞またはライゼート中の溶解された細胞でよい。

30

生物学的サンプルは癌に対する感受性を有するか、または第 1 段階の癌を有するヒト対象からのものでよい。サンプルはヒト対象の癌からのものでよく、ここで癌は黒色腫、白血病、アストチトーム、神経膠芽細胞腫、リンパ腫、神経膠腫、ホジキンリンパ腫、慢性リンパ球性白血病及び膵臓、乳房、卵巣、子宮、精巣、脳下垂体、腎臓、胃、食道及び直腸の癌から選択される。生物学的サンプルはヒト胎児または成人からのものでよく、血液サンプル、対象及び細胞系からの組織の生検サンプルのいずれかに由来するものでよい。生物学的サンプルは心臓、脳、胎盤、肝臓、骨格筋、腎臓、膵臓、脾臓、胸腺、前立腺、精巣、子宮、小腸、結腸、末梢血またはリンパ球に由来してよい。マーカーは蛍光マーカー、場合によっては F A N C D 2 - L 抗体に抱合されていてもよい蛍光マーカー、場合によっては F A N C D 2 - L 抗体に抱合されていてもよい化学ルミネセンスマーカーでよく、第 1 及び第 2 複合体を基質に抱合された第 3 抗体に結合させることができる。サンプルがライゼートである場合、これを分離手段に供して F A N C D 2 アイソフォームを分離でき、第 1 または第 2 F A N C D 2 抗体への結合を決定することにより分離されたアイソフォームを同定できる。

40

【 0 0 1 8 】

本発明の 1 つの実施形態では、対象からの細胞集団のファンコーニ貧血経路における欠

50

損を同定するための診断試験が提供され、試験にはFANCD2タンパク質に対する抗体の選択及び、野生型細胞集団における量と比較してFANCD2-Lアイソフォームの量が細胞増殖において低下するかどうかを決定すること；例えばFANCD2-Lタンパク質の量が低下する場合、次いでFANCA、FANCB、FANCC、FANCD1、FANCE、FANCFまたはFANCGタンパク質のいずれかの量が野生型と比較して細胞集団において変化するかどうかを決定して、細胞集団におけるファンコーニ貧血経路における欠損を同定することが含まれる。1つの事例では、アイソフォームの量はFANCD2-L及びFANCD2-Sアイソフォームの分離に依存し、ここで分離はゲル電気泳動によるかまたは遊走結合バンド試験ストリップにより達成できる。

#### 【0019】

本発明の1つの実施形態では、治療薬を同定するためのスクリーニングアッセイが提供され、このアッセイにはFANCD2-Lが低減された量で生成される細胞集団を選択すること；細胞集団を候補治療分子のライブラリーの個々のメンバーに暴露すること；及び細胞集団においてFANCD2-Lの量の増加を引き起こすこれらの個々のメンバー分子を同定することが含まれる。1つの事例では、細胞集団はインビトロ細胞集団である。別の事例では、細胞集団はインビボ細胞集団、実験動物内に在るインビボ集団、変異FANCD2遺伝子を有する実験動物である。さらなる事例では、実験動物はマウスFANCD2遺伝子がヒト変異FANCD2遺伝子により置換されているノックアウトマウスである。別の事例では、FANCD2-Lが低減された量で生成される細胞集団に化学発癌物質を添加し、いずれかのメンバー分子がFANCD2-Lの量を増加させて化学発癌物質の有害な影響から細胞を保護するかどうかを決定する。

#### 【0020】

本発明の1つの実施形態では、動物FANCD2遺伝子が除去されており、場合によっては前記したいずれかの核酸分子で置換されていてもよい実験動物モデルが提供される。

#### 【0021】

本発明の1つの実施形態では、変異が疑われるFANCD2アレルのヌクレオチド配列を野生型FANCD2ヌクレオチド配列と比較することからなる、対象からの細胞サンプル中の変異が疑われるFANCD2アレルにおける変異FANCD2ヌクレオチド配列を同定するための方法が提供され、ここで疑わしい変異体と野生型配列の差異により細胞サンプル中の変異FANCD2ヌクレオチド配列が同定される。1つの事例では、疑わしい変異体アレルは生殖細胞系アレルである。別の事例では、変異FANCD2ヌクレオチド配列の同定は、対象における癌の素因及び対象がファンコーニ貧血の子孫を産む危険性の高まりを診断するためのものである。別の事例では、変異が疑わしいアレルは腫瘍型の体細胞アレルであり、変異FANCD2ヌクレオチド配列の同定は腫瘍型を診断するためのものである。別の事例では、野生型ヌクレオチド配列及び変異が疑われるFANCD2ヌクレオチド配列は遺伝子、mRNA及びmRNAから作製されたcDNAから選択される。別の事例では変異が疑わしいFANCD2アレルと野生型FANCD2ポリヌクレオチド配列との比較はさらに変異体FANCD2ヌクレオチド配列に特異的にハイブリダイズするFANCD2プローブを選択し、プローブでのハイブリダイゼーションにより変異体配列の存在を検出することを含む。

別の事例では、変異が疑われるFANCD2アレルと野生型FANCD2ポリヌクレオチド配列との比較は、さらに野生型FANCD2 DNAに特異的なプライマーのセットを用いてFANCD2遺伝子の全てまたは一部を増幅して増幅されたFANCD2 DNAを生成し、FANCD2 DNAを解読して変異体配列を贈呈することを含む。別の事例では、変異体FANCD2ヌクレオチド配列がヒト対象のFANCD2アレルにおける生殖細胞系の変化である場合、変化は表3に示す変化から選択され、変異体FANCD2ヌクレオチド配列がヒト対象のFANCD2アレルにおける体細胞の変化である場合、変化は表3に示す変化から選択される。

#### 【0022】

本発明の1つの実施形態では、対象における癌の感受性を診断するための方法が提供さ

10

20

30

40

50

れ、この方法は F A N C D 2 遺伝子の生殖細胞系配列または対象からの組織サンプルのその m R N A の配列を F A N C D 2 遺伝子の生殖細胞系配列またはその m R N A の配列と比較することからなり、ここで F A N C D 2 遺伝子の生殖細胞系配列または対象のその m R N A の配列における変化は癌に対する感受性を示す。F A N C D 2 遺伝子の制御領域において変化を検出できる。生殖細胞系配列における変化を、( a ) 非変性ポリアクリルアミドゲルにおける一本鎖 D N A の電気泳動可動性のシフトを観察すること、( b ) F A N C D 2 遺伝子プローブを組織サンプルから単離されたゲノム D N A にハイブリダイズすること、( c ) アレル特異性プローブを組織サンプルのゲノム D N A にハイブリダイズすること、( d ) 組織サンプルからの F A N C D 2 遺伝子の全てまたは一部を増幅して増幅された配列を製造し、増幅された配列をシーケンシングすること、( e ) 特異的 F A N C D 2 変異体アレルのプライマーを用いて組織サンプルから F A N C D 2 遺伝子の全てまたは一部を増幅すること、( f ) 組織サンプルからの F A N C D 2 遺伝子の全てまたは一部を分子クローニングしてクローン化された配列を製造し、クローン化された配列をシーケンシングすること、( g ) ( i ) 組織サンプルから単離した F A N C D 2 遺伝子または F A N C D 2 m R N A、及び ( i i ) ヒト野生型 F A N C D 2 遺伝子配列に相補的な核酸プローブ間の誤対合を同定することであって、この場合、分子 ( i ) 及び ( i i ) は互いにハイブリダイズして二重鎖を形成する、( h ) 組織サンプル中の F A N C D 2 遺伝子配列の増幅及び増幅された配列の野生型 F A N C D 2 遺伝子配列を含んでなる核酸プローブへのハイブリダイゼーション、( i ) 組織サンプル中の F A N C D 2 遺伝子配列の増幅及び増幅された配列の、変異 F A N C D 2 遺伝子配列を含んでなる核酸プローブへのハイブリダイゼーション、( j ) 組織サンプル中の欠失変異に関するスクリーニング、( k ) 組織サンプル中の点変異に関するスクリーニング、( l ) 組織サンプル中の挿入変異に関するスクリーニング、( m ) 組織サンプル中の F A N C D 2 遺伝子の、F A N C D 2 遺伝子を含んでなる核酸プローブとのインサイチュウ・ハイブリダイゼーションからなる群から選択されるアッセイにより決定することができる。

#### 【 0 0 2 3 】

本発明の 1 つの実施形態では、対象における癌の感受性を診断するための方法が提供され、この方法は ( a ) 対象からの遺伝物質を評価し、欠損 D N A 修復を決定すること；( b ) F A N C D 2 及び少なくとも 1 つの F A N C A、F A N C B、F A N C C、F A N C D 1、F A N C D E、F A N D F、F A N D G、B R A C A 1 及び A T M を含んでなる遺伝子のセットにおける変異の存在を決定すること；並びに ( c ) 遺伝子のセットにおける変異の存在から癌の感受性を診断することを含む。

#### 【 0 0 2 4 】

本発明の 1 つの実施形態では、ヒト対象における F A N C D 2 遺伝子での腫瘍性損傷の変異を検出する方法が提供され、この方法は；

対象の損傷からの組織サンプルの F A N C D 2 遺伝子の配列またはその m R N A 配列を野生型 F A N C D 2 遺伝子の配列またはその m R N A 配列と比較することからなり、ここで対象の F A N C D 2 遺伝子の配列またはその m R N A 配列における変化は腫瘍性損傷の F A N C D 2 遺伝子での変異を示している。腫瘍性損傷の F A N C D 2 遺伝子での変異に従って腫瘍性損傷を処置するための治療用プロトコルを提供できる。

#### 【 0 0 2 5 】

本発明の 1 つの実施形態では、該対象の損傷からの組織サンプルの F A N C D 2 遺伝子の配列またはその m R N A 配列を野生型 F A N C D 2 遺伝子の配列またはその m R N A 配列と比較することからなる、ヒト対象からの腫瘍性損傷における F A N C D 2 変異の欠如を確認するための方法が提供され、ここで組織サンプルでの野生型配列の存在は F A N C D 2 遺伝子での変異の欠如を示している。

#### 【 0 0 2 6 】

本発明の 1 つの実施形態では、癌を有する対象のための治療用プロトコルを決定するための方法が提供され、これは ( a ) 特異的抗体を用いて F A N C D 2 アイソフォームを測定することにより対象からの細胞サンプルで F A N C D 2 - L の欠損が生じるかどうかを

決定すること、(b)(a)で欠損が検出された場合、次いで欠損が非癌細胞における遺伝的欠損に結果であるかどうかを決定すること；及び(c)(b)が陽性であった場合、DNA損傷の増加を引き起こす治療用プロトコルの使用を減らし、対象における正常な組織を保護し、(b)が陰性で、欠損が癌細胞における遺伝的欠損に含まれる場合、DNA損傷の増加を引き起こす治療用プロトコルの使用を増やし、癌細胞に悪影響を与えること；が含まれる。

【0027】

本発明の1つの実施形態では、細胞標的におけるFA経路欠損を処置する方法が提供され、この方法には；有効量のFANCD2タンパク質または外因性の核酸を標的に投与することが含まれる。FA経路欠損は欠損FANCD2遺伝子でよく、及び外因性核酸ベクターはさらに前記に従ってベクターを導入することを含んでよい。ベクターを変異ヘルペスウイルス、E1/E4欠失組換えアデノウイルス、変異レトロウイルス、感染性新規ウイルス粒子の生成に必須であるウイルス遺伝子に関して欠損性であるウイルスベクターから選択できる。ベクターは脂質ミセルに含まれてよい。

10

【0028】

本発明の1つの実施形態では、欠損FANCD2遺伝子を有する患者を処置する方法が提供され、その方法には欠損FANCD2遺伝子から生じる欠損に機能的に修正するための、配列番号：4に記載されたポリペプチドを提供することが含まれる。

【0029】

本発明の1つの実施形態では、FA経路欠損を検出するための細胞基盤のアッセイが提供され、その方法には対象から細胞サンプルを入手すること；細胞をDNA損傷物質に暴露すること；及びFANCD2-Lが上方制御されているかどうかを検出することが含まれ、上方制御の不在はFA経路欠損を示している。細胞基盤のアッセイでは、核フォーカスを検出するためのイムノプロットング；FANCD2アイソフォームの量を検出するためのウェスタンプロット及びDNAプローブでのハイブリダイズすることによるmRNAの定量化から選択される分析技術によりFANCD2の量を測定することができる。

20

本発明の1つの実施形態では、生物学的サンプルにおいて癌細胞を検出するのに用いられるキットが提供され、そのキットは(a)高ストリンジェント条件下でFANCD2遺伝子の配列に結合するプライマー対であって、プライマー対は表7に記載される、変化した核酸配列を特異的に増幅するように選択される；及び各プライマーのための容器を含む。

30

【図面の簡単な説明】

【0030】

本発明の前記の特徴は、添付の図面を参考にして以下の詳細な記載を参照することにより容易に理解されよう。ここで；

【図1A】図1Aはファンコーニ貧血タンパク質複合体がFANCD2のモノユビキチン化に必要であることを示すウェスタンプロットを提供する。正常な(WT)細胞(レーン1)はFANCD2タンパク質の2つのアイソフォームである低分子量アイソフォーム(FANCD2-S)(155kD)及び高分子量アイソフォーム(FANCD2-L)(162kD)を発現する。レーン3、7、9、11はFA細胞系がFANCD2-Sアイソフォームのみを発現するA、C、G及びF型患者に由来することを示している。レーン4、8、10、12は対応するFA cDNAを有する細胞系のトランスフェクションの後の高分子量アイソフォームFANCD2-Lの回復を示している。

40

【図1B】図1Bは、HeLa細胞をHA-ユビキチンをコードするcDNAでトランスフェクトした後に得られたウェスタンプロットを示している。トランスフェクションの後、指示された用量のマイトマイシンC(MMC)で細胞を処置した。指示されるようにFANCD2に対するポリクローナル抗体(E35)で細胞性タンパク質を免疫沈降させた。FANCD2を免疫沈降させ、免疫複合体を抗FANCD2または抗HAMONOクローナル抗体でプロットした。

【図1C】図1Cは、HeLa細胞をHA-ユビキチンをコードするcDNAでトランス

50

フェクトした後に得られたウェスタンブロットを示している。トランスフェクションの後、指示された用量の電磁放射線（IR）で細胞を処置した。FANCD2を免疫沈降させ、免疫複合体を抗FANCD2または抗HAモノクローナル抗体でブロットした。

【図1D】図1DはFA-G繊維芽細胞系（FAG326SV）または修正された細胞（FAG326SVプラスFANCG cDNA）をHA-Ub cDNAでトランスフェクトし、FANCD2を免疫沈降し、免疫複合体を抗FANCD2または抗HA抗血清でブロットした後に得られたウェスタンブロットを示す。

【図1E】図1EはHeLa細胞を1mMヒドロキシウレアで24時間処置した後に得られたウェスタンブロットを示す。HeLa細胞ライゼートを抽出し、指示された温度で指示された時間、2.5μMユビキチンアルデヒドを伴って、または伴わずにインキュベートした。モノクローナル抗FANCD2（FI17）でのイムノブロットによりFANCD2タンパク質を検出した。

【図2】図2はファンコーニ貧血経路がFANCD2核フォーカスの形成に必要であることを示している。上部パネルは全細胞抽出物として調製された、SV40形質転換された繊維芽細胞の抗FANCD2イムノブロットを示す。パネルaからhはアフィニティー精製された抗FANCD2抗血清での免疫蛍光を示す。未修正（変異体、M）FA繊維芽細胞はFA-A（GM6914）、FA-G（FAG326SV）、FA-C（PD426）、及びFA-D（PD20F）であった。FA-A、FA-G及びFA-C繊維芽細胞を対応するFA cDNAで機能的に補足した。FA-D細胞をネオマイシンタグ化ヒト染色体3pで補足した（Whitneyら、1995）。

【図3A-E】図3はFANCD2タンパク質の2つのアイソフォームの細胞周期依存性発現を示す。（a）HeLa細胞、FA-A患者からのSV40形質転換された繊維芽細胞（GM6914）、及びFANCA cDNAで修正したGM6914細胞を二重チミン遮断法により同調させた。細胞周期の指示された相に対応する細胞を溶解し、FANCD2イムノブロットングのために処理した。（b）ノコダゾール遮断による同調（c）ミモシン遮断による同調（d）ノコダゾールまたは（e）ミモシンを用いてHeLa細胞を細胞周期で同調させ、細胞周期の指示された相に対応する細胞を抗FANCD2抗体で免疫染色し、免疫蛍光により分析した。

【図4A-E】図4はMMC、電離放射線、または紫外線に対する細胞性暴露の後の活性化されたFANCD2核フォーカスの形成を示す。指数関数的に成長したHeLa細胞は未処置であるかまたは指示されたDNA損傷物質、（a）マイトマイシンC（MMC）、（b）線放射（IR）、または（c）紫外線（UV）に暴露し、FANCD2イムノブロットングまたはFANCD2免疫染色のために処理した。（a）細胞を指示するように40ng/ml MMCに0から72時間連続的に暴露するか、または24時間処置し、免疫蛍光のために固定した。（b）及び（c）細胞を線放射（10Gy、B）またはUV光（60J/m<sup>2</sup>）に暴露し、指示された時間の後収集する（上部パネル）かまたは指示された用量で照射し、1時間後に収穫した（下部パネル）。免疫蛍光分析のために、細胞を処置後8時間に固定した（B、10Gy、C、60J/m<sup>2</sup>）。（d）正常な固体（PD7）または種々のファンコーニ貧血患者からの、指示されたEBV形質転換されたリンパ芽球系を40ng/mlマイトマイシンCで連続的に処置するか（レーン1から21）、または15Gyの線放射（レーン22から33）に暴露し、FANCD2イムノブロットングのために処理した。MMCまたはIR処置後のFANCD-Lの上方制御がPD7及び修正されたFA-A細胞（レーン28から33）で認められたが、変異体ファンコーニ貧血細胞系のいずれにおいても観察されなかった。同様に、IR誘起のFANCD2核フォーカスはFA繊維芽細胞（FA-G+IR）では検出されなかったが、機能的補足（FA-G+FANCG）の後には回復した。

【図5A-B】図5はDNA損傷後の散在性核フォーカスにおける活性化FANCD2及びBRCA1の同時局在化を示している。HeLa細胞は未処置であるかまたは指示されるように電離放射線（10Gy）に暴露し、8時間後に固定した。（a）細胞をD-9モノクローナル抗BRCA1抗体（緑色、パネルa、d、g、h）及びウサギポリクローナ

10

20

30

40

50

ル抗FANCD2抗体(赤色、パネルb、e、h、k)で二重染色し、染色した細胞を免疫蛍光で分析した。緑色及び赤色のシグナルが重複する場合(融合、パネルc、f、i、l)、黄色のパターンが認められ、これはBRCA1及びFANCD2の同時局在化を示している。(b)FANCD2及びBRCA1の同時免疫沈降。HeLa細胞は未処置(-IR)かまたは15Gyの線放射(+IR)に暴露し、12時間後に収集した。細胞ライゼートを調製し、モノクローナルFANCD2抗体(FI-17、レーン9から10)か、またはヒトBRCA1に対して別個に誘導された3つのモノクローナル抗体(レーン3から8):D-9(サンタ・クルーズ)、Ab-1及びAb-3(オンコジーン・リサーチ・プロダクツ)のいずれか1つのいずれかで細胞タンパク質を免疫沈降した。同量の精製されたマウスIgG(シグマ)を対照サンプルに用いた(レーン1から2)。免疫複合体をSDS-PAGEにより分析し、抗FANCD2または抗BRCA1抗血清でイムノプロットした。FANCD-LアイソフォームはBRCA1で優先的に同時免疫沈降された。

10

【図6】図6はS期の間の別個の核フォーカスでの活性化されたFANCD2及びBRCA1の同時局在を示す。(a)HeLa細胞をミモシンでG1後期に同調させ、S期に開放された。S期細胞をモノクローナル抗BRCA1抗体(緑色、パネルa、d)及びウサギポリクローナル抗FANCD2抗体(赤色、パネルb、e)で二重染色し、染色した細胞を免疫蛍光により分析した。緑色及び赤色のシグナルが重複する場合(融合、パネルc、f)、黄色のパターンが認められ、これはBRCA1及びFANCD2の同時局在化を示している。(b)S期で同調されたHeLa細胞を指示されたように、未処置(a、b、k、l)またはIRに暴露(50Gy、パネルc、d、m、n)、MMC(20µg/ml、パネルe、f、o、p)、またはUV(100j/m<sup>2</sup>、パネルg、h、q、r)に暴露し、1時間後に固定した。続いて細胞をFANCD2またはBRCA1に特異的な抗体で免疫染色した。

20

【図7】図7はマウス精母細胞の減数分裂の間にBRCA1で同時局在化できる対合複合体でFANCD2がフォーカスを形成することを示している。(a)後期パキテン期マウス核における対合複合体の抗SCP-3(白色)及び抗FANCD2(赤色)染色。(b)後期パキテン期染色体のSPC3染色。(c)この染色を抗FANCD2 E35抗体に関する免疫前血清で広げたもの。(d)マウスディプロテン期核の対合複合体の抗SCP3染色。(e)この同時染色をE35抗FANCD2抗体で広げたもの。対になっていない性染色体及び常染色体のテロメアの双方を抗FANCD2で染色することに注目すべきである。(f)この染色を抗BRCA1抗体で広げたもの。性染色体は優先的に染色される。(g)後期パキテン期性染色体対合複合体の抗FANCD2染色。(h)同一の複合構造の抗BRCA1染色。(i)抗FANCD2(赤色)及び抗BRCA1(緑色)同時染色(黄色の部分により反映される同時局在化)。

30

【図8】図8は細胞経路におけるFAタンパク質の相互作用の概略図を提供する。FAタンパク質(A、C及びG)は機能的核複合体で結合する。この複合体の活性化時にS期への参加またはDNA損傷のいずれかにより、この複合体がDタンパク質を酵素的に修飾(モノユビキチン化)する。このモデルに従って、活性化されたDタンパク質が続いて核フォーカスに標的化され、そこでこれがBRCA1タンパク質及びDNA修復に關与するその他のタンパク質と相互作用する。

40

【図9】図9はFANCD2 cDNA全長でプローブし、24時間暴露したヒト成人の心臓、脳、胎盤、肝臓、骨格筋、腎臓、膵臓、脾臓、胸腺、前立腺、精巣、子宮、小腸、結腸及び末梢血リンパ球、並びにヒト胎児の脳、配列、肝臓及び腎臓からの細胞のノーザンプロットを示す。

【図10A-F】図10は2つのFANCD2ファミリーの変異分析のためのアレル特異的アッセイを示し、ここでファミリー系統図(a、d)並びにパネルb、c、e及びfを垂直に並べ、対応する変異が問題の個体の下になるようにする。パネルaからcはPD20を表し、パネルdからfはVU008ファミリーを表す。パネルb及びeは、新規MspI部位(PD20)またはDdeI部位(VU008)の作製により検出される母性変

50

異の分離を示す。双方のファミリーで父性遺伝される変異はアレル特異的オリゴヌクレオチドハイブリダイゼーションにより検出された（パネルc及びf）。

【図11】図11はヒトファンコーニ貧血細胞系におけるFANCD2タンパク質のウェスタンブロット分析を示す。指示された繊維芽細胞及びリンパ芽球系から全細胞ライゼートを作製した。抗FANCD2抗血清を用いるイムノブロッティングによりタンパク質ライゼート（70 µg）をプローブした。FANCD2タンパク質（155 kD及び162 kD）は矢印で示す。イムノブロットのその他のバンドは非特異的である。（a）試験した細胞系には野生型細胞（レーン1、7）、PD20、繊維芽細胞（レーン2）、PD20リンパ芽球（レーン4）、復帰変異体MMC抵抗性PD20リンパ芽球（レーン5、6）及びPD20繊維芽細胞で補足された染色体3p（レーン3）などがある。HSC62（レーン8）及びVU008（レーン9）などのいくつかのその他のFAのD群細胞系が分析された。FA-A細胞はHSC72（レーン10）、FA-C細胞はPD4（レーン11）及びFA-G細胞はEUF316（レーン12）であった。（b）第3のFANCD2患者の同定。FANCD2タンパク質は野生型及びFA群G細胞において容易に検出できたが、PD733細胞では検出されなかった。（c）抗体の特異性。レトロウイルスFANCD2発現ベクターで形質導入されたPD20i細胞は、空ベクター対照（レーン3）及びトランスフェクトされていないPD20i細胞（レーン2）と対照的に、FANCD2タンパク質の双方のアイソフォームを展示する（レーン4）。野生型細胞では、内因性FANCD2タンパク質（2つのアイソフォーム）もまた抗体と免疫反応性であった（レーン1）。

10

20

【図12】図12はクローン化FANCD2 cDNAでのFA-D2細胞の機能的補足を示す。SV40形質転換したFA-D2繊維芽細胞系であるPD20iをpMMP-puro（PD20+ベクター）またはpMMP-FANCD2（PD20+FANCD2wt）で形質導入した。ピューロマイシン選択細胞をMMC感受性分析に供した。分析した細胞は親PD20F細胞（○）、ヒト染色体3pで修正したPD20（○）、及びpMMP-puro（○）またはpMMP-FANCD2（wt）-puro（○）で形質導入したPD20細胞であった。

【図13】図13はPD20リンパ芽球の復帰のための分子基盤を示す。（a）エクソン5または6に対するPCRプライマーを用いてcDNAを増幅した。対照サンプル（右レーン）は114 bpで単一のバンドを生じたが、一方PD20 cDNA（左レーン）は2つのバンドを示し、大きい方は13 bpのイントロン配列の母性アレルへの挿入を反映している。PD20から復帰したMMC抵抗性リンパ芽球（中央レーン）は第3の114 + 36 bpのフレーム内スプライシング変種を示している。（b）FANCD2エクソン5/イントロン5境界でのスプライシングの該略図。野生型cDNAでは100%のスプライシング事象が適切なエクソン/イントロン境界で生じており、一方母性A->G変異（矢印で示す）はこれもまた100%でスプライシング異常を招く。復帰した細胞では母性変異を有する全てのcDNAはまた第2の配列変化をも有し、13 bp（~40%のmRNA）または36 bp（~60%のmRNA）のいずれかの挿入との混ざったスプライシングパターンを示した。

30

【図14】図14は卵巣癌の患者に由来する癌細胞系のFANCD2ウェスタンブロットを示す。

40

【図15】図15はヒトFANCD2のアミノ酸配列及びBEAUTYアルゴリズムを用いたハエ及び植物相同体とのアラインメントに関する配列表を示す（Worleyら、Genome Res. 5:173-184（1995））（配列番号：1から3）。黒色四角はアミノ酸同一性を示し、灰色は類似性を示す。最良のアラインメントスコアはキイロシヨウジヨウバエ（D.melanogaster）（ $p = 8.4 \times 10^{-58}$ 、受け入れ番号AAF55806）及びシロイヌナズナ（A.thaliana）（ $p = 9.4 \times 10^{-45}$ 、受け入れ番号B71413）での仮説タンパク質で観察された。

【図16】図16はFANCD2 cDNA配列（63から5127ヌクレオチド）（配列番号：5）及びこの配列によりコードされるポリペプチド（アミノ酸1から1472）（

50

配列番号：４）である。

【図１７】図１７はFANCD cDNA（配列番号：１８８）と比較したFANCD - S . ORF（配列番号：１８７）に関するヌクレオチド配列である。

【図１８】図１８はヒトFANCD 2 - Lに関するヌクレオチド配列である（配列番号：６）。

【図１９】図１９はヒトFANCD 2 - Sに関するヌクレオチド配列である（配列番号：７）。

【図２０】図２０はヒトFANCD 2に関するヌクレオチド配列である（配列番号：８）。

【発明を実施するための形態】

10

【００３１】

定義：

この記載及び添付の請求の範囲で用いる以下の用語は、前後関係から別に必要とされる以外は指示する意味を有する：

「FANCD 2 - L治療薬」はいずれかのタンパク質異性体を意味し、FANCD 2 - Lタンパク質のペプチド、ペプチド誘導體、類似体または異性体を含み、さらにFANCD 2 - Lタンパク質に機能的に等価であるいずれかの小型分子誘導體、類似体、異性体またはアゴニストを含む。定義には全長もしくは部分長遺伝子配列またはcDNAでよいが、または核酸を活性化する遺伝子もしくは遺伝子配列を変調するように作用できるアンチセンス分子のアプタマー（aptamer）などの分子に結合する核酸でよい、FANCD 2をコードする核酸も含まれる。

20

【００３２】

「FANCD - 2をコードする核酸」には、前記で定義したFANCD 2 - Lタンパク質を発現するためのFANCD 2の完全なcDNAもしくはゲノム配列またはその部分が含まれる。核酸はさらに核酸キャリアまたはベクターに含まれてよく、標的部位に有効に分配するために適当に修飾されている核酸を含む。

【００３３】

「ハイブリダイゼーションのストリンジェント条件」には一般に30 以上、典型的には37 以上、好ましくは45 以上の温度が含まれる。ストリンジェント塩条件は通常1000mM未満、典型的には500mM未満、好ましくは200mM未満である。

30

【００３４】

核酸に関する「実質的相同性または類似性」は、別の核酸（またはその相補鎖）と最適に整列化（適当な核酸挿入または欠失を伴う）した場合に少なくとも約60%、通常少なくとも約70%、より通常的には少なくとも約80%のヌクレオチド塩基でヌクレオチド配列同一性があれば、核酸またはそのフラグメントが別のものに「実質的に相同である」（「または実質的に類似している」）場合のことある。

【００３５】

「抗体」にはポリクローナル及びモノクローナル抗体及び、一本鎖抗体を含み、一本鎖抗体及びFabフラグメントを含むそのフラグメント、並びにタンパク質のアイソフォームを区別するのに十分な結合特異性を有するその免疫結合性等価物が含まれる。これらの抗体はアッセイ及び医薬品において有用である。

40

【００３６】

「単離された」とは、その天然の状態においてそれに付随する成分から分離されているタンパク質、ポリペプチドまたは核酸を記載するのに用いられる。サンプルの少なくとも約60%から75%が単一のアミノ酸またはヌクレオチド配列を呈する場合、「単離された」タンパク質または核酸は実質的には純粋である。

【００３７】

「制御配列」は通常遺伝子座のコード化領域の100kb以内のこれらの配列を意味するが、コード化領域からさらに遠位でもよく、遺伝子の発現（遺伝子の転写、及び翻訳、スプライシング、安定性またはメッセンジャーRNAのようなものを含む）に影響する。

50

## 【0038】

「ポリヌクレオチド」はRNA、cDNA、ゲノムDNA、合成体、及び混合ポリマー、センス及びアンチセンス鎖の双方を含み、技術分野の技術者に容易に理解されるように、化学的または生化学的に修飾されていてよく、または非天然または誘導ヌクレオチド塩基を含んでよい。かかる修飾には、例えば標識、メチル化、1つまたはそれ以上の天然発生ヌクレオチドの類似体との置換、ヌクレオチド内修飾、例えば非荷電性結合（例えばホスホン酸メチル、ホスホトリエステル、ホスホアミデート、カルバメート等）、荷電性結合（例えばホスホロチオエート、ホスホロジチオエート等）、懸垂基（例えばポリペプチド）及び修飾された結合（例えばアルファ・アノマー核酸）などがある。水素結合またはその他の化学的相互作用を介して指定された配列に結合する能力において核酸を擬似する合成分子もまた含まれる。

10

## 【0039】

「変異」は、野生型に比較して、制御配列の遺伝子内または遺伝子外のヌクレオチド配列における変化である。変化は欠失、置換、点変異、複数のヌクレオチドの変異、転位、逆位、フレームシフト、ナンセンス変異、または機能細胞において正常に発現された遺伝子のものから核酸またはタンパク質を区別する染色体異常の別の形態でよく、この場合発現及び機能性は正常に発生する範囲内である。

## 【0040】

「対象」とはヒトなどの哺乳動物を含む動物を意味する。

## 【0041】

「野性型FANCD2」とは細胞内でモノユビキチン化してFANCD-SからFANCD2-Lを形成することができるタンパク質または発現されたタンパク質をコードする遺伝子を意味する。

20

## 【0042】

いくつかのファンコーニ貧血は毛細血管拡張性運動失調(AT)、色素性乾皮症(XP)、コケーン症候群(CS)、ブルーム症候群、骨髄異形成症候群、再生不良性貧血、癌感受性症候群及びHNPCなどの症候群の群と類似性を有する(表2参照)。これらの症候群はDNA修復に基礎的欠損を有し、染色体の完全性の維持における欠損に関連する。DNA修復及び染色体の完全性の維持に関連する経路における欠損の結果、ゲノム不安定性、及びDNA損傷物質、例えば鎖間架橋を引き起こす2価性アルキル化剤に対する細胞感受性を招く。さらに、DNA修復メカニズムにおける障害は、遺伝子再配列により癌の域を誘発する可能性を実質的に増大させるようだ。この観察はマイトマイシンC、シスプラチン、シクロホスファミド、ソラレン(pсорalen)及びUVA照射などのDNA架橋剤の臨床使用に関して直接関係する。

30

## 【0043】

ファンコーニ貧血は稀な疾患であるが、FAの多形質発現効果はゲノム安定性、アポトーシス、細胞サイクル調節及びDNA架橋に対する抵抗性などの多様な細胞性過程の経路におけるFAタンパク質の野生型機能の重要性を示している。

FAの細胞性異常には、架橋化剤に対する感受性、細胞周期G2期の延長、周囲O<sub>2</sub>下での成長低下などの酸素に対する感受性、O<sub>2</sub>ラジカルの過剰生成、O<sub>2</sub>ラジカル防御の欠如、スーパーオキシドジスムターゼの欠乏；電離放射線に対する感受性(G2特異的)；腫瘍壊死因子の過剰生成、DNA付加物の蓄積などのDNA修復の直接的な欠損、及びDNA架橋の修復における欠損、自発的染色体切断などのゲノム不安定性、及び欠失メカニズムによる過剰易変性、アポトーシスの増加、p53誘導の欠損、インビトロコロニー成長の低下などの内因性幹細胞欠損；生殖腺幹細胞生存性の低下などがある。

40

## 【0044】

これらの特徴は造血及び生殖腺幹細胞の維持、並びに骨格及び泌尿性器系などの多くの異なる構造での正常な胚発生におけるFAの関与を反映している。患者からの細胞サンプルを分析してFAの相補D群における欠損を決定した。1人の患者からのリンパ芽球はPD20細胞系を上昇させ、これはD相補群の欠損を有する別の1人の患者に由来するHS

50

C62からの異なる遺伝子において変異していることが解った。双方の患者からの変異をD相補群にマッピングされたが、異なる遺伝子にはマッピングされず、よってFANCDタンパク質・FANCD1(HSC62)及びFANCD2(PD20)と命名された(Timmersら、Molecular Cell 7:241-248(2001))。我々はFANCD2がFA経路の終点であり、FA核複合体の部分ではなく、その集合または安定性に必要でもなく、FANCD2は2つのアイソフォームであるFANCD2-S及びFANCD2-Lで存在することを示した。我々はまたタンパク質の短い形態(FANCD2-S)のタンパク質の長い形態(FANCD2-L)への形質転換がFA複合体に応答して生じることをも示した(図8)。FA経路に関連する特定のタンパク質の欠損の結果、FANCD2-Lとして同定される、FANCD2の重要な翻訳後修飾形態の作製に失敗する。FANCD2の2つのアイソフォームは短形態及び長形態として同定される。

10

## 【0045】

FANCD2-Lの作製の失敗はDNA修復におけるエラー及び前記で列挙した疾患に関連する細胞周期異常に相関する。

## 【0046】

前記した症候群におけるFANCD2の役割についてさらに理解するために、我々はFANCD2遺伝子をクローン化し、タンパク質配列を決定した。FANCD2遺伝子は4353塩基対のオープン・リーディング・フレーム及び、推定分子量166kDaの新規1451アミノ酸核タンパク質をコードする44個のエクソンを有する。ウェスタンブロット分析により162及び155kDの2つのタンパク質アイソフォームの存在が示された。44イントロン/エクソン接合部に対応する配列を表6(配列番号9から94)で提示する。

20

## 【0047】

以前にクローン化されたFAタンパク質とは異なって、いくつかの無脊椎動物真核細胞からのFANCD2タンパク質はキイロショウジョウバエ(D.melanogaster)、シロイヌナズナ(A.thaliana)及び線虫(C.elegans)のタンパク質と高度に有意なアラインメントスコアを示した。ショウジョウバエ相同体はFANCD2に対して28%のアミノ酸同一性及び50%の類似性を有し(図及び配列番号:1から3)、各々の種では機能実験は実施しなかった。大腸菌(E.coli)及びパン酵母(S.cerevisiae)ではFANCD2に類似するタンパク質は見出されなかった。

30

## 【0048】

我々は、FANCD2遺伝子においてピアレリック変異を有する2つのFA細胞系であるPD20及びVU008の染色体3p遺伝子座を分析することによりFANCD2DNA配列(配列番号:5)を得た(図10)。患者からのリンパ芽球はHSC62を補足し損なっていたので、細胞系をD群の参照細胞系であるD相補群に割り当てた。FANCD2変異はこのD群参照細胞系では検出されず、これはHSC62において変異した遺伝子がFANCD1をコードする遺伝子であり、PD20及びVU008ではFANCD2であることを示している(図11)。微小核体媒介染色体移動を用いて変異を同定した(Whitneyら、Blood 88:49-58(1995))。遺伝子座を包含する小さな重複欠失を含有する5つの微小核体ハイブリッドの詳細な分析により、FANCD2遺伝子の候補領域が200kbに狭められた。FANCD2遺伝子は以下の様に単離された:3つの候補ESTをこのFANCD2臨界領域内または近くに局在させた。cDNA全長を得るために、5'及び3'RACEを用いて遺伝子をシーケンシングし、各々の発現パターンをノーザンブロットにより分析した。ESTSGC34603は以前にクローン化されたFA遺伝子に類似して偏在性であり、5kb及び7kbmRNAの発現は低レベルであった。TIGER-A004X28、AA609512及びSGC34603でオープン・リーディング・フレームが見出され、各々の長さは234、531及び4413bpであった。クローン化RT-PCR生成物をシーケンシングすることよりPD

40

50

20細胞における変異に関して3つ全てを分析した。TIGER-A004X28及びA609512では配列変化は検出されなかったが、SGC34603では5つの配列変化が見出された。次に、我々は臨界領域からのBAC177N7におけるcDNAシーケンシングプライマーを用いてSGC34603遺伝子の構造を決定した。

#### 【0049】

ゲノム配列情報に基づいてPCRプライマー対を設計し(表7)、推定される変異を含有するエクソンを増幅し、アレルト異のアッセイを發展させてPD20ファミリー及び568対照染色体をスクリーニングした。アレルの3つが共通の多型性であった;しかしながら、対照において2つの変化が見出されず、従って、変異の可能性が示された(表3)。第1は母性遺伝したnt376におけるA->G変化であった。アミノ酸の変化(S126G)に加えて、この変異はスプライシングミス及びイントロン5からmRNAへの13bpの挿入に随伴された。母性変異を有する43/43(100%)の別個にクローン化されたRT-PCR生成物はこの挿入を有し、一方対照cDNAクローンの3%(1/31)のみでmRNAのスプライシングミスが示された。13bpの挿入によりフレームシフトを生じ、たった180個のアミノ酸長の重度にトランケートされたタンパク質が予測される。第2の変異は1236位置で父性遺伝されたミスセンス変化であった(R1236H)。PD20コアファミリーにおける変異の分離を図10に示す。PD20のSGC34603遺伝子は568対照染色体には存在しない母性及び父性アレルの双方を含有し、母性変異は分析したcDNAの100%でスプライシングミスに随伴されたので、我々はSGC34603がFANCD2遺伝子であると結論した。

10

20

#### 【0050】

FANCD2にコードされるタンパク質はPD20には存在しない:SGC34603のFANCD2としての同定をさらに確実にするために、タンパク質に対して抗体を上昇させ、ウェスタンブロット分析を実施した(図11)。レトロウイルス形質導入及びPD20細胞におけるFANCD2の安定発現により抗体の特異性が示された(図11c)。野生型細胞ではこの抗体は2つのバンド(155及び162kD)で検出され、我々はFANCD2-S及びLと称する(図11cで最もよく認められる)。FANCD2タンパク質レベルは患者のPD20からの全てのMMC感受性細胞系で著明に減少したが(図11a、レーン2、4)、野生型細胞系及び別の相補群からのFA細胞には存在した。さらに、染色体3の微小核体媒介移動により修正されたPD20細胞はまた正常量のタンパク質を生成した(図11a、レーン3)。

30

#### 【0051】

FANCD2 cDNAでのFA-D2細胞の機能的補足:我々は次にクローン化されたFANCD2 cDNAがFA-D2細胞のMMC感受性を補足する能力を評価した(図12)。FANCD2 cDNA全長を前記したレトロウイルス発現ベクターであるpMMP-puroにサブクローニングした(Mol. Med. 4:468-479(1998))。形質導入されたPD-20細胞はFANCD2タンパク質のアイソフォームであるFANCD2-S及びFANCD2-Lの双方を発現した(図12c)。pMMP-FANCD2によるFA-D2(PD20)細胞の形質導入は細胞のMMC感受性を修正した。これらの結果はさらにクローン化されたFANCD2 cDNAがFANCD2-Sタンパク質をコードし、これをFANCD2-Lアイソフォームに翻訳後修飾することができる。この重要な修飾は以下で詳細に論じる。

40

#### 【0052】

表現型的に復帰したPD20クローンの分析:我々は次に、PD20細胞における配列変化は機能的に中立である多型性ではないことを示すさらなる証拠を作製した。この目的のために、我々はもはやMMCに感受性を示さない患者PD20からの復帰変異リンパ芽球クローン(PD20-c1.1)の分子分析を実施した。復帰及び体細胞モザイク現象はFAにおいてしばしば見出され、遺伝子内事象、例えば有糸分裂組換えまたは代償性フレームシフトに随伴された。実際に、母性由来するSGC34603 cDNAの~60%が、通常観察される13bpよりむしろイントロン5配列の36bpを挿入する新規ス

50

プライシング変種を有した(図13)。IVS5+6でGからAへのデノボ変化に相関するフレーム内プライシング変種の出現(図13)及び正確な読み枠の回復がウェスタンブロット分析により確認された。患者PD20からの全てのMMC感受性繊維芽細胞及びリンパ芽球であるPD20-c1.1は容易に検出できる量の、正常タンパク質よりもわずかに高分子量のFANCD2タンパク質を生成した。

#### 【0053】

別の「FANCD」患者からの細胞系の分析：また抗体を用いて、FAのD群の参照細胞系、HSC及び欧州ファンコーニ貧血登録所(EUFAR)によりD群として同定された2つの別の細胞系などのさらなるFA患者細胞系をスクリーニングした。VU008はFANCD2タンパク質を発現せず、双方共にエクソン12でミスセンス及びナンセンス変異を有する異型接合性化合物であることが見出され、370対照染色体では見出されなかった(表3、図11)。VU008細胞からのライゼートではFANCD2タンパク質を検出できないので、ミスセンス変異はFANCD2タンパク質を不安定化するようである。第3の患者PD733もまたFANCD2タンパク質を欠き(図11b、レーン3)、エクソン17の不在に至るプライシング変異及びタンパク質の内部欠失が見出された。変異とこれらの患者に由来する細胞ライゼート中のFANCD2タンパク質不在との相関は、FANCD2のFA遺伝子との同一性を実証する。対照的に、容易に検出できる量のFANCD2タンパク質の双方のアイソフォームがHSC62(図11a、レーン8)及びVU423 cDNAで見出され、双方の細胞系からのゲノムDNAが変異に関して詳細に分析され、何も見出されなかった。加えて、VU423及びPD20繊維芽細胞間の全細胞融合が染色体切断表現型の補足を示した(表5)。これらのデータを合わせて考慮すると、FAのD群は遺伝的に異種性であり、HSC62及びVU423において欠損している(複数の)遺伝子がFANCD2とは区別されることが示される。

#### 【0054】

FANCD2遺伝子及びタンパク質の同定及びシーケンシングにより、非限定例としては表2に列挙するものなどのDNA修復の失敗及び細胞周期異常に関連する疾患の治療の開発、診断試験及びスクリーニングアッセイのための新規目標が提供される。

#### 【0055】

以下の記載によりFA経路におけるFANCD2の生物学的役割に新規で有用な洞察が提供され、前記の症候群の診断及び処置のための基礎が提供される。

#### 【0056】

FA細胞が細胞周期制御における基本的な分子性欠損を有するという証拠には以下のものなどがある：(a)FA細胞は化学的架橋化剤での処置により増強される4N DNA含量を伴う細胞周期遅延を示し、(b)FA細胞の細胞周期停止及び増殖低下はタンパク質のサイクリンファミリーのメンバーであるタンパク質、SPHARの過剰発現により部分的に修正されることができ、並びに(c)カフェインはFA細胞のG2停止を止めさせる。これらの結果に合致して、カフェインは構造的にcdc2を活性化し、FA細胞の正常なG2細胞周期チェックポイントを覆うことができる。最終的にFANCCタンパク質はサイクリン依存性キナーゼであるcdc2に結合する。我々はFA複合体がサイクリンB/cdc2複合体の基質またはモデュレーターであり得ることを提示している。

#### 【0057】

加えて、FA細胞がDNA修復において基礎的な欠損を有しているという証拠は、(a)架橋DNAまたは二本鎖切断の修復における特異的欠損を示唆する、DNA架橋化剤及び電離放射線(IR)に感受性のあるFA細胞、(b)修復欠損でしかも無傷のチェックポイント活性の存在を示唆する、過剰反応性p53応答に至るFA細胞のDNA損傷、及び(c)非相同性末端結合の適合性の欠損及び相同組換え率の増加を伴うFA細胞により示唆される(Garcia-Higueraら、Mol. Cell. 7: 249-262(2001))、(Grompeら、Hum. Mol. Genet. 10: 1-7(2001))。

#### 【0058】

10

20

30

40

50

これらの細胞周期及びDNA修復の一般的な異常にかかわらず、FA経路がこれらの活性を制御するメカニズムは解りにくい。我々はここでFANCD2タンパク質がFAタンパク質複合体の下流で機能することを示す。集合したFAタンパク質複合体の存在下で、FANCD2タンパク質は活性化されて高分子量のモノユビキチン化アイソフォームになり、これはS期特異的DNA修復応答を変調するようである。活性化されたFANCD2タンパク質はDNA損傷物質に反応して核フォーカスに蓄積し、公知のDNA修復タンパク質であるBRCA1と同時局在化及び同時免疫沈降する。これらの結果は以前矛盾していたFA経路(D'Andreaら、1997)のモデルを解決し、FAタンパク質がDNA損傷に対する細胞性応答と協働することを示している。

#### 【0059】

FA経路にはFAマルチサブユニット核複合体の形成が含まれ、A/C/Gに加えて我々は複合体のサブユニットとしてFANCFもまた含まれることを示した(図8)。FA経路はS期の間に「活性化」し、S期特異的修復応答またはチェックポイント応答を提供する。FAマルチサブユニット複合体に依存するFA経路の正常な活性化の結果、FANCD2-Lとして同定される高分子量活性化アイソフォームへのリン酸化工程を介するホスホタンパク質FANCD2のモノユビキチン化の制御に至る(図1)。モノユビキチン化は細胞輸送に関連する。FANCD2-LはS期特異的DNA修復応答を変調するようである(図3)。FA細胞がFANCD2のS期特異的活性化を活性化し損なうことは細胞周期特異的異常に関連する。活性化されたFANCD2タンパク質はDNA損傷物質、MMC及びIRに反応して核フォーカスに蓄積し、公知のDNA修復タンパク質であるBRCA1と同時局在化及び同時免疫沈降する(図4から6)。これらの結果は以前矛盾していたFA経路のモデルを解決し(D'Andreaら、1997)、DNA修復におけるFA経路の役割を強く支持する。

#### 【0060】

我々はFA経路に関するFANCD2アイソフォーム間の会合及び種々の癌の公知の診断用分子であるタンパク質を初めて同定した。S期においてまたはDNA損傷に反応して翻訳後修飾された高分子量のアイソフォームに活性化されるBRCA1タンパク質の、DNA損傷に関する類似の経路は、活性化されたFANCD2タンパク質がBRCA1と相互作用することを示唆している。さらにとりわけ、FANCD2のモノユビキチン化の制御は、FANCD2タンパク質を、BRCA1を含有する核フォーカスに標的化するのである。FANCD2はBRCA1と同時免疫沈降し、別の「ドット」タンパク質、例えばRAD50、Mre11、NBS、またはRAD51とさらに結合できる。最近の研究ではBRCA1フォーカスが大きな(2メガダルトン)のマルチタンパク質複合体から構成されることが示されている(Wangら、Genes Dev. 14:927-939(2000))。この複合体にはATM、DNA修復機能に関与するATM基質(BRCA1)、及びチェックポイント機能に関与するATM基質(NBS)などがある。さらに損傷認識及びFA経路活性化はATM、ATR、CHK1またはCHK2などのDNA損傷に反応するキナーゼに関与することが示唆されている。

#### 【0061】

我々はDNA損傷物質、IR及びMMCがFANCD2の独立した翻訳後修飾を活性化し、異なる機能的結果に至ることを見出した。IRはセリン222でATM依存的にFANCD2のリン酸化を活性化し、S期チェックポイント応答に至る。MMCはリジン561でBRCA1依存的及びFA経路依存的にFANCD2のモノユビキチン化を活性化し、FANCD2/BRCA1核フォーカスの集合及びMMC抵抗に至る。従って、FANCD2は、共通の経路における2つの別の癌感受性遺伝子(ATM及びBRCA1)間の連結を提供する2つの異なる翻訳後修飾の結果、染色体安定性の維持において2つの独立した機能的役割を有する。いくつかの別の系の証拠によりFANCD2及びBRCA1間の相互作用が支持される。第1に、BRCA1(-/-)細胞系であるNCC1937(Scullyら、Mol. Cell. 4:1093-1099(1999))は、染色体不安定で、トリラジカル及びテトララジカル染色体形成の増加を伴う「ファンコーニ

10

20

30

40

50

貧血様」表現型を有する。第2にBRCA1フォーカス(及びRAD51フォーカス)からのFA細胞は通常IRに応答するが、BRCA1(-/-)細胞は検出可能なBRCA1フォーカスを有さず、正常細胞に比較してFANCD2フォーカスの数が非常に増加している。BRCA1(-/-)細胞の機能的補足によりBRCA1フォーカス及びFANCD2フォーカスが正常レベルに回復し、正常なMMC抵抗性が回復した。

#### 【0062】

FANCD2-Lの量は一部fancd2遺伝子から合成されるFANCD2-Sの量により決定され、一部はFANCD2-Lを形成するモノユビキチン化FANCD2-Sに対するFA複合体の利用性により決定される。BRCA及びATMを含む核フォーカスとのFANCD2-Lの会合及びDNA修復におけるFANCD2-Lの役割の決定により、このタンパク質を広範な癌の患者における癌進行の可能性を観察するための強力な標的にする。かかる癌には染色体3pにおける損傷から生じるもの、及び変異が結果的にFA経路の上流のメンバー、例えばFANCG、FANCCまたはFANCAの生成を干渉するような別の染色体上の癌などがある。癌系及びバイオプシなどの癌患者からの1次細胞をFANCD-Lに関してスクリーニングし、このタンパク質のレベル異常を予測して疾患の早期診断に相関させる。FANCD2タンパク質はDNA修復への経路の最終工程であるので、FANCD2-SのFANCD-Lへの変換に至る1つまたはそれ以上の経路におけるタンパク質のいずれかの異常はFANCD2のレベルを測定することにより容易に検出されると考えられる。さらに、FANCD2のレベルは別のタンパク質、例えば核内のBRCA及びATMがどのように機能的に患者の結果と相互作用するかに影響する。患者のFANCD2レベルの分析は、患者に提示される癌のクラスに関する臨床決定において医師を補助すると予測される。例えば、癌細胞がモノユビキチン化FANCD2-Lアイソフォームの作製に失敗する場合、細胞は染色体不安定性を増し、恐らく照射または化学療法剤に対する感受性を増すであろう。この情報は患者の処置を改善する方法において医師を補助するであろう。

10

20

#### 【0063】

ファンコーニ貧血は異なる癌の広範なスペクトルのみならず、先天的異常にも関連する。胎児の発達は複雑ではあるが秩序だった過程である。特定のタンパク質は、この秩序だった発達の進行を破壊させるので、とりわけ広いスペクトルの影響を有する。FA経路は発達において重要な役割を果たし、FA経路の破壊の結果、多数の悪影響に至る。FA経路におけるエラーは胎児細胞からのFANCD2-Lタンパク質を分析することにより検出できる。FANCD2は正常な胎児発達のための診断マーカー及び治療的介入に可能な標的に相当する。

30

#### 【0064】

前記に一致して、我々は生存精子の生成においてFANCD2が役割を果たすことを示した。FANCD2は後期パキテン及びディプロテン期のネズミ精母細胞においてXY2価染色体の不对軸でフォーカスを形成する。(図7)興味深いことに、FANCD2フォーカスはまたディプロネマの常染色体テロメアでも認められる。FA患者及びFA-Cノックアウトマウスにおいて知られている生殖障害を考慮に入れて、我々の観察により、減数分裂Iを経る精母細胞の正常な発達に、活性化されたFANCD2タンパク質が必要であることが示唆される。XY軸で認められるたいていのFANCD2フォーカスはBRCA1フォーカスと同時局在していることが見出され、これは2つのタンパク質が減数分裂細胞で一緒に機能し得ることを示唆している。BRCA1と同様、FANCD2は発達中の対合複合体の軸(対合していない)エレメントにおいて検出された。組換えは対合している領域で生じるので、FANCD2は組換えが開始される前に機能して、恐らく染色体がシナプシスの準備をするのを助けるかまたは続く組換え事象を制御することができる。FANCD2は減数分裂染色体において集合し、有糸分裂細胞でドット構造を形成し、これは有糸分裂及び減数分裂の双方の細胞周期調節におけるFANCD2の役割を示唆している。

40

#### 【0065】

50

本発明の実施形態は翻訳後修飾されたアイソフォーム：FANCD-2LのFA経路の完全性を決定するための診断標的として使用することを志向する。FANCD2のユビキチン化及びFANCD2核フォーカスの形成はFA経路の下流の事象であり、いくつかのFA遺伝子の機能を必要とする。我々はいずれかの上流FA遺伝子（FANCA、FANCB、FANCC、FANCE、FANCF及びFANG）のピアリック変異が、ユビキチン化されていないFANCD2（FANCD2-S）形態からユビキチン化（FANCD2-L）へのFANCD2の翻訳後修飾を遮断することを見出した。これらの上流の欠損のいずれかはFANCD2 cDNAで細胞をトランスフェクトすることにより覆される。（図1a）

我々は初めてFANCD2の存在及びFA経路におけるその役割を示した。我々は、FANCD2がDNA損傷物質に応答して核フォーカスにおいて蓄積され、そこで別のDNA修復タンパク質、例えばBRCA1及びATMに関連することを示した。我々はFANCD2が細胞内で2つのアイソフォームで存在し、そこで2つのアイソフォームの1つの低下がファンコーニ貧血及び癌感受性の増大と関連することをも示した。我々はこれらの知見を用いて患者の看護を補助する臨床使用のための多くの診断試験を提供した。

#### 【0066】

これらの試験には：（a）将来の子孫または既存の妊娠における遺伝性ファンコーニ貧血を心配する親のための遺伝及び親のカウンセリング；（b）FA経路欠損に相関する癌に対する感受性の増大を決定するための成人のための遺伝カウンセリング及び免疫診断試験；並びに（c）対象にすでに既存の癌を診断して、副作用を最小限にし、対象に最大の効果を発揮する処置プロトコルを開発するための機会を提供すること；が含まれる。

#### 【0067】

本明細書に記載する診断試験は技術分野で公知の標準的なプロトコルに依存し、我々はその試験のためにFANCD2タンパク質及びヌクレオチド配列を試験するための新規試薬を提供した。これらの試薬はFANCD2アイソフォームに特異的な抗体、FANCD2遺伝子における遺伝子変化を検出するためのベクター、プローブ及びプライマーが誘導されたヌクレオチド配列、並びにFA経路における欠損を保護及び試験するための細胞系及び組換え細胞が含まれる。

#### 【0068】

我々は実施例1に記載するように、FANCD2-L及びFANCD2-Sタンパク質に特異的なモノクローナル及びポリクローナル抗体調製物を調製した。加えて、FANCD2アイソフォーム特異的抗体フラグメント及び一本鎖抗体を標準的技術を用いて調製できる。我々は、生物学的サンプルにおけるFANCD2アイソフォームを同定するために、湿式化学アッセイ、例えば免疫沈降アッセイ、例えばウェスタンブロットでこれらの抗体を使用した（図1）。酵素結合免疫吸着アッセイ（ELISA）、ラジオイムノアッセイ（RIA）、免疫放射定量アッセイ（IRMA）及び免疫酵素アッセイ（IEMA）を含み、さらにサンドウィッチアッセイなどを含む慣用されるイムノアッセイをも使用した。別のイムノアッセイは全細胞のサンプル、及び溶液中の抗体と反応し、場合によっては貯蔵容器内で液体状態で分析される溶解した細胞を利用することができる。細胞系、組織バイオプシ及び血液などの無傷細胞で、例えば蛍光活性化細胞分類、及び免疫蛍光細胞を検出するためのレーザーまたは光学顕微鏡を用いる免疫学的技術によりFANCD2のアイソフォームをインサイチュウ同定できる（図1から7、9から14）。例えば保存状態、例えばパラフィンに包埋してスライド上で調製した、または凍結組織切片として、組織または細胞単層のバイオプシを抗体に暴露し、FANCD2-Lを検出し、次いで蛍光顕微鏡により試験できる。

#### 【0069】

本発明の実施形態では、患者に由来する細胞系または癌細胞系をイムノプロットティング及び免疫蛍光により分析し、FANCD2アイソフォームの量の変化を検出するための新規の単純診断試験を提供する。FA経路の上流に欠損を有する個体はFANCD2をユビキチン化することができないので、診断試験はまたFA経路の上流の欠損をスクリーニン

10

20

30

40

50

グするための手段、及び現在用いられているF AのD E B / M M C染色体切断試験の実質的代替方法をも提供する。形質転換された患者由来細胞系を形成するレトロウイルス遺伝子移動(Pulsipherら、Mol. Med. 4: 468 - 179 (1998))をF A N C D 2イムノプロットングと一緒に組み合わせたアッセイなどの別のアッセイを用いて、表2に記載するいずれかの症候群、とりわけF Aの症候群を有する、新規に診断された患者を迅速にサブタイピング分析することができる。

【0070】

前記のアッセイは診断用研究室により実施され得るか、また別に診断キットが製造され、医療介護提供者または自己診断のための私用に個人に販売され得る。これらの試験の結果及び解釈の情報は患者の症状の診断及び処置において医療介護提供者に有用である。

10

【0071】

遺伝学的試験は対象に、疫学的基礎に対する特定の症状に関する迅速で信頼できる危険性分析を提供できる。我々のデータは、遺伝的異種性がF A N C D 2相補群の中のF Aを有する患者に生じることを示唆している。我々は遺伝的異種性及び疾患、並びに遺伝的異種性及びF A N C D 2 - Lの存在または不在に至る翻訳後修飾異常の間に相関性を見出した。この相関性により予後診断試験及び診断試験の基礎、並びにD N A修復異常により特徴付けられるいずれかの症候群の処置が提供される。例えば、採血した血液から得られた細胞サンプルまたは対象に由来する別の細胞からの核酸を、F A N C D 2遺伝子における変異に関して分析でき、対象が癌に対する感受性を増していることを診断できる。

【0072】

我々は高頻度の癌に相関する領域の染色体3 p上の3025.3のF A N C D 2遺伝子を位置決定した。細胞遺伝学及び異型接合性損失(LOH)研究により、染色体3 pの欠失は肺癌の全ての形態において高頻度で発生することが示された(Toddら、Cancer Res. 57: 1344 - 1352)。例えば異型接合性欠失は3つの扁平上皮細胞系で3 p 21の領域内に見出された。同型接合性欠失もまた小型細胞腫瘍において3 p 12及び3 p 14.2で見出された(Franklinら、Cancer Res. 57: 1344 - 1352 (1997))。F A N C D 2のこのマッピングは、この染色体領域が重要な腫瘍抑制遺伝子を含むという理論を支持する。家族性卵巣癌の新規感受性遺伝子の染色体3 p 22からp 25への局在を報告している、Sekineら、Human Molecular Genetics 10: 1421 - 1429 (2001)の最近の文献によりこれはさらに支持されている。F A N C D 2部位(3 p 25.3)のみならず、環境物質への暴露及び正常な加齢過程から生じる細胞損傷に続くD N A修復の欠損から生じる可能性がある染色体の別の部位での変異の結果である腫瘍の危険性の増加を診断するために、F A N C D 2 - Lの低下または不在が本明細書で提案される。

20

30

【0073】

より多くの個人及び家族がF A N C D 2遺伝子の遺伝的欠損をスクリーニングされると、データベースに異なる変異のための集団頻度が集められ、これらの変異及び個体の健康プロファイル間の相関性が作成され、遺伝学的分析の予測価値が継続的に改善されるように、データベースが開発される。2つの家族のF A N C D 2に関するアレル特異的系統図分析の実例を図10に提示する。

40

【0074】

F A N C D 2遺伝子の変異の診断を、F A N C D 2 - Lタンパク質の量の低下を検出するための迅速な免疫学的アッセイにより最初に検出できる。次いでF A経路のいずれかの遺伝子の欠損に利用可能なプローブ及びプライマーで陽性サンプルをスクリーニングできる。F A N C D 2遺伝子における欠損が関係する場合、例えば表7で提供されるプライマーまたはプローブを用いて変異を検出できる。かかるプライマーまたはプローブにより変異が検出されないこれらのサンプルでは、F A N C D 2遺伝子全体をシーケンシングして遺伝子における変異の存在及び位置を決定できる。

【0075】

F A N C D 2遺伝子座における遺伝子欠損を同定するのに用いられる核酸スクリーニン

50

グアッセイにはPCR及び変異を検出するための非PCR基盤アッセイなどが含まれ得る。特定のアレルにおける変異の存在に関して細胞ゲノムを分析する多くの研究法がある。野生型FANCD2アレルの変化を、例えば点変異、欠失または挿入によるかそうでなくても、プローブを用いる標準的な方法により検出できる(米国特許第6033857号)。標準的な方法には:(a)無傷全細胞で用いることができる蛍光インサイチュウ・ハイブリダイゼーション(FISH);及び(b)アレル特異的オリゴヌクレオチド(ASO)などがあり、これらを用いて、単離された核酸におけるハイブリダイゼーションを用いて変異を検出できる(Connerら、Hum. Genet. 85:55-74(1989))。別の技術には(a)非変性ポリアクリルアミドゲル上の一本鎖DNAの電気泳動可動性のシフトを観察すること、(b)組織サンプルから単離されたゲノムDNAにFANCD2遺伝子プローブをハイブリダイズすること、(d)組織サンプルからのFANCD2遺伝子の全てまたは一部を増幅して増幅された配列を製造し、増幅された配列をシーケンシングすること、(e)特異的FANCD2変異アレルのためのプライマーを用いて組織サンプルからのFANCD2遺伝子の全てまたは一部を増幅すること、(f)組織サンプルからのFANCD2遺伝子の全てまたは一部を分子クローニングして、クローン化された配列を製造し、クローン化された配列をシーケンシングすること、(g)(i)組織サンプルから単離されたFANCD2遺伝子またはFANCD2 mRNA、及び(ii)ヒト野生型FANCD2遺伝子配列に相補的な核酸プローブ間の誤対合を同定することであって、ここで(i)及び(ii)は互いにハイブリダイズして二重鎖を形成する、(h)組織サンプル中のFANCD2遺伝子配列の増幅及び野生型FANCD2遺伝子配列を含んでなる核酸プローブに増幅された配列をハイブリダイズすること、(I)組織サンプル中のFANCD2遺伝子ハイブリダイズの増幅及び変異FANCD2遺伝子配列を含んでなる核酸プローブに増幅された配列をハイブリダイズすること、(j)組織サンプル中の欠失変異を検出するためにスクリーニングすること、(k)組織サンプル中の点変異を検出するためにスクリーニングすること、(l)組織サンプル中の挿入変異を検出するためにスクリーニングすること、並びに(m)組織サンプルのFANCD2遺伝子の、FANCD2遺伝子を含んでなる核酸プローブとのインサイチュウ・ハイブリダイゼーションなどがある。

#### 【0076】

点変異に関して遺伝子またはゲノムの比較的短い領域をスキャンするのがしばしば望ましい:配列の全ての可能性のある部位を試験するのに必要な多くのオリゴヌクレオチドを効率的な組換え方法により作製できる(Southern, E.M.ら、Nucleic Acids Res. 22:1368-1373(1994))。アレイをリガーゼまたはポリメラーゼと組み合わせて使用して標的配列の全ての部位を探ることができる(米国特許第6307039号)。ハイブリダイゼーションによる変異の分析を例えばゲル、アレイまたはドットプロットの手段により実施できる。

#### 【0077】

全遺伝子をシーケンシングして変異を同定することができる(米国特許第6033857号)。標的配列に結合したときに固体相支持体上の補体に結合するようになる最低交差ハイブリダイジングセットからのオリゴヌクレオチドタグを用いてFANCD2遺伝子座のシーケンシングを達成できる(米国特許第6280935号)。

#### 【0078】

FANCD2遺伝子の変異を検出する別の研究法には米国特許第6297010号、米国特許第6287772号及び米国特許第6300076号に記載されているものなどがある。核酸マイクロチップ技術またはチップのラボラトリーを使用する複数のサンプルの分析をアッセイに用いることができることもさらに企図されている。

#### 【0079】

FANCD2遺伝子変異またはFANCD2-Lタンパク質の欠損を検出するための核酸診断試験または抗体基盤の試験を用いて腫瘍が進行している対象をスクリーニングすることができる。かかるスクリーニングに基づいてサンプルを、黒色腫、白血病、アストチ

トーム、神経膠芽細胞腫、リンパ腫、神経膠腫、ホジキンリンパ腫、慢性リンパ球性白血病並びに膵臓、乳房、卵巣、子宮、精巣、脳下垂体、腎臓、胃、食道及び直腸の癌などの広範な癌を有する対象から入手することができる。患者の処置利益を最大にするための適当な処置プロトコル選択するための臨床医の能力が改善されている。とりわけ、遺伝的損傷またはFANCD2-Lタンパク質の欠損を種々の既存の化学療法剤及び放射線治療に対する応答性と関連させることができる。

#### 【0080】

細胞アッセイ(実施例11から12)において、及びノックアウトマウスモデル(実施例10)においてFANCD2-Lを上昇させるFANCD2-Sのモノユビキチン化を変調する分子に関してスクリーニングすることにより新規治療的処置を開発できる。かかる分子には直接FANCD2に、またはBRACA-1と相互作用すると思われる分子例えばBRACA-2に結合して、今度はFANCD2を活性化すると思われる分子などでよい。

10

#### 【0081】

FANCD2遺伝子またはタンパク質の欠損に依存するスクリーニングアッセイに加えて、FANCD2-Lの形成に必要であるユビキチン化反応の失敗が観察されるが、それはFANCD2-Sそのものを含むFANCD2の翻訳後修飾に先行する、いずれかの点でのFA経路における欠失の結果であろう。反応の終末工程を知ることにより、分子が破壊された経路を修復することが見出されるまで、小型分子が「FA経路破壊」含有する細胞において、またはインビトロでスクリーニングされるスクリーニングアッセイを構築することが可能になる。次いでこの分子をプローブとして利用して欠損の特性を同定することができる。これをさらに治療用物質として用いて欠損を修復することができる。例えば、我々は、細胞周期停止及びFA細胞の増殖低下がタンパク質のサイクリンファミリーのメンバーであるSPHARタンパク質の過剰発現により部分的に修正され得ることを示した。これは治療用小型分子を同定するためのスクリーニングとして適当であるアッセイの基礎を成すことができる。

20

#### 【0082】

翻訳後修飾されたFANCD2で欠損している細胞はとりわけDNA損傷に対して感受性が高い。化合物(毒性分子を含む)がDNAを損傷する能力を有しているかどうかを決定するための感受性スクリーニングとしてこれらの細胞を提供できる。逆にいえば、これらの細胞はまた、化合物がDNA損傷に対して細胞を保護できるかどうかを決定するための感受性スクリーニングとして提供される。

30

#### 【0083】

FA患者及びDNA修復損傷に関連する症候群を患う患者は骨髄損傷の合併症から死亡する。遺伝子移入は欠損を修正するための治療オプションである。複数の欠損がFA経路の至る所で生じ得る。我々は、終末工程が細胞及び生物の適切な機能化に非常に重要であることを示した。本発明の実施形態では、FANCD2-SのFANCD2-Lへの変換がうまく達成されるようにこの変換を標的化する遺伝子治療または治療薬により、FA経路のどこかの欠損の修正が満足できるように達成され得る。

#### 【0084】

一般に許容される方法に従って、例えばFriedmanにより「Therapy for Genetic Disease」T. Friedman編、オックスフォード・ユニバーシティ・プレス、105-121頁(1991)に記載されるように、遺伝子治療を実施できる。エクソビオまたはインビトロ遺伝子治療に関して標的化された組織には骨髄例えば貧血を発症する前の造血幹細胞及び発達異常に関与する胎児組織などがある。遺伝子治療は、変異FANCD2アレルを担持する細胞に野生型FANCD2-L機能を提供できる。かかる機能を提供することにより、受体細胞の腫瘍成長が抑制されるか、またはファンコーニ貧血の症状が改善されるはずである。

40

#### 【0085】

野生型FANCD2遺伝子または遺伝子の一部を、遺伝子が染色体外因子のままである

50

ようにベクターの細胞に導入できる。かかる状況では、遺伝子は染色体外位置から細胞により発現され得る。遺伝子部分が変異 F A N C D 2 アレルを担持する細胞に導入及び発現される場合、遺伝子部分は細胞の非腫瘍性成長に必要な F A N C D 2 タンパク質の一部をコードできる。また別に、野生型 F A N C D 2 遺伝子またはその一部を、それが細胞に存在する内因性変異 F A N C D 2 遺伝子と組換わるように変異細胞に導入することができる。

【 0 0 8 6 】

ウイルスベクターは遺伝子治療を達成するためのベクターの1つのクラスである。ウイルス媒介遺伝子治療を、リポソーム分配を用いるインビボ直接遺伝子移入と組み合わせることができ、これによりウイルスベクターを腫瘍細胞に指向させ、周辺の非分裂細胞には指向させないようにすることができる。また別に、ウイルスベクター産生細胞系を腫瘍に注入することができる (Culverら、1992)。産生細胞の注入により、次いでベクター粒子の連続的な供給源が提供される。この技術は手術不能な脳腫瘍を有するヒトへの使用が認められている。

10

【 0 0 8 7 】

ベクターを患者に、腫瘍の部位に局所的または全身的のいずれかで (別の部位に転移したいずれの腫瘍細胞にも到達するように) 注入できる。移入された遺伝子は標的化された腫瘍細胞の各々のゲノムに永続的に組み込まれない場合、処置を定期的に繰り返さなければならない。

【 0 0 8 8 】

組換え及び染色体外保持の双方のための遺伝子の導入用ベクターは技術分野で公知であり (例えば米国特許第 5 2 5 2 4 7 9 号及び P C T 9 3 / 0 7 2 8 2 及び米国特許第 6 3 0 3 3 7 9 号)、F A N C D 2 - L を含有するウイルスベクター例えばレトロウイルス、ヘルペスウイルス (米国特許第 6 2 8 7 5 5 7 号) またはアデノウイルス (米国特許第 6 2 8 1 0 1 0 号) またはプラスミドベクターなどを含む。

20

【 0 0 8 9 】

治療用遺伝子配列を担持するベクターまたは遺伝子もしくは遺伝子の断片をコードする D N A を患者に、腫瘍の部位に局所的または全身的のいずれかで注入して、転移した腫瘍細胞に到達するようにできる。ベクターをさらに操作することなく標的化を達成できるか、またはベクターを腫瘍に結合特異性を有する分子と結合させることができ、ここでかかる分子はレセプターアゴニストまたはアンタゴニストでよく、さらにペプチド、脂質 (リポソームを含む) または糖類 (オリゴ多糖類または多糖類を含む) 及び合成標的化分子を含んでよい。D N A はポリリジンを通して結合リガンドに抱合され得る。トランスフェクトされた遺伝子が永続的に標的化腫瘍細胞の各々のゲノムに組み込まれない場合、処置を定期的に繰り返さなければならない。

30

【 0 0 9 0 】

患者に導入する前に D N A を細胞に導入する方法は、技術分野で報告されているような技術 (米国特許第 6 0 3 3 8 5 7 号)、例えばエレクトロポレーション、リン酸カルシウム共沈殿及びウイルス形質導入を用いて達成でき、方法の選択は通常の実験者の能力範囲内である。

40

【 0 0 9 1 】

野生型 F A N C D 2 遺伝子または変異 F A N C D 2 遺伝子で形質転換された細胞をモデル系として用いて欠損 D N A 修復の結果である疾患の緩解及びかかる緩解を促進する薬物処置を研究することができる。

【 0 0 9 2 】

前記に一般的に論じたように、F A N C D 2 遺伝子またはそのフラグメントを、適当な場合遺伝子治療方法に用いて、異常細胞においてかかる遺伝子の発現産物の量を増加させることができる。かかる遺伝子治療はとりわけ前癌細胞での使用に適しており、ここで正常細胞と比較して F A N C D 2 - L ポリペプチドのレベルを不在または低減させ、F A N C D 2 - L のレベル増加は欠損 D N A 修復から生じる欠損の蓄積を減速させ、従って癌状

50

態の開始を遅延させる。変異遺伝子が「正常」レベルで発現されるが、FANCD2-Lアイソフォームのレベルが低下しているこれらの細胞でさえ、FANCD2遺伝子の発現レベルを増加させることはまた有用である。正常なDNA修復におけるFANCD2-Lの非常に重要な役割により、FANCD2-Lのレベル低下を引きおこす欠損を修正する治療薬を開発する機会が提供される。新規治療薬を開発するための1つの研究法は合理的な薬物設計によるものである。合理的な薬物設計により、さらに活性化もしくは安定なポリペプチドの形態に変えるために、またはインビボでポリペプチドの機能を増強もしくは干渉する小型分子を設計するために、目的の生物学的に活性化ポリペプチドの、またはこれらが相互作用する小型分子（例えばアゴニスト、アンタゴニスト、インヒビターまたはエンハンサー）の構造類似体が提供され得る。Hodgson、1991。合理的な薬物設計によりFANCD2-L活性もしくは安定性が改善されているか、またはFANCD2-L活性のエンハンサー、インヒビター、アゴニストもしくはアンタゴニストとして作用する小型分子を提供され得る。クローン化されたFANCD2配列の利用率のために、十分量のFANCD2-Lポリペプチドを、X線結晶学としてかかる分析実験を実施するために利用可能にできる。加えて、本明細書で提供するFANCD2-Lタンパク質の配列に関する知識は、X線結晶学の代わりに、またはそれに加えてこれらをコンピューターモデリング技術に使用する手引きとなる。

10

#### 【0093】

FANCD2-L活性を有するペプチドまたはその他の分子を治療用処方でタンパク質が欠乏している細胞に供給できる。FANCD2-Lタンパク質の配列はいくつかの生物（ヒト、ハエ及び植物）に関して開示されている（配列番号：1から3）。細菌におけるcDNA配列の発現により、例えばさらなる翻訳後修飾を有する公知の発現ベクターを用いてFANCD2を製造することができる。或いは、FANCD2-LポリペプチドをFANCD2-L産生哺乳動物細胞から抽出することができる。加えて、合成化学の技術を用いてFANCD2-Lタンパク質を合成することができる。FANCD2-L活性を有する別の分子（例えばペプチド、薬物または有機化合物）を治療薬として用いることもできる。実質的に類似する機能を有する修飾されたポリペプチドはペプチド治療にも用いられる。

20

#### 【0094】

同様に、変異FANCD2アレルを担持するかまたは不十分なレベルのFANCD2-Lを作製する細胞及び動物をモデル系として用いて、治療薬としての可能性を有する物質に関して研究及び試験することができる。体細胞または生殖細胞系のいずれかでよい細胞をFANCD2-Lのレベルが低下した個体から単離することができる。また別に、前記したようにFANCD2-Lのレベルが低下するように細胞系を操作することができる。被験物質を細胞に適用した後、細胞のDNA修復欠陥形質転換表現型を決定する。

30

#### 【0095】

新規候補治療用分子の有効性を有効性及び毒性欠如に関する動物実験において試験できる。標準的な技術を用いて、動物全体の変異誘発の後か、または生殖細胞または接合子の遺伝子操作の後に動物を選択して、トランスジェニック動物を形成できる。かかる処置には、通常第2の動物種からの変異FANCD2アレルの挿入、及び破壊された相同性遺伝子の挿入などがある。また別に、慣用される技術を用いて、動物の内因性FANCD2遺伝子を挿入もしくは欠失変異または別の遺伝子変化により破壊することができる（Capocchi、Science 244：1288-1292（1989）；Valancius及びSmithies、（1991））。被験物質を動物に投与した後、腫瘍の成長を評価しなければならない。被験物質が欠損DNA修復から生じる病理を防御または抑制する場合、次いで被験物質は本明細書で同定された疾患の処置のための治療薬の候補となる。

40

#### 【0096】

本発明を以下の実施例を参照して記載するが、これは説明のために提示されるものであり、本発明をいかなるようにも限定することを意図するものではない。技術分野で公知の

50

標準的な技術または以下に特記する技術を利用した。

【0097】

本明細書で引用した全ての参考文献は引用により本明細書に包含される。

【0098】

(実施例)

実施例1：実施例2から8で用いられる実験プロトコル

細胞系及び培養条件

エプスタイン・バー・ウイルス(EBV)形質転換リンパ芽球を、15%熱不活化ウシ胎児血清(FCS)を補充したRPMI培地中で維持し、37℃で5%CO<sub>2</sub>含有加湿大気下で成長させた。対照リンパ芽球系(PD7)及びFAリンパ芽球系(FA-A(HSC72)、FA-C(PD-4)、FA-D(PD-20)、FA-F(EUFA121)、及びFA-G(EUFA316))が以前に報告されている(de Winterら、Nat. Genet. 20:281-283(1998))(Whitneyら、Nat. Genet. 11:341-343(1995))(Yamashitaら、P.N.A.S. 91:6712-6716(1994))(de Winterら、Am. J. Hum. Genet. 67:1306-1308(2000))。PD81はFA-A患者からのリンパ芽球細胞系である。SV40形質転換FA繊維芽細胞であるGM6914、PD426、FAG3263SV及びPD20F、並びにHeLa細胞を15%FCSを補充したDMEM中で成長させた。FA細胞(リンパ芽細胞及び繊維芽細胞の双方)は対応するFANScDNAを含有するpMMPレトロウイルスベクターで機能的に補足され、機能的な補足はMMCアッセイにより確認された(Garcia-Higueraら、Mol. Cell. Biol. 19:4866-4873(1999))(Kuangら、Blood 96:1625-1632(2000))。

10

20

【0099】

細胞周期同調化

HeLa細胞、GM6914細胞及びpMMP-FANCAレトロウイルスで修正されたGM6914を前記した二重チミジン遮断法にわずかに修正を加えた方法(Kupferら、Blood 90:1047-1054(1997))より同調化した。簡単には、細胞を2mMチミジンで18時間、チミジン不含培地で10時間、及びさらに2mMチミジンで18時間処置して、G1/S境界で細胞周期を停止させた。細胞をPBSで2回洗浄し、ついでDMEM+15%FCS中に放出させて、種々の時間間隔で分析した。

30

【0100】

また別に、HeLa細胞を0.5mMミモシン(シグマ)で24時間処置して、G1期に同調させ(Krude、1999)、PBSで2回洗浄し、DMEM+15%FCS中に放出させた。M期での同調化に関しては、ノコダゾール遮断を用いた(Ruffnerら、Mol. Cell. Biol. 19:4843-4854(1999))。細胞を0.1µg/mlノコダゾール(シグマ)で15時間処置し、非付着性細胞をPBSで2回洗浄し、DMEM+15%中に再プレートした。

40

【0101】

細胞周期分析

トリプシン処置した細胞をPBS 0.5mlに再懸濁し、氷冷エタノール5mlを加えて固定した。次に細胞を1%ウシ血清アルブミン分画Vを含むPBS(1%BSA/PBS)(シグマ)で2回洗浄し、1%BSA/PBS 0.24mlに再懸濁した。38mMクエン酸ナトリウム(pH7.0)中500µg/mlヨウ化プロピジウム(シグマ)30µl及び10mg/mlDNアーゼ不含RNアーゼA(シグマ)30µlを加えた後、サンプルを37℃で30分間インキュベートした。DNA含量をFACSキャン(ベクトン・ディッキンソン)により測定し、データをセル・クエスト及びモディフィットLTプログラム(ベクトン・ディッキンソン)により分析した。

50

【0102】

### 抗FANCD2抗血清の作製

抗原供給源としてGST-FANCD2(N末端)融合タンパク質を用いてFANCD2に対するウサギポリクローナル抗血清を作製した。プライマー(配列番号:95)DF4EcoRI(5'AGCCTCgaattcGTTTCCAAAAGAAGACTGTCA-3')及び(配列番号:96)DR816Xh(5'-GGTATCctcgaagTCAAGACGACAACCTTATCCATCA-3')を用いて、FANCD2 cDNA全長からポリメラーゼ連鎖反応(PCR)により5'フラグメントを増幅した。得られた841bpのPCR産物はFANCD2ポリペプチドのアミノ末端272アミノ酸をコードし、これをEcoRI/XhoIで消化し、プラスミドGEX4T-1(ファルマシア)のEcoRI/XhoI部位にサブクローニングした。予想された大きさ(54kD)のGST-FANCD2(N末端)融合タンパク質が大腸菌(E.coli)DH5株で発現され、グルタチオン-S-セファロースで精製し、ニュージーランド白ウサギを免疫するのに使用した。GSTタンパク質を負荷したアミノリンク・プラス・カラム(ピヤース)で継代することにより、及びGST-FANCD2(N末端)融合タンパク質を負荷したアミノリンク・プラス・カラムで継代することにより、FANCD2特異的免疫抗血清をアフィニティー精製した。

10

#### 【0103】

### 抗FANCD2モノクローナル抗体の作製

2つの抗FANCD2モノクローナル抗体を以下のように作製した。Balb/cマウスをGST-FANCD2(N末端)融合タンパク質で免疫し、これはFANCD2に対するウサギポリクローナル抗血清(E35)の作製に用いたのと同じ融合タンパク質である。融合精母細胞を収穫する前に動物を4に置換免疫原で追加免疫し、骨髓腫細胞でハイブリダイゼーションを実施した。ハイブリドーマ上澄を収集し、最初のスクリーニングとして標準的なELISAアッセイを、第2のスクリーニングとしてFANCD2のイムノブロット分析を用いて検定した。2つの抗ヒトFANCD2モノクローナル抗体(MoAbs)(FI17及びFI14)をさらなる研究のために選択した。2つの陽性細胞系からのハイブリドーマ上澄を遠心により清澄化した。上澄をMoAbsとしてウェスタンブロットに使用した。IgGに関するアフィニティカラムを用いてMoAbsを精製した。リン酸緩衝生理食塩水(PBS)中MoAbsを0.5mg/mlストックとして保存した。抗HA抗体(HA.11)はパブコのものである。

20

30

#### 【0104】

### イムノブロットティング

細胞を1Xサンプルバッファー(50mM トリスHCl、pH6.8、86mM 2-メルカプトエタノール、2%ドデシル硫酸ナトリウム(SDS))で溶解し、5分間沸騰させ、7.5%ポリアクリルアミドSDSゲル電気泳動に供した。電気泳動の後、25mM トリス塩基、200mM グリシン、20%メタノールで充填した浸漬移動装置(バイオラッド)を用いてタンパク質をニトロセルロースに移した。TBS-T(50mM トリスHCl、pH8.0、150mM NaCl、0.1%トゥイーン20)中5%脱脂粉乳で遮断した後、膜をTBS-Tで希釈した1次抗体(アフィニティー精製した抗FANCD2ポリクローナル抗体(E35)または抗HA(HA.11)には1:1000希釈、抗FANCD2マウスモノクローナル抗体(FI17)には1:200希釈)と共にインキュベートし、何度も洗浄し、適当なセイヨウワサビペルオキシダーゼ結合2次抗体(アマシャム)と共にインキュベートした。検出に化学ルミネセンスを用いた。

40

#### 【0105】

### DNA損傷の作製

ガンマセル40装置を用いてガンマ線を照射した。穏やかに培養培地を吸引した後、ストラタリンカー(ストラタジーン)を用いてUV暴露を達成した。マイトマイシンC処置用に細胞を継続的に指示された時間薬物に暴露した。ヒドロキシウレア(シグマ)を最終濃度1mMまで24時間添加した。

#### 【0106】

50

## モノユビキチン化FANCD2の検出

製造者のプロトコルに従って、HeLa細胞（またはFA-G繊維芽細胞、FAG326SV）をFuGENE6（ロッシュ）を用いてトランスフェクトした。HeLa細胞を15cm組織培養皿にプレートし、皿あたり15μgのHAタグ化ユビキチン発現ベクター（pMT123）（Treierら、Cell 78:787-798（1994））でトランスフェクトした。トランスフェクション後12時間、指示された濃度のMMC（0、10、40、160ng/ml）または指示された用量のIR（0、5、10、20Gy）で細胞を処置した。

MMCと共に12時間のインキュベーション、またはIR処置後2時間の後、プロテアーゼインヒビター（1μg/ml ロイペプチン及びペプスタチン、21μg/ml アプロチニン、1mM フッ化フェニルメチルスルホニル）及びホスファターゼインヒビター（1mM オルソバナジウム酸ナトリウム、10mM フッ化ナトリウム）を補充した溶解バッファー（50mM トリスHCl、pH7.4、150mM NaCl、1（容量/容量）%トリトンX-100）中で全細胞抽出物を調製した。各々の免疫沈降（IP）を4mgのタンパク質を含有するように標準化したこと以外は本質的には記載されたように（Kupferら、1997）、FANCD2（E35）に対するモノクローナル抗体を用いて免疫沈降（IP）を実施した。陰性対照として同一のウサギからの免疫前血清をIP反応において使用した。抗HA（HA.11）または抗FANCD2（FI17）モノクローナル抗体を用いてイムノプロットングを行った。

## 【0107】

## ユビキチンアルデヒド処置

HeLa細胞を1mM ヒドロキシウレアで24時間処置し、プロテアーゼインヒビター及びホスファターゼインヒビターを補充した溶解バッファー中で全細胞抽出物を調製した。DMSO中15μM ユビキチンアルデヒド（ボストン・バイオケム）6.7μlを伴うかまたはDMSO6.7μlを伴う反応物67μl中細胞ライゼート200μgを指示された期間30 または37 でインキュベーションした。2Xサンプルバッファー67μlを各サンプルに加え、サンプルを5分間沸騰させ、7.5%SDS-PAGEで分離し、FI17モノクローナル抗ヒトFANCD2抗体を用いてFANCD2に関してイムノプロットングした。

## 【0108】

## 免疫蛍光顕微鏡法

PBS中2%パラホルムアルデヒドで細胞を固定し、続いてPBS中0.3%トリトンX-100で透過処理した（10分間）。10%ヤギ血清で遮断した後、PBS中0.1%NP-40（遮断バッファー）、特異的抗体を遮断バッファー中適当な希釈率で加え、室温で2から4時間インキュベーションした。アフィニティー精製されたE35ポリクローナル抗体（1/100）を用いてFANCD2を検出した。BRCA1検出に関しては、我々は市販のモノクローナル抗体（D-9、サンタ・クルーズ）を2μg/mlで用いた。続いて細胞をPBS+0.1%NP-40で3回洗浄し（各洗浄10から15分）、種特異的蛍光またはテキサス・レッド抱合2次抗体（ジャクソン・イムノリサーチ）を遮断バッファー（抗マウス 1/200、抗ウサギ 1/1000）で希釈し、添加した。室温で1時間の後、10から15分の洗浄をさらに3回適用し、ベクタシールド（ベクター・ラボラトリーズ）中でスライドをマウントした。画像をニコン顕微鏡で捕らえ、アドビ・フォトショップ・ソフトウェアを用いて処理した。

## 【0109】

## 減数分裂染色体染色

16から28日齢の雄マウスからのパキテン期及びディプロテン期精母細胞の表面塗抹を記載されるように（Petersら、1997）調製した。マウスSCP3タンパク質に対するポリクローナルヤギ抗体を用いて、減数分裂調製物の軸エレメント及び対合複合体を可視化した。マウスBRCA1に対するM118マウスモノクローナル抗体を標準的な技術により、マウスをネズミBRCA1タンパク質で免疫することにより作製した。ア

フィニティー精製 E 3 5 ウサギポリクローナル抗体を 1 : 2 0 0 希釈で用いて F A C D を検出した。抗体インキュベーション及び検出方法は ( K e e g a n ら、 G e n e s D e v . 1 0 : 2 4 2 3 - 2 4 3 7 ( 1 9 9 6 ) ) に記載されるように、 ( M o n e n s ら、 J . C e l l . B i o l . 1 0 5 : 9 3 - 1 0 3 ( 1 9 8 7 ) ) のプロトコルを修飾したものであった。ロバ抗マウス I g G - F I T C 抱合、ロバ抗ウサギ I g G T R I T C 抱合、及びロバ抗ヤギ I g G - C y 5 抱合 2 次抗体を検出に使用した ( ジャクソン・イムノリサーチ・ラボラトリーズ )。全ての調製物を 4 ' , 6 ' ジアミノ - 2 - フェニルインドール ( D A P I、シグマ ) で対比染色し、 D A B C O ( シグマ ) 退色防止溶液にマウントした。調製物をニコン E 1 0 0 0 顕微鏡 ( 6 0 X C F I プラン・アポクロマツ及び 1 0 0 X C F I プラン・フルオロオイル・イマージョン・オブジェクトブズ ) で試験した。各々の蛍光色素 ( F I T C、T R I T C、C y 5 及び D A P I ) 画像を冷却 C C D カメラ ( プリンストン・インスツルメンツ・マイクロマックス ) を調節する I P ラブソフトウェア ( スキャナリティック ) による 8 0 0 x 1 0 0 0 ピクセル 1 2 ビットソースイメージとして別個に捕らえ、アドビ・フォトショップを用いて別個の 1 2 ビット・グレー・スケール・イメージを再度サンプリングし、2 4 ビット擬似着色し、マージした。

10

20

30

40

50

#### 【 0 1 1 0 】

##### 実施例 2 : 共通の細胞経路における F A 遺伝子相互作用

正常なリンパ芽球は F A N C D 2 の 2 つのアイソフォームである短形態 ( F A N C D 2 - S、1 5 5 k D ) 及び長形態 ( F A N C D 2 - L、1 6 2 k D ) を発現する。図 1 は全細胞抽出物をリンパ芽球系から調製し、細胞性タンパク質を抗 F A N C D 2 抗血清で免疫沈降した場合に何が生じるかを示している。正常な野生型細胞は F A N C D 2 の 2 つのアイソフォームである低分子量アイソフォーム、F A N C D 2 - S ( 1 5 5 k D ) 及び高分子量アイソフォーム、F A N C D 2 - L ( 1 6 2 k D ) を発現する。F A N C D 2 - S はクローン化された F A N C D 2 c D N A の 1 次翻訳産物である。我々は次に F A N C D 2 アイソフォームの発現に関して大きな一連の F A リンパ芽球及び繊維芽細胞を評価した ( 表 5 )。F A 細胞系の対応する F A c D N A での修正により機能的補正及び高分子量アイソフォーム、F A N C D 2 - L の回復に至った。

#### 【 0 1 1 1 】

以前に報告されているように、F A 細胞は D N A 架橋試薬である M M C に、及びある場合には電離放射線 ( I R ) に感受性を有する。興味深いことに、複数の相補群 ( A、C、G 及び F ) からの F A 細胞は F A N C D 2 - S アイソフォームのみを発現した ( 図 1 A、レーン 3、7、9、1 1 )。相補群 B 及び E からの F A 細胞もまた F A N C D 2 - S のみを発現する。対応する F A N C c D N A を有する F A 細胞の M M C 及び I R 感受性の機能的修正により F A タンパク質複合体を回復させ、高分子量アイソフォーム ( F A N C D 2 - L ) を回復させた ( 図 1 A、レーン 4、8、1 0、1 2 )。これらを考慮に入れて、これらの結果から F A N C A、F A N C C、F A N C F 及び F A N C G を含む F A タンパク質複合体は直接的または間接的に F A N C D 2 の 2 つのアイソフォームの発現を制御することが示された。従って、6 個のクローン化 F A 遺伝子は共通の経路で相互作用するようである。

#### 【 0 1 1 2 】

実施例 3 : F A タンパク質複合体は F A N C D 2 のモノユビキチン化に必要である。

#### 【 0 1 1 3 】

F A N C D 2 の高分子量アイソフォームは F A N C D 2 m R N A の代替スプライシングまたは F A N C D 2 タンパク質の ( 複数の ) 翻訳後修飾などの 1 つまたはそれ以上のメカニズムの結果である。リン酸塩での処置は F A N C D 2 - L を F A N C D 2 - S に変換せず、これはリン酸化単独では観察された分子量の差に寄与しないことを示している。

#### 【 0 1 1 4 】

別の可能な F A N C D 2 の翻訳後修飾を同定するために、我々は最初に F A N C D 2 - S の F A N C D 2 - L への変換を制御する細胞条件を捜した ( 図 1 B、C )。F A 細胞は M M C 及び I R に感受性を有するので、これらの物質が正常な脂肪において F A N C D 2

- SのFANCD2-Lへの変換を制御するのではと推論した。興味深いことに、MMC (図1B、レーン1から6)またはIR (図1C、レーン1から6)で処置したHeLa細胞はFANCD2-Lアイソフォームの発現の濃度依存的増加を示した。

【0115】

FANCD2-LがFANCD2-Sのユビキチン化アイソフォームであるかどうかを決定するために、HA-ユビキチンをコードするcDNAでHeLa細胞をトランスフェクトした(Treierら、1994)。MMC (図1B、レーン7から10)またはIR (図1C、レーン7から10)への細胞暴露によりFANCD2のHA-ユビキチン抱合体が用量依存的に増加した。FANCD2-Lアイソフォームのみが抗HA抗体と免疫反応し、FANCD2-Sアイソフォームは反応しなかった。FANCD2はFA細胞においてユビキチン化されなかったが、FANCD2ユビキチン化はこれらの細胞の機能的補足で回復した。

FANCD2-S及びFANCD2-Lアイソフォームは7kDだけ異なり、FANCD2-Lは恐らくアミド結合によりFANCD2の内部リジン残基に共有結合した単一のユビキチン部分を含む。

【0116】

モノユビキチン化を確認するために、我々はHeLa細胞からFANCD2-Lタンパク質を単離し、質量分析によりそのトリプシンフラグメントを分析した(Wuら、Science 289:11a)。明らかにユビキチントリプシンフラグメントが同定され、モノユビキチン化部位(FANCD2のK561)もまた同定された。興味深いことに、このリジン残基はヒト、ショウジョウバエ(Drosophila)、及び線虫(Celegans)からのFANCD2配列間で保存されており、これはこの部位のユビキチン化が複数の生物におけるFA経路に非常に重要であることを示唆している。このリジン残基であるFANCD2(K561)の変異によりFANCD2モノユビキチン化の喪失に至る。

【0117】

実施例4：FANCD2を含む核フォーカスの形成は無傷のFA経路に必要である。

【0118】

我々は修正されていない、MMC感受性FA繊維芽細胞及び機能的に補足された繊維芽細胞におけるFANCD2タンパク質の免疫蛍光パターンを試験した(図2)。

【0119】

修正されたFA細胞はFANCD2-S及びFANCD2-Lアイソフォームの双方を発現した(図2A、レーン2、4、6、8)。内因性FANCD2タンパク質はヒト細胞の核でもっぱら観察され、細胞質染色が明らかであるものはなかった(図2B、aからh)。PD-20(FA-D)細胞では核免疫蛍光が低下し(図2B、d)、これはイムノプロットによるこれらの細胞におけるFANCD2タンパク質の発現の低下に合致する(図2A、レーン7)。染色体移動によりFANCD2遺伝子で機能的に修正されたPD-20細胞では、FANCD2タンパク質が2つの核パターンで染色された。最も修正された細胞は染色の拡散核パターンを有し、細胞の小分画は核フォーカスが染色された(パネルhのドットを参照)。双方の核パターンは3つの別個に誘導された抗FANCD2抗血清で観察された(1ポリクローナル、2モノクローナル抗血清)。サブタイプA、G及びCからのFA繊維芽細胞はFANCD2核免疫蛍光の拡散パターンのみを示した。これらの細胞のFANCA、FANCGまたはFANCC cDNA各々での機能的補足によりこれらの細胞のMMC抵抗性が回復し(表6)、いくつかの細胞では核フォーカスが回復した。従って、高分子量FANCD2-Lアイソフォームの存在FANCD2核フォーカスの存在と関連し、これはモノユビキチン化されたFANCD2-Lアイソフォームのみが選択的にこれらのフォーカスに局在する。

【0120】

実施例5：FANCD2タンパク質は細胞周期のS期の間核フォーカスに局在する。

## 【 0 1 2 1 】

非同調性機能補足細胞のフラクシヨンのみが F A N C D 2 核フォーカスを含有したので、我々はこれらのフォーカスが細胞周期の異なる時期に集合するのかもしれないと推論した。この仮説を試験するために、同調化した細胞の F A N C D 2 - L アイソフォームの形成及び F A N C D 2 核フォーカスを試験した ( 図 3 )。He L a 細胞を G 1 / S 境界で同調化し、S 期に放出し、F A N C D 2 - L アイソフォームの形成に関して分析した ( 図 3 A )。F A N C D 2 - L アイソフォームは G 1 後期及び S 期の間中に特異的に発現された。同調化し、補足されていない F A 細胞 ( F A - A 繊維芽細胞、G M 6 9 1 4 ) は正常から増加レベルの F A N C D 2 - S タンパク質を発現したが、細胞周期のいずれの時期にも F A N C D 2 - L を発現し損ねた。F A N C D 2 c D N A での安定したトランスフェクシヨンによるこれらの F A - A 細胞の機能的補足は F A N C D 2 - L の S 期特異的発現を回復させた。He L a 細胞を別の方法、例えばノコダゾール停止 ( 図 3 B ) またはミモシン暴露 ( 図 3 B ) で同調化した場合、F A N C D 2 - L アイソフォームの S 期特異的発現が確認された。有糸分裂で停止した細胞は F A N C D 2 - L を発現せず、これは F A N C D 2 - L アイソフォームが細胞分割の前に除去または分解されることを示唆している ( 図 3 B、有糸分裂 )。これらを考慮して、これらの結果は F A N C D 2 タンパク質のモノユビキチン化は細胞周期の間に高度に制御され、この修飾には無傷の F A 経路が必要であることが示唆される。

10

## 【 0 1 2 2 】

F A N C D 2 - L アイソフォームの細胞周期依存的発現は F A N C D 2 核フォーカスの形成とも相関する ( 図 3 C )。ノコダゾール停止 ( 有糸分裂 ) 細胞は F A N C D 2 - L アイソフォームを発現せず、F A N C D 2 核フォーカスを呈しない ( 図 3 C、0 時間 )。これらの同調化細胞が S 期を横切ることができる場合 ( 1 5 から 1 8 時間 )、F A N C D 2 核フォーカスの増加が観察された。

20

## 【 0 1 2 3 】

実施例 6 : D N A 損傷に应答して F A N C D 2 タンパク質が核フォーカスに局在する。

## 【 0 1 2 4 】

我々は D N A 損傷に应答した F A N C D 2 - L アイソフォーム及び F A N C D 2 核フォーカスの蓄積を試験した ( 図 4 )。以前の研究で、F A 細胞が D N A 鎖間架橋 ( M M C ) または二本鎖切断 ( I R ) を引き起こす作用物質に対して感受性があるが、紫外線 ( U V ) 及び単官能基アルキル化剤に対して比較的抵抗することが示されている。同調化された He L a 細胞において M M C は F A N C D 2 - S の F A N C D 2 - L への変換を活性化した ( 図 4 A )。F A N C D 2 - L への最大変換は M M C 暴露の 2 4 時間後に生じ、これは F A N C D 2 核フォーカス形成が最大になる時期に相関した。F A N C D 2 - L での増加に対応して F A N C D 2 核フォーカスが増加した。電離放射線もまた He L a 細胞における F A N C D 2 - L の時期依存的及び用量依存的増加を活性化し、これは F A N C D 2 フォーカスの対応する増加を伴った ( 図 4 B )。驚くべきことに、紫外線 ( U V ) は時期依存的及び用量依存的な F A N C D 2 - S の F A N C D 2 - L の変換を活性化しこれは F A N C D 2 フォーカスの対応する増加を伴った ( 図 4 C )。

30

## 【 0 1 2 5 】

我々は F A 細胞における D N A 損傷の影響を試験した ( 図 4 D )。複数の相補群からの F A 細胞 ( A、C 及び G ) は M M C または I R 暴露に应答した F A N C D 2 - L アイソフォームを活性化し損ない、F A N C D 2 核フォーカスを活性化し損なった。これらのデータは、F A 細胞の細胞感受性は少なくとも部分的に F A N C D 2 - L 及び F A N C D 2 核フォーカスの活性化の失敗の結果であることを示唆している。

40

## 【 0 1 2 6 】

実施例 7 : 活性化された F A N C D 2 及び B R C A 1 タンパク質の同時局在化

F A N C D 2 と同様に、乳癌感受性タンパク質である B R C A 1 は細胞の増殖において上方制御され、S 期にまたは D N A 損傷に应答して翻訳後修飾により活性化される。B R C A は高度な酸性 H M G 様ドメインを含有するカルボキシ末端の 2 0 個のアミノ酸を有し

50

、これはクロマチン修復の可能なメカニズムを示唆している。BRCA1タンパク質はIR誘導フォーカス(IRIF)でDNA修復に関する別のタンパク質、例えばRAD51またはNBS/Mre11/RAD50複合体と同時局在化する。BRCA1のピアリック変異を有する細胞はDNA修復で欠損を有し、DNA損傷物質、例えばIR及びMMCに対して感受性を有する(表5)。これらを考慮に入れて、これらのデータはFANCD2及びBRCA1タンパク質間の可能な機能的相互作用を示唆している。BRCA1フォーカスはDNA修復(BRCA1)及びチェックポイント機能(NBS)に関するATM及びATM基質を含む大きな(2mda)多タンパク質複合体である。

#### 【0127】

活性化されたFANCD2タンパク質がBRCA1タンパク質と共に同時局在化するかどうかを決定するために、HeLa細胞の二重免疫標識を実施した(図5)。電離放射線の不在下で、およそ30から50%の細胞がBRCA1核フォーカスを含有した(図5A)。対照的に、S期を横切るわずかの細胞のみがFANCD2ドットを含有した(b, e)。これらの核フォーカスはまたBRCA1及びFANCD2の双方に対する抗血清と免疫反応性であった(c, f)。IR暴露の結果、核フォーカスを含有する細胞数及び細胞あたりのフォーカス数が増加した。これらの核フォーカスはIR不在時に観察されたフォーカスよりも大きく、より蛍光性であった。またこれらのフォーカスはBRCA1及びFANCD2タンパク質の双方を含有した(i, l)。FANCD2-L及びBRCA1の相互作用はさらにIRに暴露され、指数関数的に成長しているHeLa細胞からのタンパク質の同時免疫沈降(図5B)により確認された。

10

20

#### 【0128】

我々はFANCD2-L及び核フォーカスの形成に及ぼすBRCA1発現の影響を試験した(図6)。BRCA1(-/-)細胞系であるHCC1937はカルボキシ末端トランケーションを有するBRCA1タンパク質の変異体を発現する。これらの細胞は低レベルのFANCD2-Lを発現したが(図6A)、IRはFANCD2-Lレベル増加を活性化し損ねた。また、これらの細胞ではIR誘導FANCD2フォーカス数が低下した(図6B、パネルc, d)。BRCA1 cDNAでの安定したトランスフェクションによるこれらのBRCA1(-/-)細胞の修正によりIR誘導FANCD2ユビキチン化及び核フォーカスが回復した(図6B、パネルk, l)。これらのデータは野生型BRCA1タンパク質がIR誘起FANCD2ドット形成の「オーガナイザー」として必要であることを示唆し、さらにタンパク質間の機能的相互作用を示唆する。

30

#### 【0129】

実施例8：減数分裂染色体におけるFANCD2及びBRCA1の同時局在化

有糸分裂におけるFANCD2及びBRCA1の会合により、これらのタンパク質が減数分裂前期にも同時局在するかもしれないことが示唆された。以前の研究により、BRCA1タンパク質が接合子及びパキテン期精母細胞におけるヒト対合複合体の非対合/軸エレメントに集中することが示された。減数分裂細胞におけるFANCD2及びBRCA1の可能な同時局在化を試験するために、我々はパキテン後期及びディプロテン初期のマウス精母細胞の表面塗抹をFANCD2及びBRCA1タンパク質の存在に関して試験した(図7)。我々はウサギポリクローナル抗FANCD2抗体E35が、パキネマ後期(図7a)及びディプロネマ期(図7d, 7e及び7g)のX及びY染色体の非対軸を特異的に染色することを見出した。同一の実験条件下で、免疫前血清は対合複合体を染色しなかった(図7b及び7c)。M118抗BRCA1抗体はMスツパキテン期及びディプロテン期精母細胞の非対性染色体を染色した(図7f及び7h)。性染色体の対合していない軸のFANCD2抗体染色が中断され、系に通したビーズ様の外観を呈した(図7g)。20個のパキテン期の核の一貫した試験により、これらの抗FANCD2核フォーカスの多くが(~65%)強度な抗BRCA1染色の領域と共に同時局在することが示され、さらにこれらのタンパク質間の相互作用が支持された(図7g, 7h及び7i)。これらの結果により、クロマチンに結合するFANCD2タンパク質(活性化FANCD2)の第1の実例が提供される。

40

50

## 【 0 1 3 0 】

実施例 9 : F A N C D 2 の D N A 及びタンパク質配列を入手及び分析するための実験プロトコル

ノーザンハイブリダイゼーション。ヒト成人及び胎児の多組織 m R N A プロットをクローンテックから購入した ( パロ・アルト、カリフォルニア州 )。プロットを E S T クローン S G C 3 4 6 0 3 からの <sup>3</sup> <sup>2</sup> P 標識 D N A でプローブした。標準的なハイブリダイゼーション及び洗浄条件を使用した。プロットをアクチン c D N A プローブで再ハイブリダイズすることにより同等の負荷が確認された。

## 【 0 1 3 1 】

変異分析。全細胞 R N A を市販のキット ( ギブコ / B R L ) を用いて逆転写した。ネステッド P C R プロトコルを用いて得られた患者及び対照 c D N A から F A N C D 2 の 5 ' 末端部分を増幅した。第 1 ラウンドをプライマー ( 配列番号 : 9 7 ) M G 4 7 1 5 ' - A A T C G A A A A C T A C G G G C G - 3 ' 及び ( 配列番号 : 9 8 ) M G 4 5 7 5 ' - G A G A A C A C A T G A A T G A A C G C - 3 ' を用いて実施した。このラウンドからの P C R 産物を、プライマー ( 配列番号 : 9 9 ) M G 4 9 2 5 ' - G G C G A C G G C T T C T C G G A A G T A A T T T A A G - 3 ' 及び ( 配列番号 : 1 0 0 ) M G 4 7 2 5 ' - A G C G G C A G G A G G T T T A T G - 3 ' を用いる次のラウンドのために 1 : 5 0 に希釈した。P C R 条件は以下のとおりであった : 9 4 で 3 分、並びに 9 4 で 4 5 秒、5 0 で 4 5 秒、7 2 で 3 分及び最後に 7 2 で 5 分の 2 5 サイクル。プライマー ( 配列番号 : 1 0 1 ) M G 4 7 4 5 ' - T G G C G G C A G A C A G A A G T G - 3 ' 及び ( 配列番号 : 1 0 2 ) M G 4 7 5 5 ' - T G G C G G C A G A C A G A A G T G - 3 ' を用いる以外は前記したように遺伝子の 3 ' 部分を増幅した。P C R の第 2 ラウンドを ( 配列番号 : 1 0 3 ) M G 4 9 1 5 ' - A G A G A G C C A A C C T G A G C G A T G - 3 ' 及び ( 配列番号 : 1 0 4 ) N G 4 7 6 5 ' - G T G C C A G A C T C T G G T G G G - 3 ' を用いて実施した。P C R 産物をゲル精製し、p T - A d v ベクター ( クローンテック ) にクローン化し、内部プライマーを用いてシーケンシングした。

## 【 0 1 3 2 】

アレル特異的アッセイ : アレル特異的アッセイを P D 2 0 ファミリー及び 2 9 0 個の対照サンプル ( = 5 8 0 染色体 ) で実施した。P D 2 0 ファミリーは混合北欧系の家系であり、V U 0 0 8 はオランダの家系である。対照 D N A サンプルは C E P H ファミリー ( n = 9 5 ) の関連性のない個体からのものであり、サンプルは外胚葉性異形成症 ( n = 9 5 ) またはファンコーニ貧血 ( n = 9 4 ) のいずれかを有する関連性のない北米の家系からのものであった。P D 2 0 ファミリーの母性 n t 3 7 6 a g 変異は新規 M s p I 制限部位を作製した。ゲノム D N A に関しては、アッセイはエクソン 4 に位置する ( 配列番号 : 1 0 5 ) M G 7 9 2 5 ' - A G G A G A C A C C C T T C C T A T C C - 3 ' 及びイントロン 5 にある ( 配列番号 : 1 0 6 ) M G 8 0 3 5 ' - G A A G T T G G C A A A A C A G A C T G - 3 ' プライマーを用いてゲノム D N A を増幅することを含んだ。P C R 産物の大きさは 3 4 0 b p であり、これは変異が存在する場合、M s p I 消化時に 2 8 3 b p 及び 5 7 b p の 2 つのフラグメントを生じた。復帰した c D N A クローンの分析のために、プライマー ( 配列番号 : 1 0 7 ) M G 9 2 4 5 ' - T G T C T T G T G A G C G T C T G C A G G - 3 ' 及び ( 配列番号 : 1 0 8 ) M G 7 5 3 5 ' - A G G T T T T G A T A A T G G C A G G C - 3 ' を用いて P C R を実施した。P D 2 0 の父性エクソン 3 7 変異 ( R 1 2 3 6 H ) 及び V U 0 0 8 のエクソン 1 2 ミスセンス変異 ( R 3 0 2 W ) をアレル特異的オリゴヌクレオチド ( A S O ) ハイブリダイゼーションにより試験した ( W u ら、D N A 8 : 1 3 5 - 1 4 2 ( 1 9 8 9 ) )。エクソン 1 2 アッセイのために、ゲノム D N A をプライマー ( 配列番号 : 1 0 9 ) M G 9 7 9 5 ' - A C T G G A C T G T G C C T A C C C A C T A T G - 3 ' 及び ( 配列番号 : 1 1 0 ) M G 9 8 4 5 ' - C C T G T G T G A G G A T G A G C T C T - 3 ' で増幅した。プライマー ( 配列番号 : 1 7 1 ) M G 8 1 8 5 ' - A G A G G T A G G G A A G G A A G C T A C - 3 ' 及び ( 配列番号 : 1 7 2 ) M G 8 1 3 5 ' - C C A A A G T C C A C T T C T T G A A G - 3 ' を工

10

20

30

40

50

クソン37に関して使用した。野生型(配列番号:111)(R302Wに関しては5'-TTCTCCCGAAGCTCAG-3'及びR1236Hに関しては(配列番号:112)5'-TTTCTTCCGTGTGATGA-3')及び変異体(配列番号:111(R302Wに関しては5'-TTCTCCCAAGCTGAG-3'及びR1236Hに関しては配列番号:112)5'-TTTCTTCCATGTGATGA-3')オリゴヌクレオチドを<sup>3</sup>2P-[ATP]で末端標識し、新規DdeI部位として以前のようにドットプロット標的PCR産物にハイブリダイズした。野生型PCR産物を117及び71bp産物に消化し、一方変異アレルは56、61及び71bpの長さ3つのフラグメントを生じた、前記のアッセイの全てでPCRはゲノムDNA 50ngで、94で25秒、50で25秒及び72で35秒の37サイクルを実施した。

10

## 【0133】

抗FANCD2抗血清の作製:抗原供給源としてGST-FANCD2(N-末端)融合タンパク質を用いて、FANCD2に対するウサギポリクローナル抗血清を作製した。ポリメラーゼ連鎖反応(PCR)によりFANCD2 cDNA全長から、プライマー(配列番号:113)DF4EcoRI(5'-AGCCTCgaattcGTTTCCA A A A G A A G A C T G T C A - 3')及び(配列番号:114)DR816Xh(5'-GGTATCctcgaGTC A A G A C G A C A A C T T A T C C A T C A - 3')を用いて5'フラグメントを増幅した。FANCD2ポリペプチドのアミノ末端272個のアミノ酸をコードする、得られた841bpのPCR産物をEcoRI/XhoIで消化し、プラスミドpGEX4T-1(ファルマシア)のEcoRI/XhoI部位にサブクローニングした。予測した大きさ(54kD)のGST-FANCD2(N末端)融合タンパク質を大腸菌(E.coli)DH5株に発現し、グルタチオン-S-セファロース上で精製し、ニュージーランド白ウサギの免疫に使用した。FANCD2特異的免疫抗血清を、GSTタンパク質を負荷したアミノリンク・プラス・カラム(ピアース)上で、及びGST-FANCD2(N末端)融合タンパク質を負荷したアミノリンク・プラス・カラム(ピアース)上でアフィニティー精製した。

20

## 【0134】

実施例1に記載したのと同様のイムノプロットティング

細胞系及びトランスフェクション:オレガノ・ヘルス・サイエンシズ・ファンコーニ貧血細胞貯蔵所により作成された、PD20iは不死化された、及びPD733は1次繊維芽細胞系である(Jakobsら、Somert.Cell.Mol.Genet.22:151-157(1996))。PD20リンパ芽球は骨髓サンプルに由来した。欧州ファンコーニ貧血登録所(EUFAR)により作製された、VU008はリンパ芽球であり、VU423は繊維芽球系である。VU423iはSV40T抗原でのトランスフェクション(Jakobsら、1996)及びテロメラーゼ(Bodnarら、Science 279:349-352(1998))により誘導された不死化系である。別のFA細胞系は以前に記載されている。ヒト繊維芽細胞をMEM及び20%ウシ胎児血清中で培養した。形質転換されたリンパ芽球を、15%熱不活性化ウシ胎児結成を補充したRPMI 1640中で培養した。

30

## 【0135】

FANCD2発現構築物を作製するために、pBluescriptでクローン化されたRT-PCR産物からcDNA全長を集め、PCR誘導変異の不在をシーケンシングにより確認した。発現ベクターpIRES-Neo、pEGFP-N1、pRevTRE及びpRevTet-offはクロンテック(パロ・アルト、カリフォルニア州)のものである。FANCD2をこれらのベクターの適当な多クローニング部位に挿入した。標準的な条件を用いてPD20細胞系及び正常な対照繊維芽細胞系であるFM639に発現構築物を電気穿孔処理した(vanderHoffら、1992)。400µg/mlの活性G418でネオマイシン選別を実施した(ギブコ)。

40

## 【0136】

全細胞融合:全細胞融合実験のために、ヒグロマイシンBに抵抗し、HPRT遺伝子座

50

を欠失したPD20細胞系(PD20i)を使用した(Jakobsら、Somat. Cell. Mol. Genet. 23: 1-7 (1997))。対照はPD24(PD20の罹患同胞からの1次繊維芽細胞)及びPD319i(Jakobsら、1997)(非A、C、DまたはG FA患者からの不死化繊維芽細胞)を含有した。各細胞系からの $2.5 \times 10^5$ セルをT25フラスコで混合し、24時間回復させた。細胞を血清不含培地で洗浄し、次いで50%PEGで1分間融合させた。PEGを除去した後、細胞を血清不含培地で3回洗浄し、選別しない完全培地中で一晚回復させた。翌日、400 $\mu$ g/mlヒグロマイシンB(ロッシュ・モレキュラー)及び1X HATを含有する選択培地に細胞を1:10で分配した。選別が完了した後、ハイブリッドを1回継代し、次いで以下に記載するように分析した。

10

## 【0137】

FA-D2細胞のレトロウイルス形質導入及び補足分析: FANCD2 cDNA全長をベクターpMMP-puroにサブクローニングした(Pulsipherら、1998)。レトロウイルス上澄を用いてPD20Fを形質導入し、ピューロマイシン抵抗性細胞を選択した。クリスタルバイオレットアッセイによりMMC感受性に関して細胞を分析した(Nafら、1998)。

## 【0138】

染色体切断分析: OHSU(ポートランド、オレゴン州)でサイトジェネティクス・コア・ラブにより染色体切断分析を実施した。分析のために(Cohenら、1982)、細胞をT<sub>25</sub>フラスコにプレートし、回復させ、次いで300ng/mlのDEBで2日間処置した。処置後、細胞をコルセミドに3時間暴露し、0.075M KCl及び3:1メタノール:酢酸を用いて収穫した。スライドをライト染料で染色し、50-100分裂中期をラジアルに関してスコアを取った。

20

## 【0139】

実施例10: 可能性のある治療薬のスクリーニングに使用するためのFAに関するマウスモデル

D'Andreaら、90: 1725-1736(1997)、及びYangら、Blood 98: 1-6(2001)に記載されるように胚幹細胞における相同性組換えまたは標的破壊を用いてFANCD2のネズミモデルを作製できる。マウスのFANCD2遺伝子座のノックアウトは致死変異ではない。これらのノックアウト動物は癌に対する感受性が増しており、さらにFAに特徴的な別の症状を展示する。特定の治療薬をノックアウトマウスに投与することはマウスの癌に対する感受性を低減させると予測される。さらに、特定の確立された化学療法剤は特定の遺伝子欠損の結果として癌を進行させたノックアウトマウスを処置するのにさらに有効であると同定され、これはまた癌に対する感受性を有するヒト対象またはFANCD2遺伝子座の変異の結果として癌を進行させたヒト対象の処置に有用であると予測される。

30

## 【0140】

我々は例えば遺伝子のエクソンに破壊を創ったFanc cに関して、Chenら、Nat. Genet. 12: 448-451(1996)により、及び相同性組換えを用いて遺伝子のエクソンに破壊を創ったWhitneyら、88: 49-58(1996)により記載された研究法を用いて、FANCD2の標的破壊で実験マウスモデルを作製できる。双方の動物モデルでは自発的染色体切断及び2官能基アルキル化剤に応答した脾臓リンパ球における染色体切断の増加が観察される。双方のモデルでは、Fanc d2 - / -マウスは生殖細胞欠損及び生殖能力の低下を有する。Fanc d2ネズミノックアウトモデルは(1)DNA損傷に対する造血細胞の生理学的応答におけるFanc d2遺伝子の役割、(2)FA骨髓細胞における抑制性サイトカインのインビボの影響を試験するのに有用である。129/Sv及びC57BLを以下の標準的なプロトコルに従って作製できる。マウステールゲノムDNAを以前に記載したように調製し、ポリメラーゼ連鎖反応(PCR)遺伝子型同定の鋳型として使用できる。

40

## 【0141】

50

公知の F a n c d 2 遺伝子型の 6 週齢マウスから脾臓細胞を調製できる。脾臓を切開し、R P M I 培地中ですり潰し、単一細胞懸濁液にし、70 μm フィルターを通した。赤血球を低張性塩化アンモニウム中で溶解する。得られた脾臓リンパ球をリン酸塩緩衝生理食塩水で洗浄し、R P M I / 10% ウシ胎児血清 + フィトヘマグルチニンに懸濁した。トリパンブルー排除アッセイにより生存性に関して細胞を試験する。培地中細胞を 24 時間培養し、さらに 48 時間 M M C または D E B に暴露する。また別に 50 時間細胞を培養し、I R (指示するように 1 または 4 G y) に暴露し、染色体切断またはトリパンブルー排除 (生存性) 分析の前に 12 時間回復させる。

#### 【0142】

以前に記載したように、4 から 6 週齢の F a n c d 2 + / - または F a n c d 2 - / - マウスの大腿及び脛骨から単核細胞を単離できる。M e t h o C u l t M 3 4 3 培地 (ステムセル・テクノロジーズ、バンクーバー、ブリティッシュコロンビア州) 1 m l 中、M M C 処置を伴うかまたは伴わずに、2 × 10<sup>4</sup> セルの全てを培養した。7 日にコロニーを計数し、この時たいていのコロニーは顆粒球・マクロファージコロニー形成ユニットまたは赤芽球バースト形成ユニット系列に属する。各々の数は 2 検体ずつのプレートからの平均であり、データは 2 つの別個の実験から誘導された。

#### 【0143】

胸腺、脾臓及び末梢リンパ節から単離されたリンパ球を、フルオレセインイソチオシアネート抱合抗 C D 3、C D 4 及び C D 19、並びに P E 抱合抗 C D 8、C D 44、C D 45 B、免疫グロブリン M、及び B 220 (B D ファルミンゲン、カリフォルニア州) で T - または B - リンパ球表面分子に関して染色する。染色された細胞をカウンター・エピックス X L フロー・サイトメトリー・システムで分析した。

#### 【0144】

マウス卵巣及び精巣を単離し、4% パラホルムアルデヒドで固定し、マサチューセッツ総合病院の病理部門の中心的施設でさらに処理した。

#### 【0145】

実施例 11: ヒト対象における癌感受性の増加を検出するための抗体試薬を用いるスクリーニングアッセイ

F A N C D 2 - L に比較した F A N C D 2 - S の相対量及び F A N C D 2 - L の存在または不在に関して試験するために、血液サンプルまたは組織サンプルを対象から採取できる。F A N C D 2 - S 及び F A N C D 2 - L タンパク質に特異的な抗体試薬を用いて (実施例 1)、図 14 に示すウェスタンプロットで陽性サンプルを同定できる。別の抗体アッセイ、例えば 5654162 及び 5073484 に記載される 1 工程遊走結合バンドアッセイを利用できる。酵素結合免疫吸着アッセイ (E L I S A)、サンドウィッチアッセイ、ラジオイムノアッセイ及びその他の技術分野で公知の免疫診断アッセイを用いて F A N C D 2 - S 及び F A N C D 2 - L の相対結合濃度を決定できる。

#### 【0146】

この研究法の実行可能性を以下に説明する:

ヒト癌細胞系をスクリーニングするための F A N C D 2 診断ウェスタンプロット

ヒト癌細胞系を電離放射線を伴って、または伴わずに (図 14 に示すように) 処置し、全細胞タンパク質を電気穿孔処置し、ニトロセルロースに移し、実施例 1 の抗 F A N C D 2 モノクローナル抗体でイムノプロットした。卵巣癌細胞系 (T O V 2 1 G) は F A N C D 2 - S を発現したが、F A N C D 2 - L を発現しなかった (レーン 9、10 参照)。この細胞系は F A N C D 2 遺伝子を重複するヒト染色体 3 p の欠失を有し、F A N C D 2 に関して半接合性であり、第 2 の F A N C D 2 アレルで変異を有すると予測され、従って F A 複合体によりモノユビキチン化に失敗し、よって F A N C D 2 - L が不在 (レーン 9、10)。この実施例は、抗体基盤の試験が癌感受性の増加に至る F A N C D 2 遺伝子の損傷を決定するのに適していることを示している。

#### 【0147】

実施例 12: ヒト対象における癌感受性の増加を検出するための核酸試薬を用いるスク

10

20

30

40

50

### リーニングアッセイ

血液サンプルまたは組織サンプルを対象から採取し、シーケンシング技術または核酸プローブを用いてスクリーニングし、対象のゲノムにあるとすれば、遺伝子損傷の大きさ及び位置を決定する。スクリーニング方法には全遺伝子のシーケンシング、またはプローブのセットもしくは単一のプローブを使用して損傷を同定することによるものなどがある。単一の損傷は集団において優勢であるが、別の遺伝的条件、例えば嚢胞性繊維症及びP53腫瘍抑制遺伝子の場合にあるように、低頻度で遺伝子中いたるところで別の損傷を生じ得る。

#### 【0148】

この研究法の実行可能性を以下に説明する：

標準的なフィコール・ハイパーク・グラジエントを用いて患者から末梢血リンパ球を単離し、ゲノムDNAをこれらのリンパ球から単離する。我々はゲノムPCRを用いてヒトFANCD2遺伝子の44個のエクソン(表7参照)を増幅し、2つのFANCD2アレルをシーケンシングして変異を同定する。かかる変異が見出される場合、FA-D2インディケーター細胞系の機能的補足を排除する能力によりこれらを良性多型性から区別する。

#### 【0149】

【表 1】

表 1 ファンコーニ貧血の相補群及び原因遺伝子

サブタイプ	患者の推定パーセンテージ	原因遺伝子	染色体位置	エクソン数	タンパク質産物
A	66%	FANCA	16q24.3	43	163Kd
B	4.3%	FANCB	-	-	-
C	12.7%	FANCC	9q22.3	14	63Kd
D1	まれ	FANCD1	-	-	-
D2	まれ	FANCD2	3p25.3	44	155, 162kD
E	12.7%	FANCE	6p21.2-21.3	10	60kD
F	まれ	FANCF	11p15	1	42kD
G	まれ	FANCG (XRCC9)	9p13	14	68kD

【 0 1 5 0 】

10

20

30

【表2】

表2

ゲノム不安定性の疾患

疾患	損傷手段	腫瘍	機能
FA	架橋化剤、 白血病、肝臓、胃腸 及び婦人科の腫瘍	急性骨髄芽球性	未知
XP	紫外線	扁平上皮癌	除去修復
AT	電離放射線	リンパ腫	p53への 求心性経路
ブルーム 症候群	アルキル化剤	急性リンパ芽球白血病	細胞周期制御
コケーン 症候群	紫外線	基底細胞癌修復	転写共役
遺伝性非 ホリニア性 結腸癌 (HNPCC)	未知	結腸の腺癌、卵巣癌	DNA誤対合修復

10

20

30

【表 3】

表 3

FANCD2 配列変化	
変異	
PD20	nt376a→g S126G/ 接合
	nt3707g→a R1236H
VU008	nt904c→t R302W
	nt958c→t Q320X
PD733	エクソン17の欠失
多型性	
	nt1122a→g V374V
	nt1440t→c <sup>*</sup> H480H
	nt1509c→t <sup>†</sup> N503N
	nt2141c→t <sup>*</sup> L714P
	nt2259t→c D753D
	nt4098t→g <sup>†</sup> L1366L
	nt4453g→a <sup>†</sup> 3UTR

\* PD20 は異型接合性である；  
† VU008 は異型接合性である。

【 0 1 5 2 】

10

20

30

【表 4】

表 4  
全細胞融合の染色体切断分析

細胞系/ハイブリット	DEB (ng/ml)	MMC (ng/ml)	ラジアルを有する 細胞の%	表現型
PD20i	300		58	S
PD24p	300		na*	S
VU423p	300		na*	S
PD319i	300		52	S
PD20i/VU423p	300		6	R
PD20i/PD24p	300		30	S
PD20i/PD319i	300		0	R
PD20i		40	48	S
VU423i		40	78	S
PD20i/VU423i		40	10	R
VU423i + chr. 3, クローン1		40	74	S
VU423i + chr. 3, クローン2		40	68	S
VU423i + chr. 3, クローン3		40	88	S
PD20i + 空 ベクター	0	0	2	
		40	24	S
	200		62	S
PD20i + FANCD2ベクター	0	0	0	
		40	2	R
	200		10	R

実験の群を行間で分ける。S、架橋化剤感受性；R、架橋化剤抵抗性；

i = 不死化繊維芽細胞系；p = 1次繊維芽細胞

\* この濃度での細胞生存性が低すぎてラジアル形成に関してスコアを取れず、これはDNA鎖間架橋に対する1次繊維芽細胞の鋭敏な感受性を示している。

【 0 1 5 3 】

【表 5】

表 5					
細胞系/プラスミド		F A 群	F A タンパク質 複合体 (1)	MMC 感受性 (2)	IR/ブレオマイシン 感受性 (3)
リンパ芽球	PD7	Wt	+	R	R
	HSC72	A	-	S	
	HSC72+A	A	+	R	
	PD4	C	-	S	
	PD4+C	C	+	R	
	EUFA316	G	-	S	
	EUFA316+G	G	+	R	
	EUFA121	F	-	S	S
	EUFA121+F	F	+	R	R
	PD20	D	+	S	S
	PD20(R)	D	+	R	R
繊維芽細胞	GM0637	Wt	+	R	R
	GM6914	A	-	S	S
	GM694+A	A	+	R	R
	PD426	C	-	S	
	PD426+C	C	+	R	
	FAG326SV	G	-	S	
	FAG326SV+G	G	+	R	
	PD20F	D	+	S	S
	20-3-15 (+D)	D	+	R	R
	NBS (-/-)	NBS	+	S	S
	ATM (-/-)	ATM	+	S	S
	BRCA1 (-/-)	BRCA1	+	S	S

- 1) F A タンパク質複合体 (FANCA/FANCG/FANCC) の存在を以前に報告された様に決定した。(Garcia-Higueraら、MCB 19: 4866-4873、1999)
- 2) X T T アッセイにより決定されたリンパ芽球に関する、またはクリスタルバイオレットアッセイにより決定された繊維芽細胞に関するMMC感受性。
- 2) 染色体切断の分析によりIR/ブレオマイシン感受性を決定した(材料及び方法を参照)。

【 0 1 5 4 】

10

20

30

40

【表 6】

表 6 FANCDのイントロン/エクソン接合部

エクソン	サイズ*	配列番号	5'-供与部位	イントロン	3'-受容部位	エクソン
1	30	9	TTC	gtgagtaagt	gtttcccgatttttgctctag	85
2	97	10	CAA	gtaagratcta	gaaatattttctatttttccag	83
3	141	11	TAG	gtaaatattta	ctctctttttttctgcctag	88
4	68	12	AAA	gtagtattttt	atthtttaaatctctctaaag	78
5	104	13	CAG	gtgtggagagg	gattttctttttttttttacag	91
6	61	14	CAG	gtaagactgtc	ccctatgctctctttttttttag	86
7	53	15	AAA	gtaegyggcgt	ttctctcttaaacattttttag	80
8	79	16	AAG	gtaggctttatg	aatagtgctttctactgcag	85
9	125	17	CAG	gtggataaac	tthttttctaccattctcacag	86
10	88	18	AAG	gtgaaaagac	tttggtcttttaatttttttag	85
11	105	19	GAG	gtagtctctta	ctaaatatttactttctgcag	87
12	101	20	AAG	gtaagagcttc	ttctctctctgctactgttag	84
13	101	21	AAG	gtgagatcttt	actctctctgttttttttag	92
14	36	22	AAG	gtaatgttcat	tgcatatttatagcaaatag	73
15	144	23	TTA	gtaegtgtcag	ttctactctcccccactcaag	86
16	135	24	CAG	gtatgttgaat	gttgactctccctgtatag	84
17	132	25	AAG	gtatctttatg	tggeatcatttttttccacag	89
18	111	26	CAG	gttagagagaa	ttctctactctcattgcag	87
19	110	27	CAG	gtacacgtgga	aaaaaattttttgttttttag	83
20	61	28	CAG	gtgagtctttt	attttctctctttttgtctcag	93
21	120	29	CTG	gtaaagccaat	tgtttggttgtctctcag	85
22	74	30	AGG	gtaggatattg	attctgggtttttctcag	88
23	147	31	AAA	gtcagtaagt	aatattctctctctctccag	89
24	101	32	TAG	gtatgggatga	aaatgtttgttctctctcag	86
25	116	33	GAG	gtgagccaggt	atgttaattgttactttgcag	82
26	109	34	CAG	gtaagagaggt	caagctgtgtttttgtttcag	81
27	111	35	TAG	gtaagtattgt	ttctcttttttaataaaag	73
28	110	36	AAG	gtattggagtg	ttgctgtgacttccccatag	85
29	144	37	GAA	gtaegtgtcag	ttctctctccatgtgcaag	84
30	117	38	AAG	gttaggttagg	taactctgcattttatttag	80
31	129	39	CAG	gtcagaagccl	aaaaatcatttttttttttag	79
32	119	40	TTG	gtaggtattgt	ttttacccttgactctcttag	85
33	111	41	CAG	gtgagtataaa	tttttctgtctctctcag	91
34	131	42	TTG	gtgatgggccc	tttgctctctttttctaacag	89
35	94	43	CTG	gtgagatgltt	atatttgactctcattgcag	78
36	123	44	CAG	gtaagggaggt	atgctttttccgctctcttag	88
37	94	45	CAG	gtgagtaagat	catatattttgtgtgccccag	81
38	72	46	AAG	gtgagtagga	cttgctctttccacctctccag	93
39	39	47	AAG	gtgagagattt	agtggtctctctctctcag	86
40	75	48	CGG	gtaagagctaa	kataaactattgggtatag	77
41	75	49	AAG	gtaagaagggg	tgattattttctcattcag	86
42	147	50	CAG	gtaagccttgg	cttggtccatctcaaatattag	80
43	228	51	CCA	taa + 3'UTR	atthttttttgtgcccccttag	80
44	72	51	GAG	gtatctcttaca		
			GAT	tag + 3'UTR		

【 0 1 5 5 】

10

20

30

【表 7】

表 7 : FANCDの44エクソンを増幅するためのプライマー

エクソン	プライマー名	配列番号	プライマー配列 (5'→3')	生成物のサイズ (bp)	アニーリング温度
1	MG914	115	F: CTAGCACAGAAGCTCTGCTGC	372	54
	MG837	116	R: CTAGCACAGAAGCTCTGCTGC		
2	MG746	117	F:	422	50
	MG747	118	CTTCAGCAACAGCGAAGTAGTCT G R: GATTCTCAGCACTTGAAAAGCAG G		
3	MG773	119	F: GGACACATCAGTTTTCTCTC	309	50
	MG789	120	R: GAAAACCCATGATTCAGTCC		
4-5	MG816	121	F: TCATCAGGCAAGAACTTGG	467	50
	MG803	122	R: GAAAGTTGGCAAAACAGACTG		
6	MG804	123	F: GAGCCATCTGCTCATTTCTG	283	50
	MG812	124	R: CCCGCTATTTAGACTTGAGC		
7	MG775	125	F: CAAAGTGTTTATTCCAGGAGC	343	50
	MG802	126	R: CATCAGGGTACTTTGAACATTC		
8-9	MG727	127	F: TTGACCAGAAAAGGCTCAGTTCC	640	50
	MG915	128	R: AGATGATGCCAGAGGGTTTATCC		
10	MG790	129	F: TGCCCAGCTCTGTTCAAACC	222	50
	MG774	130	R: AGGCAATGACTGACTGACAC		
11	MG805	131	F:	392	50
	MG791	132	TGCCCGTCTATTTTGTATGAAGC R: TCTCAGTTAGTCTGGGGACAG		
12	MG751	133	F:	432	50
	MG972	134	TCATGGTAGAGAGACTGGACTGT GC R: ACCCTGGAGCAAATGACAACC		
13-14	MG973	135	F: ATTTGCTCCAGGTTACATGGC	555	50
	MG974	136	R: GAAAGACAGTGGGAAGGCAAGC		
15	MG975	137	F:	513	50
	MG976	138	GGGAGTGTGTGGAACAAATGAGC R: AGTTTCTACAGGCTGGTCCTATTC C		
16	MG755	139	F: AACGTGGAATCCCATTGATGC	379	48
	MG730	140	R: TTTCGTGTTCCGTCCTTGC		
17	MG794	141	F: GATGGTCAAGTTACACTGGC	382	50
	MG778	142	R: CACCTCCCACCAATTATAGTATTC		
18	MG808	143	F: CTATGIGTGTCTTTTACAGGG	234	48
	MG817	144	R: AATCTTCCACCATATTGC		
19	MG779	145	F: CATACCTCTTTTGTCTGTGC	199	48
	MG795	146	R: CCACAGAAGTCAGAATCTCCACG		
20	MG731	147	F: TGTAACAAACCTGCACGTTG	632	56
	MG732	148	R: TGCTACCCAAGCCAGTAGTTCC		
21	MG788	149	F: GAGTTGGGAAAGATTGGCAGC	232	50
	MG772	150	R: TGTAGTAAAGCAGCTCTCATGC		
22-23	MG733	151	F: CAAGTACACTCTGCACTGCC	652	50
	MG758	152	R: TGACTCAACTTCCCCACCAAGAG		

10

20

30

40

24-25	MG736 MG737	153 154	F: R:	CTCCCTATGTACGTGGAGTAATAC GGGAGTCTGTGGAACTAAG	732	50
26	MG780 MG759	155 156	F: R:	TTCATAGACATCTCTCAGCTCTG GTTTTGGTATCAGGGAAAGC	284	50
27-28	MG760 MG781	157 158	F: R:	AGCCATGCTTGGAAATTTGG CTCACTGGGATGTCACAAAC	653	50
29	MG740 MG741	159 160	F: R:	GGTCTTGATGTGTGACTTGTATCC C CCTCAGTGTACAGTGTCTTTGT G	447	50
30	MG809 MG797	161 162	F: R:	CATGAAATGACTAGGACATTCC CTACCCAGTGACCCAAACAC	281	48
31-32	MG761 MG742	163 164	F: R:	CGAACCTTAGTTTCTGAGAGGC TCACTGCCCTTGGTACTGTC	503	50
33	MG916 MG810	165 166	F: R:	TTGATGGTACAGACTGGAGGC AAGAAAGTTGCCAATCCTGTCC	274	50
34	MG762 MG743	167 168	F: R:	AGCACCTGAAAAAAGGAGG GCCCAAAGTTGTAAAGTGTGAG	343	50
35-36	MG787 MG806	169 170	F: R:	AGCAAGAATGAGGTCAAGTTC GGGAAAACTGGAGGAAAGAACT C	590	50
37	MG818 MG813	171 172	F: R:	AGAGCTAGGGAAGGAAGCTAC GCCAAAGTCCACTTCTGAAG	233	50
38	MG834 MG836	173 174	F: R:	GATGCACTGGTTGCTACATC CCAGGCACTTGGTTTCTGC	275	50
39	MG839 MG871	175 176	F: R:	ACACTCCCAGTTGGAATCAG CTTGTGGGCAAGAAATTGAG	370	50
40	MG829 MG870	177 178	F: R:	TGGGCTGGATGAGACTATTC CCAAGGACATATCTTCTGAGCAA C	223	50
41	MG820 MG811	179 180	F: R:	TGATTATCAGCATAGGCTGG GATCCCCCAATAGCAACTGC	271	50
42	MG763 MG782	181 182	F: R:	CATTCAGATTCACCAGGACAC CCTTACATGCCATCTGATGC	227	50
43	MG764 3'UTR MG835	183 184	F: R:	AAGCTTCGCCCTATTACCC GGAAAATGAGAGGCTATAATGC	435	50
44	MG1006 3'UTR MG1005	185 186	F: R:	TGTATTCCAGAGGTCACCCAGAG C CCAGTAAGAAAGGCAAACAGCG	234	50

10

20

30

【 図 1 A 】

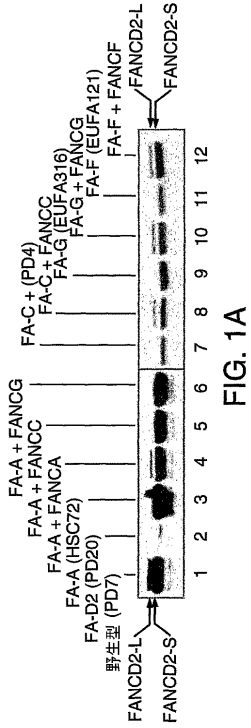


FIG. 1A

【 図 1 B 】

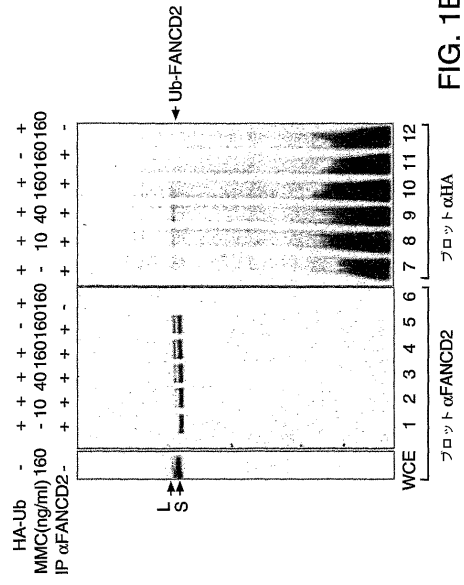


FIG. 1B

【 図 1 C 】

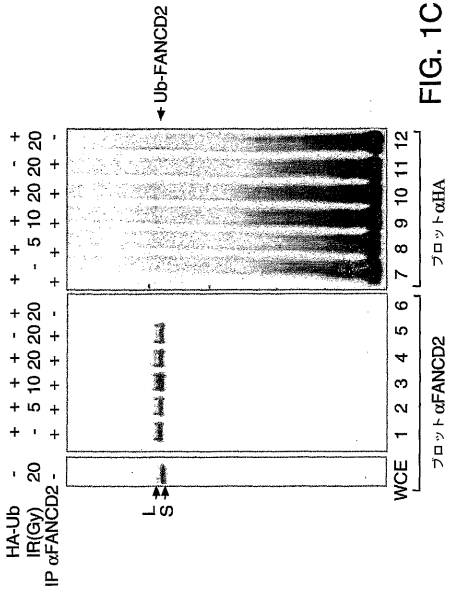


FIG. 1C

【 図 1 D 】

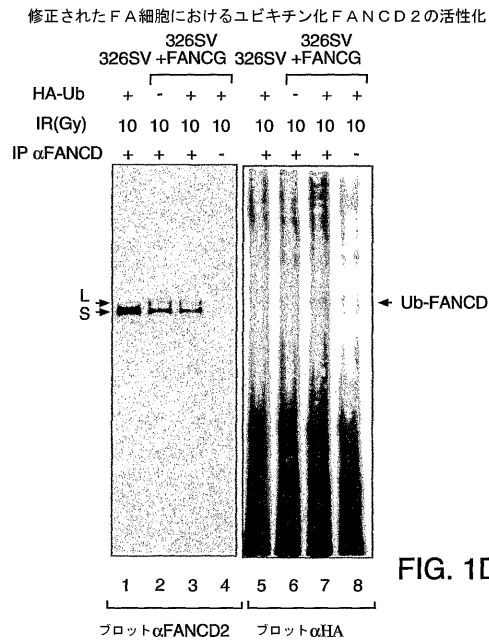


FIG. 1D

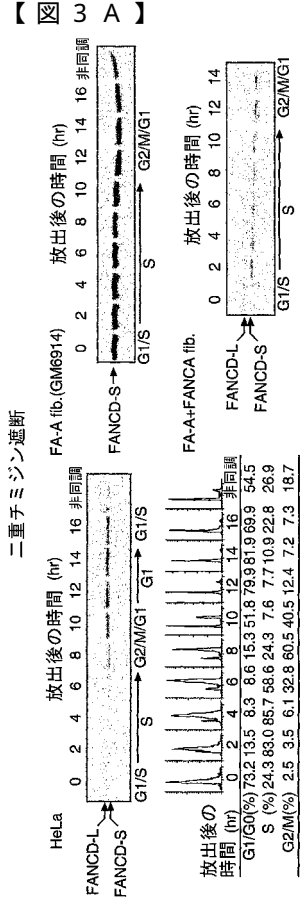
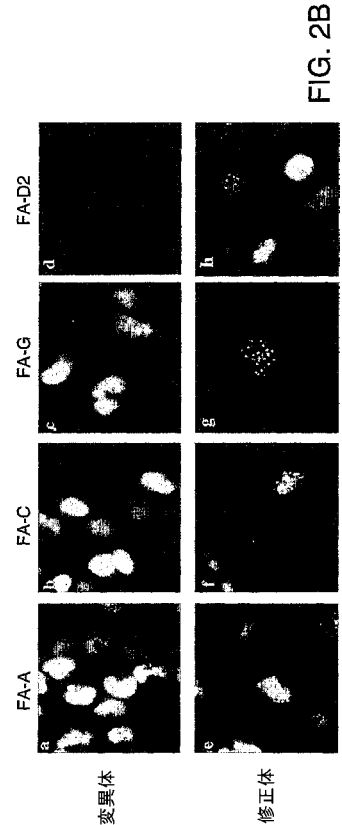
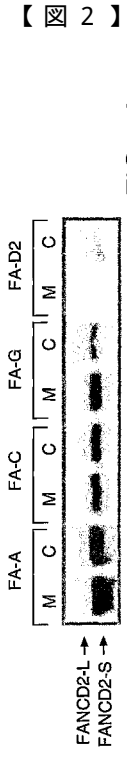
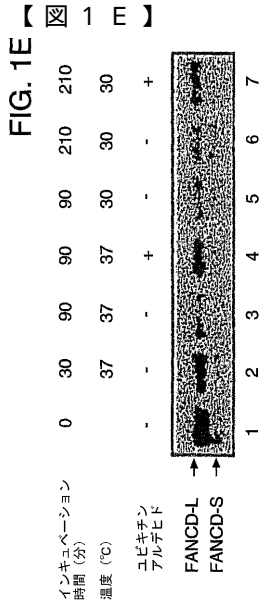


FIG. 3A

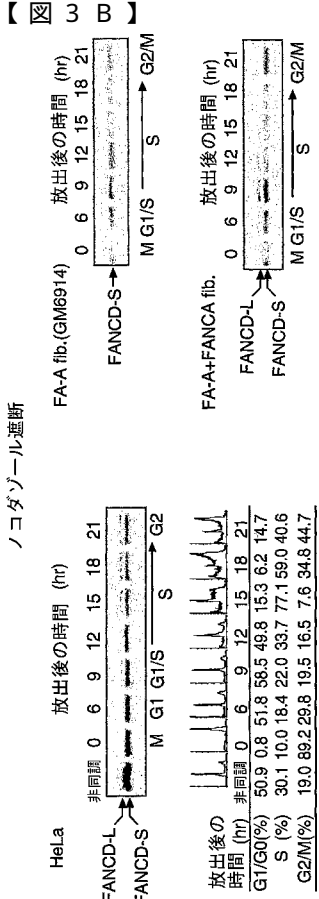
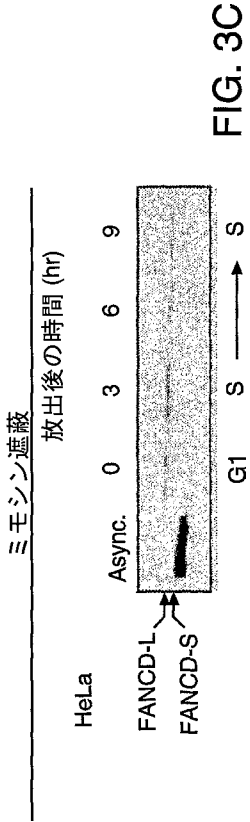
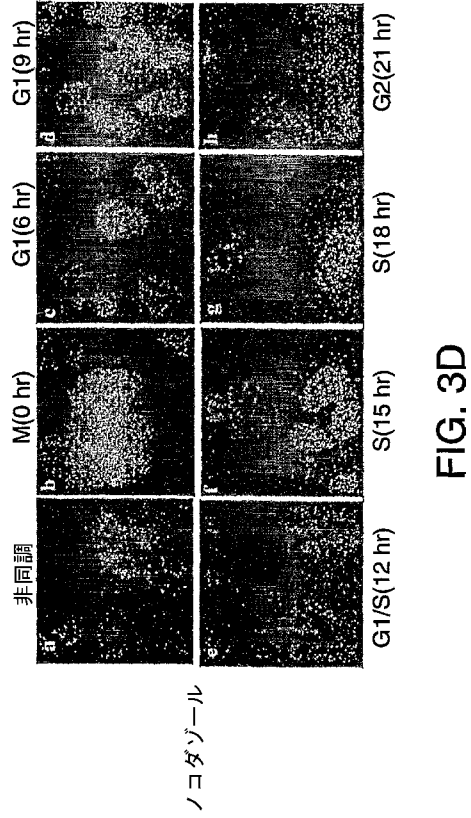


FIG. 3B

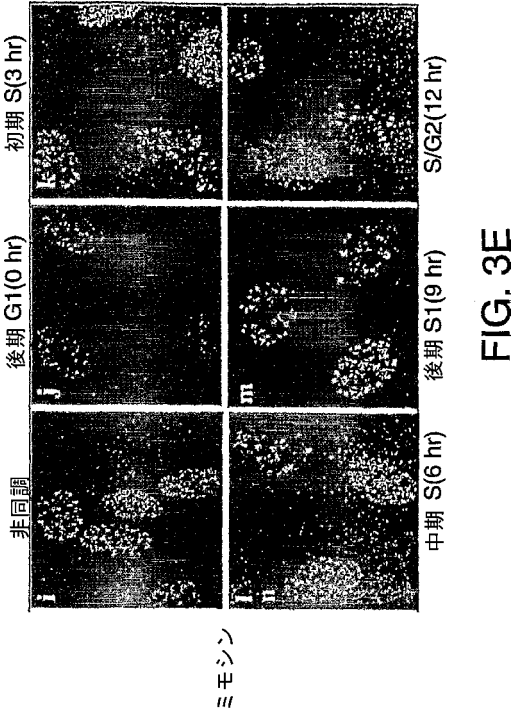
【 図 3 C 】



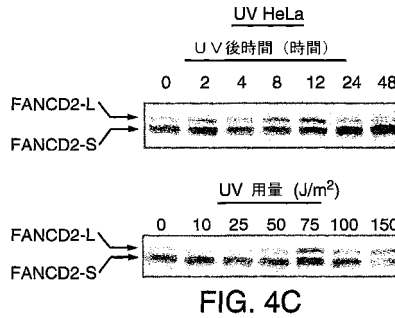
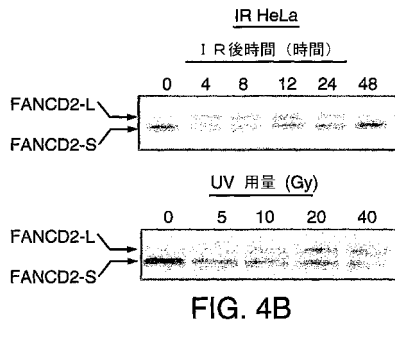
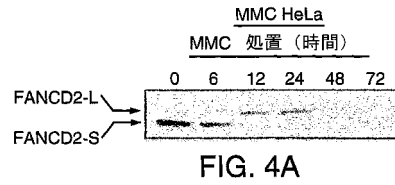
【 図 3 D 】



【 図 3 E 】



【 図 4 A - C 】



【 図 4 D 】

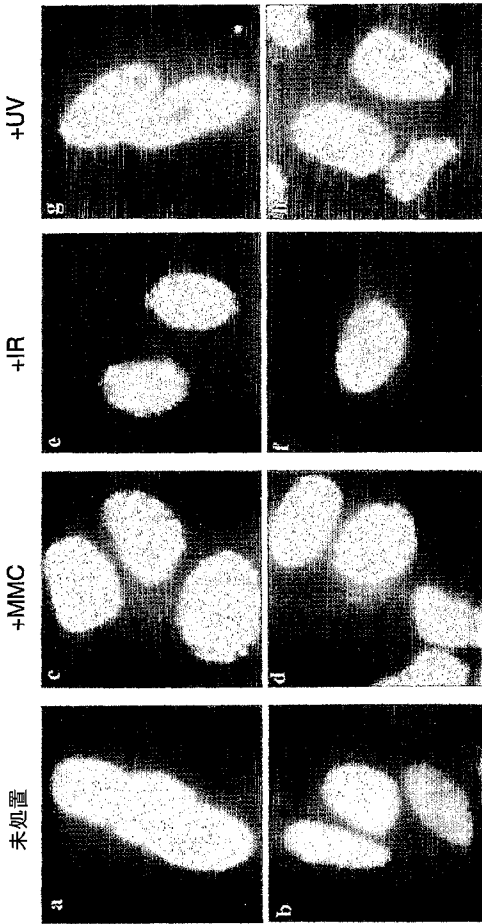


FIG. 4D

【 図 4 E 】

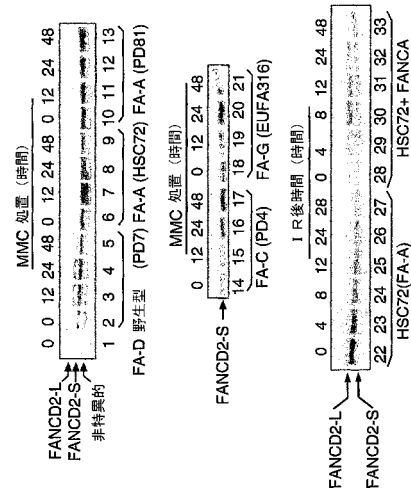
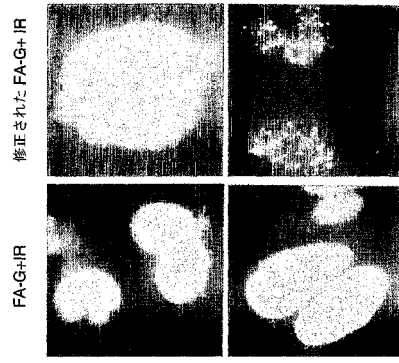


FIG. 4E

【 図 5 A 】

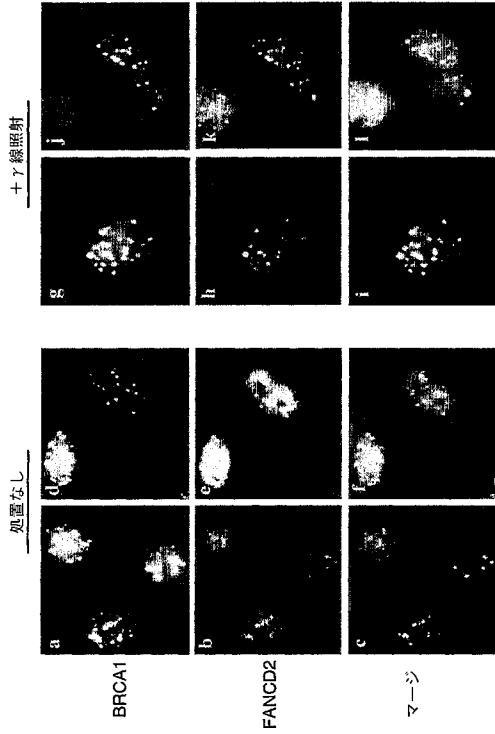


FIG. 5A

【 図 5 B 】

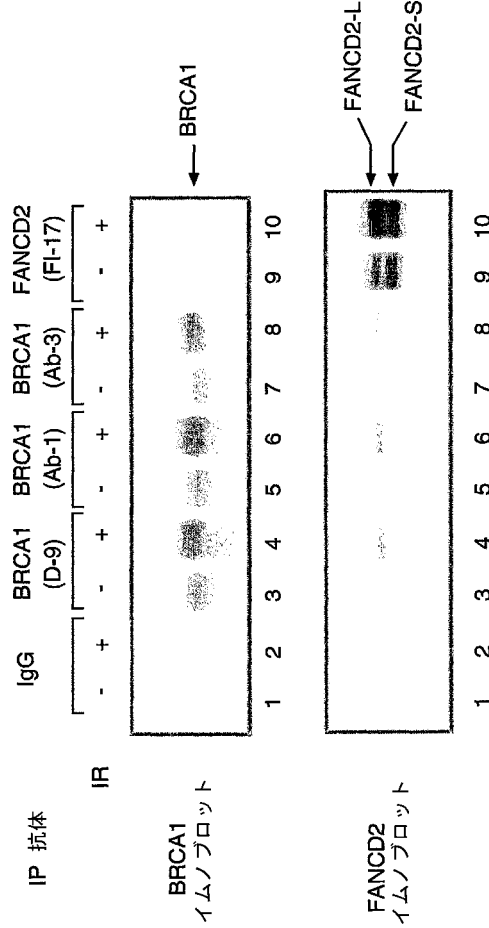


FIG. 5B

【 図 6 】

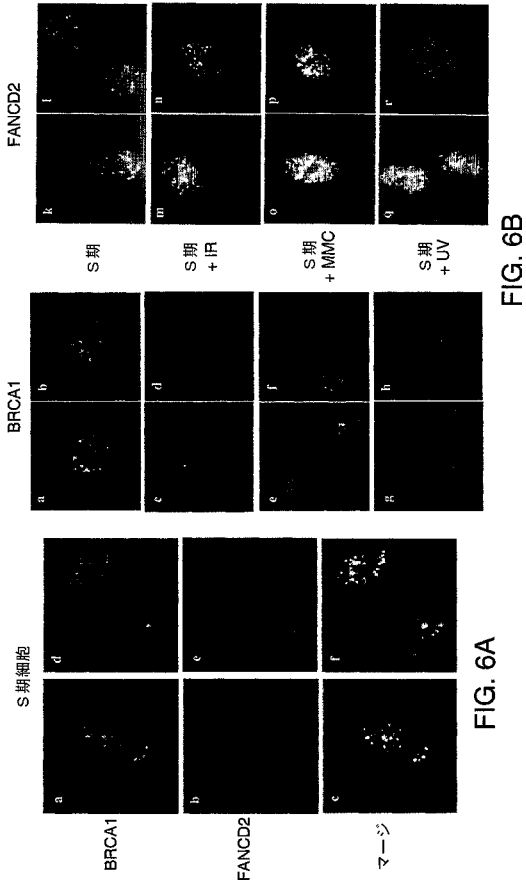


FIG. 6B

FIG. 6A

【 図 7 】

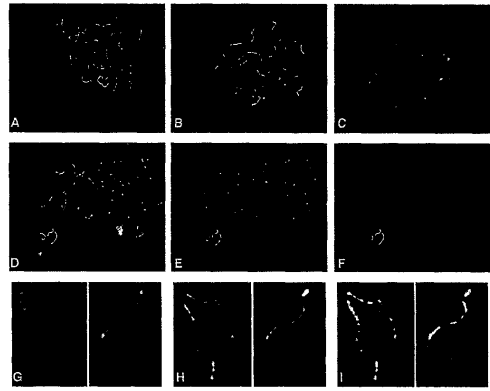


FIG. 7

【 図 8 】

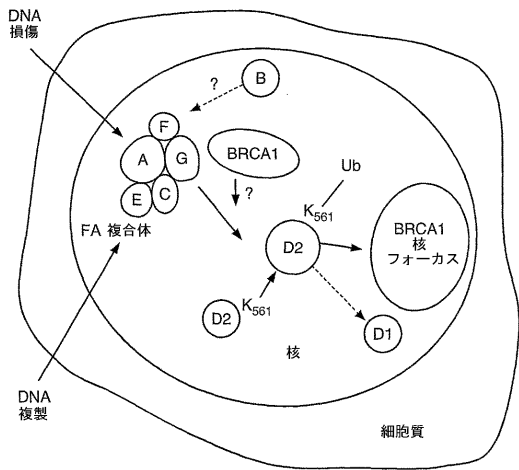


FIG. 8

【 図 9 】

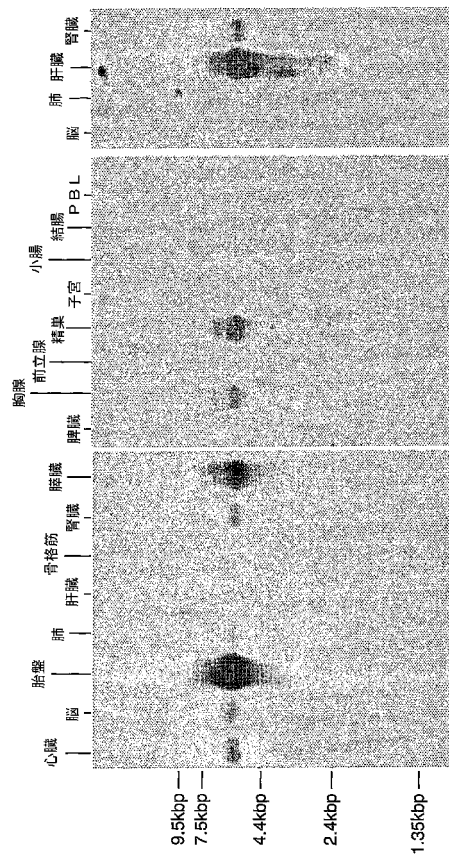


FIG. 9

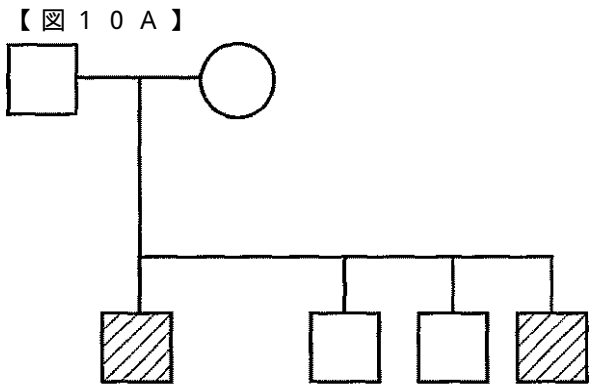


FIG. 10A

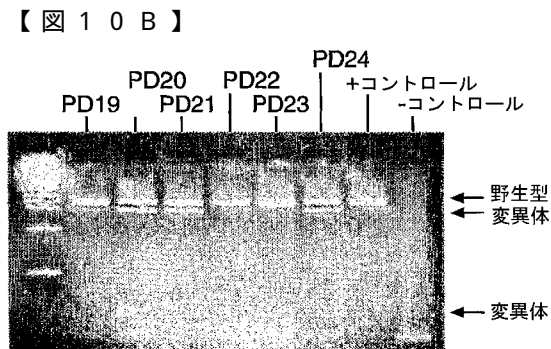


FIG. 10B

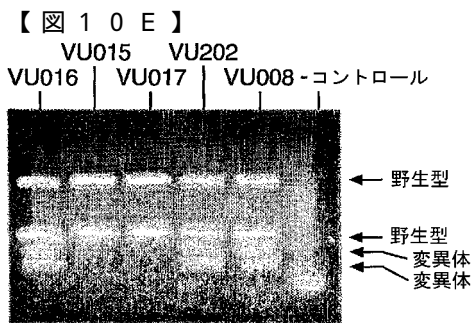


FIG. 10E

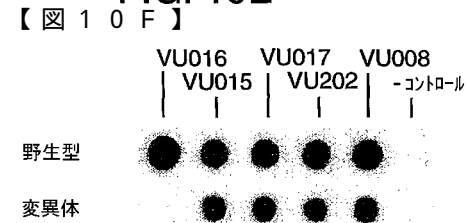


FIG. 10F

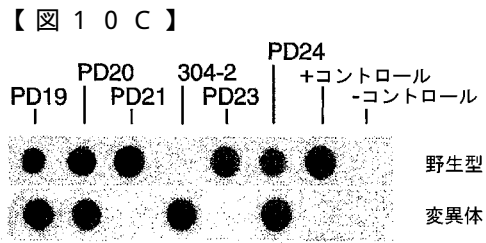


FIG. 10C

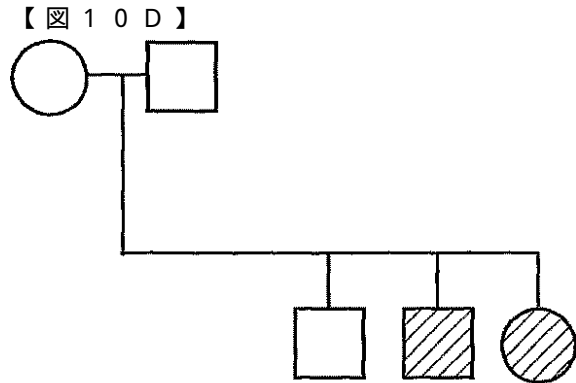
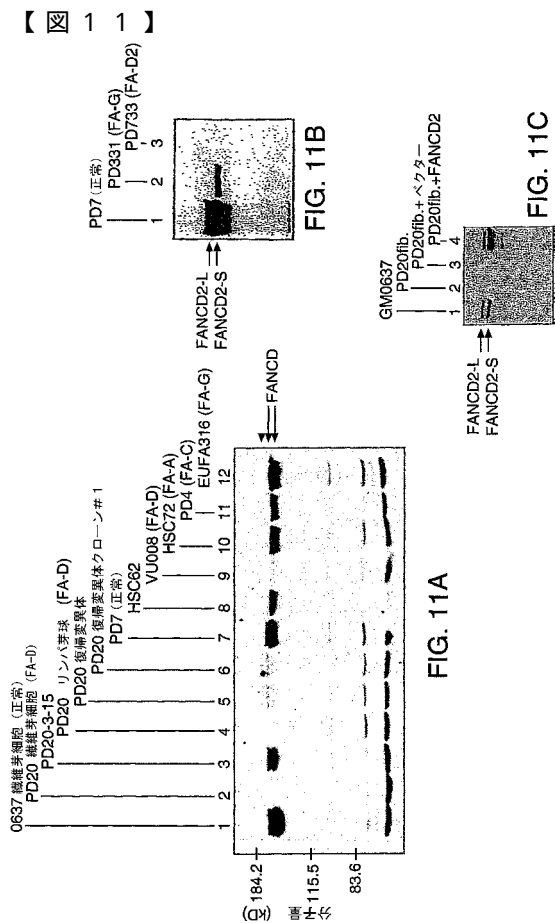


FIG. 10D



【 図 1 2 】

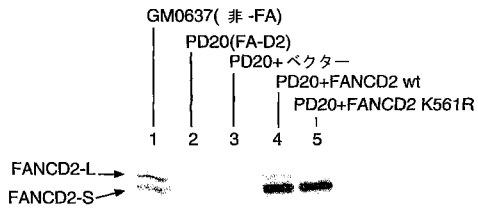


FIG. 12A

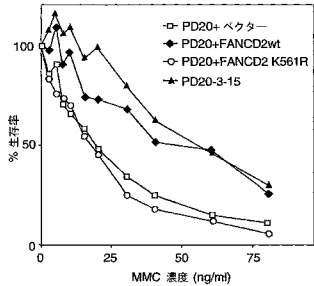


FIG. 12B

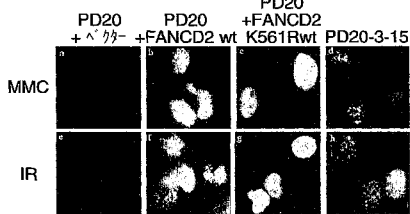
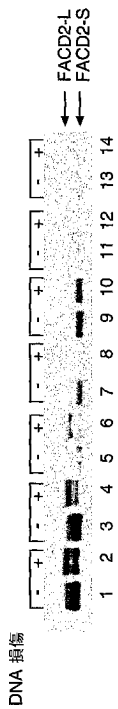


FIG. 12C

【 図 1 4 】

FIG. 14



【 図 1 3 】

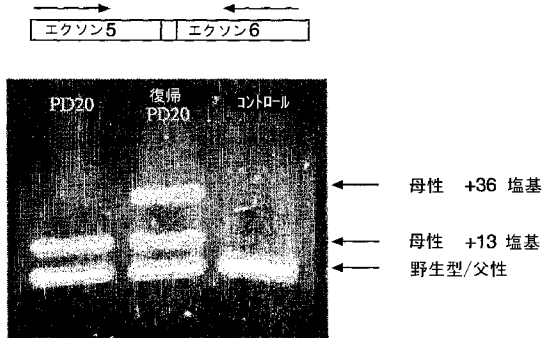
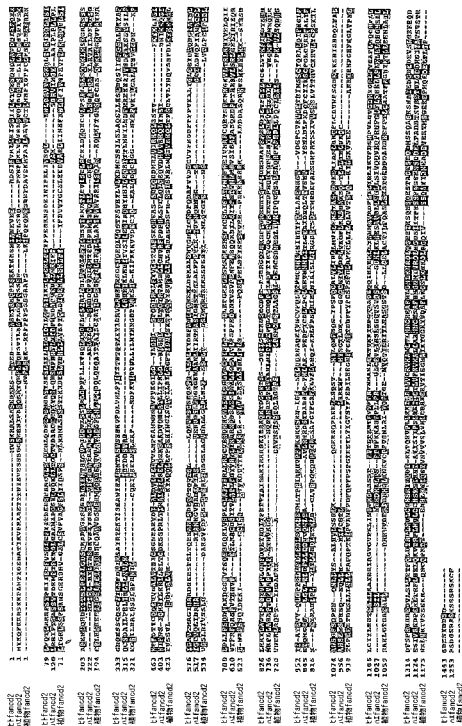


FIG. 13A



FIG. 13B

【 図 1 5 】



配列番号: 1-3

Figure 15



【 図 17 】

Genomic map showing sequence coordinates and annotations for Figure 17. The map includes coordinates from 1114 to 4114 and various gene models such as FANCD1, FANCD2, and FANCD3. The sequence is presented in a standard 60-base line format.

配列番号 : 187-188

Figure 17

【 図 18 - 1 】

配列番号 : 6

HTFANCD-2長形態

Sequence for Figure 18-1 (HTFANCD-2 long form). The sequence is shown in a standard 60-base line format, starting with TAGAATCGAA... and ending with CTCACAGGAATCAGGCCAAA...

Figure 18

【 図 18 - 2 】

Sequence for Figure 18-2. The sequence is shown in a standard 60-base line format, starting with GGTTACTGAGACTTTGTTGGAGAGACGATAACGGAACCTGGAGGAGATTGA...

【 図 18 - 3 】

Sequence for Figure 18-3. The sequence is shown in a standard 60-base line format, starting with AACATACAAAATAGAAATGTGAAGCATTGGCTGAAAGCCAAAGCACAGATC...

Figure 19

【 図 19 - 1 】

配列番号 : 7

HTFANCD2短形態

Sequence for Figure 19-1 (HTFANCD2 short form). The sequence is shown in a standard 60-base line format, starting with TCGAAACACTCGGGCCGGCAGCGCTTCTCGGAAGTAATTTAAGTGCACAA...

【 図 19 - 2 】

CTGACACTTGTGTGGAGACACAGCATAACCGAACTTGGAGGAGATTGATGGTC
TACTAGATTGCTATATTCTCTAACTGACCTGGACCTGGAGAGAAGTTGGAGTCC
ATGCTGCTAAAGAGCGGTCAITCACTGTTCTCTCATATTTCTTCTCAACTTGG
TTCCGAGAGATTGTAATAGCTTCTTGGAGGAAACAATCACTGGAGATGAAGGGGA
AGTGCTCAGTCTGGTTAAAGGCACATTTAGAAATTCGAAATAATCTGGAAAGAAT
CTTGGCAGTCCACCCAGACTGTGCTCTTGGAAACTTTGATGTGAAACTTT
TAGATATAACACTCATACTGTTACTGCTATTTCAGCAAAAATCAGAAAGAAAG
AAAAATAGAAAGGAAACAAAACAGATGGCAGCAAGACATCTCTCTCAACA
CACTTTCAGAAAGAGAAAATTTCAGAAATGTGACCTTACGCCATCTCATAGGCCA
GCTAAACAAGGAGTTCACAGGAAGGAAGAAAGACATCATGTTACTACATAA
TTCCATGCTTTTTTCCGAGAGCTGGACATTTAGAGGCTTCTCTATTTCTACATAA
ACTTGTGACGAAGTTCATCTAGATACCTGAAATGGCACTGAAGCTACAGAACTT
GTGAACTTGGCCCTTGTAGCTTTTCTTCTGGAGATCTCTCCAGAACTT
GGAGAGTATGGTACACTCTCTATTGCCAGGAGTCCCTTTCTCAAGAACAAA
GGAAGCCGGAATATTGGATTCTCACATCTCCAAACAGAGATCTGCCCAAGAAATTG
TTCAATTTGTTTTCAACTGCTACCCCAATCTGTAAACCCTGGAGAACATTCAC
AACTATTTTCAGTGTTTAGCTGCTGAGCAATTCAGGCTGTAGTTGATGACCCAGGAT
GAAAGTTCAGGAGTACACATTAATGTTCTCTCAACTGCAAAAATCAGAAATTAAGT
TTCAATGGCTTTTTGCTTGGAGTGGATTTCTCAAGCTTAAAGTGTATTTTGGAG
AAATCAACAGCTTCTGCTCAGCAAAAGAAAATTTGCTTCCCTGCCAGCAAT
TCCCTGTGCGGTCTGCTCTGTAATCTGAGCAAGACAGCAATCTCTAATGA
CCAGCTCTATGCTCTGCTGTAATCTGAGCAAGACAGCAATCTCTGAAAG
GCAATAGAGGAGATTGCTGGTGTGCTGCTGAGCACTTGAATGCTCTGAAAG
ATGATCTTCTCCACATCTCCCTACACTGACAGGCAATCTTTGTTGTTCTTCC
GTGCTATGATGGCTGAACTAGAGAAAGACGGTGAATAAGCAAGCAATCTCTAATGA
CAGCACTCTGAGGAGTCTGCTGTAATCTGAGCAAGACAGCAATCTCTGAAAG
TTAGATATGATGCTGAGGAGCACTTCTGCTGTGAGCAAAAACAGAAACAAAG
GGAAAGACCGGGGAAAGAAACAAAAGCTGATAGCAAAAGCACTCTGTTTC
GGACACACTTTAAACAGAGCACTTCCAGAGTGTGACATGGCCCACTCTGGGAG
AAGCCACTGATGACAAGGATCCACAGGGAAGGAAGGAAGACGTTTGTGCTACT
GCAGAAATACCCGCTTTTTTCCGAGAGCTGCACTTGAAGGCTTCTCTAATCTAC
ATTCTGGACTTGTGACCAAGTTACTTGTAGACACTGAAATGCACACTGAAAGTAC
AGAGGCTGTACAGCTTGGGGCTGTGAGCTGCTCTTCTGCTGGAAGATCTTTCCC
AGAAAGTAGAAATATGCTGACTGCTCTTTTCCAAAGAGAAATCTGCTGCTTAAAG
AATAAAGGAAGGCAGAAATATGGCTTCTCACATCTTCTACAGAGATCTGTCCAG
ACATTTGTGACAGTGTGGTTGAGTCTAAACCCGATGTGTAACCACTTGGAGAAC
ATTCACAACTTCTTCACTGCTTGTGCTGAGCTGAGCACTCACTGACAGATGCAAGGC
GAGAGGCACAGCTCAGGAGCAGCACCAATGGCTGCTGCTACCAAGAAAGCTGCT
GCAGGCTTGTGACCGCTCTTTGGCTGGAAGGATTTACTCAACCAATCAAAAGCAC
CGCTCTGCACTCAGCCCTTGGAGTCTCTCGAACCGCAATAAGCAGATGGAAC
AGGACCCAGCCCTTGGAGAACTGGTACGCCAGAGCTTCACTTACTGACAGAACTT
CCAACATAGTGTCCAGTTTCCAGTGTGGTCTTCACTTCTCAGACTTCTGATGGC
CCTTCTGGAGAAAGTCTGAGTACTCAACAGAGAAAGAAAACCTTCCCTCTGCT
GCCAAACAGCTGCTTTGCCGAGCATGGCTCATGGGGAAAAGAGAAAGAACCC
ACTTTAATGACCACTGATGATGCTTACTACTACTTGGAGCACAAGAGCA
GTTCTGAAAGGCATAGAGGAGATYACTGGTGTGGTTCGCCAGAACTGGTCACT
GCTCCGAAAGAGCGCGCTCTCTACATCTCCCTACGTTGACCGRGCACACTTTTGT
CATATTTCTGCTGTGATGATGCTGAACTCGAAGAGCGGTGAAGGGTCTYACG
CTGGCACAGCAGAGATTCCAGCAGGTTACAGAAAGAGAAAGCTCTCTATTGGA
ACATGGCTGCTCCGAGATTCAGYATCTTCTCAATCTGATGAAAGTATTGACAGT
TACTGTTCTGCAATGTTGTTAAAGTATGGCCGTGCTTTTGGAGGCAATTTCTG
AAGCAATGATGCCACTCTGACTTCACTTTAGAAAGCATGGGGAAGATGTTCT
TGAGCTTGTGCAAAACCTTCACTGAAACAGGAGCTACTCATCACTTTTGTGGA
CACTCCAAAGATTCGCCAGGACACAGACTCAACAGCAAGTGTCCCTTACTCAAAA
AGTCACTGGAACCTGTTAGTTGAGAGTCAAAAGCCATGCTTGTCTCAAACTGT
AGAGAGGCTTCTGGTGGTACTTCAAAAACCGAGACTTACAGGGTGAAGAAA
TTATTTCAGGATCCCTCTTCTCAGAGAGCAATCCAGAGGACAGTGAAGGATGG
CGTGACATCTACGCTCTCAGGAAACAGAGCAACAGAGGATGGGGAAGTGAAGC
AAGTGTGAACAGAAAGGACAGGACAGTGTGAAAGTGAAGCAAGCTCCAGTTA
GAGCCGAGTGGATGGTGGCTGCTCACTGTACAGACTCTCATCTTTTGGGG
TTGAAAGTCAAGTGTCTTTTTTCTAGTCAAGAGCACTCTGTTGTCTCAAGAA
GGGGTTTATAATTTCCAGGGTTCACAGGTTGCTCAACCTCCAGGTCCCT
GCTTCAGGAGTCCAGTGTAGCAATCCATCTGACTAGGAYAAACATGGCTGGGCT
CCAGTGCAGTGCAGTGCAGGCTGCTGGCCCTTGGGTTTGAACATAAAA

【 図 20 - 1 】

配列番号：8

マウス FANCD2

GGAAAGTCGAAAACGAAGGGAAGCAACTGGCGGGTCCCAAGAAATATAA
GTGGCAGAAAGAGCTTAGTCAAATGATTTCCAAAAGAGCTGGCTAGATTCTGAG
GATAAAGAAAACCTGACAGAGAGATGCTTCCAAAACATGCCCTTCCAAAGCTG
GCAAAGAACTCTCAAAATTTCTCAAGAAGTGAAGAAAATGGCAGTCTTTGTAA
AGCTTCTTAAGGCTTCAGACTCCTCITAAAACCTGGAGAGAAACAAATCACT
AGGTGTGGATCAGGTAATCTTCCAAAGGAAGCTCTTTCAGGCTTGGAGGAAGCAT
CTGTCTTATCCAAAGTAAAGAGAGCTTTGTAATGGCTGGAGTCTTACACTGA
GGACATGAGAGTCTCAGGAACTCCCTGCTGCTTCTGAGCCCTGACAGGATGAG
GAGCCAGCAGTGGCACAATTTACTCCAAAGAGTCTGATCAAGACTCTTGGGGA
TTGACATTTTACAGCTGCCATTAATAAAATGTTATTGAAAAGTGTCTCAGTTTC
TTTTGAAAAGTGAGAACAGAGATGGAATCAACATGGCCAGACTCATTAATCAATCA
ACTAAAATGGCTGGATGAAATTTGGATGGCAAGGACCTCAGCCGACAGATGAT
GCAGTTGATCAGTGTGCTCCCGTGAACCTTACAGCAATGACTTCAACAGGACCTTC
CTGAAATCTAGGGGATTCAGCAGTCTGTAATTTGGGGAAGAGCTTGGCGGAGCT
GCTGGTGCAGAAATCTCCCTGACGTGTCCAATTTTGGATGCTTTTCCAGTCTCGG
ACTTGACCCCACTCTGTGCAAGATCCCGAGTGTGGTGTGGGCAAGCTGTCACT
TCTCCCTGCTAGAGGATTTCCCTGIGATTTGAAAAGTCCCTTCTTCAATTTGTAAAC
ACACCACTTCCCTGAGGCTTTCGCCAGCTTCCGGGAGAAGTGAAGCTCCAGCA
GTTTATTGGCGTACAGAAATTCAGGCTTCCAAAAGCAAATGAAAAGTAAAAGGA
CTAGCAAGCTCTTACAGAAAATCAAGAGAACAGCTGATAAAGACTGTATTTGTTCTG
TCTTTGATGTAATAAAGTACAGCAATGAGATAAGAAAACCAATTCAGAGGGCTG
GTTAAGGCAATGAAAGCAGTGTAGTCCCGGCTGAAACATAAGGCTTTGGAGCTG
TGATGCTGCTCATCATCTACAGCACCAGCAGCAGCAAGAAAGAGGGCTGAGCA
AAGCTGCTGAGAAAACAAGATTGACTGAGACTGATTTCAAGAAAGCTGCTTGGCA
GTGCTTCTACACATTAAGGCTGTTAAGGATAAGCCCAAGCCCAACAGGACCTTC
GGCTCAGACTTGTCTCACTTCAAGACAGAGGATCATTTTGTGGAGTCTTCTT
GTACAAAATAGCTTTAAGTTTTTGTATCACTAGCCAGCAGGAAAGTGTGGTGG
CCCTAGTCCCACTGTGCAAGTGGGACTGAGGCTGAAGTCCGACACTGCACTGGA
AGTCAAGTGTGATTTGCTAAGCCCTTCTGCTCAAAAAGCCCTCAATGCTGCTTT
TGTAAGGGCACTTAGATATTGGAAAATATGTCCTCCAGCAAAATGAAAGAAA
ATTCAGGACCACTTTCAGCACTTTCGATTTACCAATGATGAGAGGCTTGA
TTCCAGGACCACTGCACTGGTGTATCCGGAAGCAGCTCTCACTGACTGTGTC
AAGTCAAGCTCAATGGGATCATTTGGTCACTTCCAGGAGGAAAGTGTGGCGG
AAGCAGAAAGTGTACCACTCACTATCCAGAGGAGGCCCAATGTGAGCAGTGT
AGCAGCGCACACAGGAGTCTTGTGTACAACACTGATTTCACTTGCAGTGCAGC
TCTCTTGGGCTTCTCTGATATGATGAAATTTGCCAAGCTGATCCAAAGAAAG
GAAGTTGGCTCAAAAACCTTGGAGTGGTGGTGTGCAAGCACTTCAATGATTTTC
CAAGATGCTTTTGGTGTGACTTCTGCTGCTCCAGAGGGTGTACTTCCATTTCT
TGAAAGCGCTTATGAGACTGGAAGAGTACAGCACTCAAGACGGCATGTGCTACA
ACCTCTGCGCTGTCTATCAGGAATGTGCAAAAGATGCCAGTCCAGGCGCACTC
ACAAGAAATCGAGCCAGAGATCAATGTCTTTTGTGCTCCCTCCATTTCCGGCT

Figure 20

【 図 20 - 2 】

TGCTGAGACTTGGCTGGCAAGACAACATGATGGAACCTGGATGAGATCCGATGG
TCTCTTAGATTTGCCCTGTTCCTCCCTGACTTGGAACTGGAGGAGAAACTGGAGT
CCATGCTGCTAAAGACCGTTCCGCTATGTTGCTGCTCACATTCCTTAACITCAACT
GGTTCGGAGAGGTTGTAATGCTTCTGCCAACAACAICTCTCTGAGATGAAGGG
CAAGGTTCTTAGTCCGCTAAAGCACTTGTAGAACTTCAAGGAAATCTAGAGAAAG
TACTTGGCAGTATCCAGACTATGTTCCGCTTTCGCAAGCGTTGACTTGGAGCAC
TTAGATATGATGCTGAGGAGCAGTCTGCTGTTGAGCAAAAACAGAAACAAAG
GGAAAGACCGGGGAAAGAAACAAAAGCTGATAGCAAAAGCACTCTGTTTC
GGACACACTTTAAACAGAGCACTTCCAGAGTGTGACATGGCCCACTCTGGGAG
AAGCCACTGATGACAAGGATCCACAGGGAAGGAAGGAAGACGTTTGTGCTACT
GCAGAAATACCCGCTTTTTTCCGAGAGCTGCACTTGAAGGCTTCTCTAATCTAC
ATTCTGGACTTGTGACCAAGTTACTTGTAGACACTGAAATGCACACTGAAAGTAC
AGAGGCTGTACAGCTTGGGGCTGTGAGCTGCTCTTCTGCTGGAAGATCTTTCCC
AGAAAGTAGAAATATGCTGACTGCTCTTTTCCAAAGAGAAATCTGCTGCTTAAAG
AATAAAGGAAGGCAGAAATATGGCTTCTCACATCTTCTACAGAGATCTGTCCAG
ACATTTGTGACAGTGTGGTTGAGTCTAAACCCGATGTGTAACCACTTGGAGAAC
ATTCACAACTTCTTCACTGCTTGTGCTGAGCTGAGCACTCACTGACAGATGCAAGGC
GAGAGGCACAGCTCAGGAGCAGCACCAATGGCTGCTGCTACCAAGAAAGCTGCT
GCAGGCTTGTGACCGCTCTTTGGCTGGAAGGATTTACTCAACCAATCAAAAGCAC
CGCTCTGCACTCAGCCCTTGGAGTCTCTCGAACCGCAATAAGCAGATGGAAC
AGGACCCAGCCCTTGGAGAACTGGTACGCCAGAGCTTCACTTACTGACAGAACTT
CCAACATAGTGTCCAGTTTCCAGTGTGGTCTTCACTTCTCAGACTTCTGATGGC
CCTTCTGGAGAAAGTCTGAGTACTCAACAGAGAAAGAAAACCTTCCCTCTGCT
GCCAAACAGCTGCTTTGCCGAGCATGGCTCATGGGGAAAAGAGAAAGAACCC
ACTTTAATGACCACTGATGATGCTTACTACTACTTGGAGCACAAGAGCA
GTTCTGAAAGGCATAGAGGAGATYACTGGTGTGGTTCGCCAGAACTGGTCACT
GCTCCGAAAGAGCGCGCTCTCTACATCTCCCTACGTTGACCGRGCACACTTTTGT
CATATTTCTGCTGTGATGATGCTGAACTCGAAGAGCGGTGAAGGGTCTYACG
CTGGCACAGCAGAGATTCCAGCAGGTTACAGAAAGAGAAAGCTCTCTATTGGA
ACATGGCTGCTCCGAGATTCAGYATCTTCTCAATCTGATGAAAGTATTGACAGT
TACTGTTCTGCAATGTTGTTAAAGTATGGCCGTGCTTTTGGAGGCAATTTCTG
AAGCAATGATGCCACTCTGACTTCACTTTAGAAAGCATGGGGAAGATGTTCT
TGAGCTTGTGCAAAACCTTCACTGAAACAGGAGCTACTCATCACTTTTGTGGA
CACTCCAAAGATTCGCCAGGACACAGACTCAACAGCAAGTGTCCCTTACTCAAAA
AGTCACTGGAACCTGTTAGTTGAGAGTCAAAAGCCATGCTTGTCTCAAACTGT
AGAGAGGCTTCTGGTGGTACTTCAAAAACCGAGACTTACAGGGTGAAGAAA
TTATTTCAGGATCCCTCTTCTCAGAGAGCAATCCAGAGGACAGTGAAGGATGG
CGTGACATCTACGCTCTCAGGAAACAGAGCAACAGAGGATGGGGAAGTGAAGC
AAGTGTGAACAGAAAGGACAGGACAGTGTGAAAGTGAAGCAAGCTCCAGTTA
GAGCCGAGTGGATGGTGGCTGCTCACTGTACAGACTCTCATCTTTTGGGG
TTGAAAGTCAAGTGTCTTTTTTCTAGTCAAGAGCACTCTGTTGTCTCAAGAA
GGGGTTTATAATTTCCAGGGTTCACAGGTTGCTCAACCTCCAGGTCCCT
GCTTCAGGAGTCCAGTGTAGCAATCCATCTGACTAGGAYAAACATGGCTGGGCT
CCAGTGCAGTGCAGTGCAGGCTGCTGGCCCTTGGGTTTGAACATAAAA

【 図 20 - 3 】

TTTAGGCTTATAATAGCTTTGTAATAAATCTGTTTCAGAGTTTGGCTCAGTACC
TTTTCTCCTTATAGATGATTTATCAAGGATCTCATATTCAAGGATTAAGTAA
TATGAGTTGAGGTTTGGCAATCGTACTGGTGGCTAAAAGTATGTTCCGACTGT
TATCTCTCGGAGGAATGACCCAACTTCTTATCAATGATCAAGTGTGGTGTGG
TCTGCTCAGGCTCTTTTACATAGTCTGGTGGTGTATTAGATATGTTGAGCC
AGGAGGGTGTGAACATTACTTTGAAATTTAAAACATTTTGTACATATGTTGATGG
GCTATATGTTGCCACTGTGCAATGTTGATGTTAGTCAAGGATGATGTTGAGGAGT
GCTCTCTCTCCACCTATGGTGGTCCAGGGTCAAACTTACAGCTTCACTGCTCT
AGCCACTTACCTTTTAAAATGTTGGTTATTAATATAAAGAAAGGAAGGAAGAC
AACTCAAAACATGTGCTGGCTTTGATGATATAATAGTTTATTTTCCACATTAAT
TGAAATGAGCAATAAATAATTTGTAATAATCACTACAAAATAAATTTGAAATTA
GAAATAGAAACATCAGGAGTAAATAGGGGATCTTCTGTCTTTCAGGAAAGC
CCAGTCTCAGAGATGCTGCCAGGCTTCTCTGCTGTTGCCATTAAGATTAATTAAT
TTTTGTAATATTTTACTCATACCCGTTAAAGGATGATGTTTGTGGAAAATAAAA
AAAAAAAAAAAAAAAAAAAAA

## 【配列表】

2010246550000001.app

## 【手続補正書】

【提出日】平成22年6月22日(2010.6.22)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

(a) 配列番号：7に示すヌクレオチド配列；(b) (a)のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列；(c) (b)のヌクレオチド配列に相補的なヌクレオチド配列；

から選択されるポリヌクレオチドを含んでなる単離された核酸分子。

【請求項2】

(a) 配列番号：6に示すヌクレオチド配列；(b) (a)のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列；(c) (b)のヌクレオチド配列に相補的なヌクレオチド配列；

から選択されるポリヌクレオチドを含んでなる単離された核酸分子。

【請求項3】

(a) 配列番号：5に示すヌクレオチド配列；(b) (a)のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列；(c) (b)のヌクレオチド配列に相補的なヌクレオチド配列；

から選択されるポリヌクレオチドを含んでなる単離された核酸分子。

【請求項4】

(a) 配列番号：8に示すヌクレオチド配列；(b) (a)のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列；(c) (b)のヌクレオチド配列に相補的なヌクレオチド配列；

から選択されるポリヌクレオチドを含んでなる単離された核酸分子。

【請求項5】

(a) 配列番号：4に示すアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列；(b) (a)のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列；(c) (b)のヌクレオチド配列に相補的なヌクレオチド配列；

から選択されるポリヌクレオチドを含んでなる単離された核酸分子。

【請求項6】

DNA修復を促進するための細胞の核においてFANCD2タンパク質の短い形態から長い形態への変換の生物学的特性を保持するのに十分に、配列番号：4のアミノ酸配列と類似したアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列から本質的に構成される単離された核酸分子。

【請求項7】

ポリヌクレオチドがDNA分子、cDNA分子、又はRNA分子である請求項1～6のいずれか1項に記載の単離された核酸分子。

【請求項8】

配列番号：9から191に少なくとも90%同一であるか、または配列番号：9から191に少なくとも90%同一であるヌクレオチド配列に相補的である、ヌクレオチド配列を有するポリヌクレオチドから本質的に構成される単離された核酸分子。

【請求項9】

配列が、配列番号：9から94より選択されるイントロン/エクソン配列である請求項

8に記載の単離された核酸分子。

【請求項10】

配列が、配列番号：115から186より選択されるPCRプライマーの配列である、請求項8に記載の単離された核酸分子。

【請求項11】

(a) 配列番号：7のヌクレオチド配列を有する核酸によりコードされるアミノ酸配列；  
(b) 配列番号：7のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列を有する核酸によりコードされるアミノ酸配列；及び  
(c) 少なくとも50アミノ酸の長さを有する、(a)又は(b)の、ポリペプチドフラグメント；  
から選択されるアミノ酸配列を含んでなる単離されたポリペプチド。

【請求項12】

(a) 配列番号：6のヌクレオチド配列を有する核酸によりコードされるアミノ酸配列；  
(b) 配列番号：6のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列を有する核酸によりコードされるアミノ酸配列；及び  
(c) 少なくとも50アミノ酸の長さを有する、(a)又は(b)の、ポリペプチドフラグメント；  
から選択されるアミノ酸配列を含んでなる単離されたポリペプチド。

【請求項13】

(a) 配列番号：5のヌクレオチド配列を有する核酸によりコードされるアミノ酸配列；  
(b) 配列番号：5のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列を有する核酸によりコードされるアミノ酸配列；及び  
(c) 少なくとも50アミノ酸の長さを有する、(a)又は(b)の、ポリペプチドフラグメント；  
から選択されるアミノ酸配列を含んでなる単離されたポリペプチド。

【請求項14】

(a) 配列番号：8のヌクレオチド配列を有する核酸によりコードされるアミノ酸配列；  
(b) 配列番号：8のヌクレオチド配列に少なくとも90%同一であるヌクレオチド配列を有する核酸によりコードされるアミノ酸配列  
(c) フラグメントが少なくとも50アミノ酸の長さである(a)又は(b)の、ポリペプチドフラグメント；  
から選択されるアミノ酸配列を含んでなる単離されたポリペプチド。

【請求項15】

(a) 配列番号：4；  
(b) (a)に少なくとも90%同一であるアミノ酸配列；  
(c) 配列番号：187から188の少なくとも1つに少なくとも90%同一であるヌクレオチド配列を有するポリヌクレオチドによりコードされるアミノ酸配列；及び  
(d) 少なくとも50アミノ酸の長さを有する、(a)から(c)の、ポリペプチドフラグメント；  
から選択されるアミノ酸配列を含んでなる単離されたポリペプチド。

【請求項16】

nt376のAからG、nt3707のGからA、nt904のCからT、及びnt958のCからTより選択される変異を有するDNAによりコードされる請求項11~15のいずれか1項に記載の単離されたポリペプチド。

【請求項17】

多型がnt1122のAからG、nt1440のTからC、nt1509のCからT、nt2141のCからT、nt2259のTからC、nt4098のTからG、nt4453のGからAより選択される、ポリペプチドをコードするDNAの多型によりポリペプチドが特徴付けられる請求項11~15のいずれか1項に記載の単離されたポリペプチド。

## 【請求項 18】

ポリペプチドがアミノ酸 222 またはアミノ酸 561 での変異により特徴付けられる請求項 11 ~ 15 のいずれか 1 項に記載の単離されたポリペプチド。

## 【請求項 19】

生物学的サンプルが、ファンコーニ貧血経路を有するか否かを決定する方法であって、  
 (a) FANCD2 タンパク質のアイソフォームである FANCD2-L と第 1 複合体を形成するために、生物学的サンプルを第 1 抗体に暴露すること  
 (b) FANCD2 タンパク質のアイソフォームである FANCD2-S と第 2 複合体を形成するために、生物学的サンプルを第 2 抗体に暴露すること；  
 (c) マーカーでサンプル中の第 1 複合体及び第 2 複合体の量を検出すること  
 (d) 第 1 複合体の量と第 2 複合体の量との比を測定すること；及び  
 (e) (d) で得られた比を野生型細胞における FANCD2-L 複合体の量と FANCD2-S 複合体の量との比と比較すること；  
 を含む方法。

## 【請求項 20】

生物学的サンプルが、無傷細胞又はライゼート中に溶解された細胞を含んでなる請求項 19 に記載の方法。

## 【請求項 21】

生物学的サンプルが癌に対する感受性を有するかまたは初期段階の癌を有するヒト被検対象からのものである請求項 19 に記載の方法。

## 【請求項 22】

生物学的サンプルがヒト対象の癌からのものであり、ここで癌が黒色腫、白血病、アストチトーム、神経膠芽細胞腫、リンパ腫、神経膠腫、ホジキンリンパ腫、慢性リンパ球性白血病及び、膵臓、乳房、卵巣、子宮、精巣、脳下垂体、腎臓、胃、食道及び直腸の癌から選択される請求項 21 に記載の方法。

## 【請求項 23】

生物学的サンプルがヒト胎児又は成人からのものである請求項 19 に記載の方法。

## 【請求項 24】

生物学的サンプルが血液サンプル、被験対象及び細胞系からの組織の生検サンプルから選択される請求項 19 に記載の方法。

## 【請求項 25】

生物学的サンプルが心臓、脳、胎盤、肝臓、骨格筋、腎臓、膵臓、脾臓、胸腺、前立腺、精巣、子宮、小腸、結腸、末梢血またはリンパ球に由来する請求項 19 に記載の方法。

## 【請求項 26】

マーカーが蛍光マーカー又は化学ルミネセンスマーカーであり、該蛍光マーカー又は化学ルミネセンスマーカーは第 1 抗体及び第 2 抗体と抱合されていてよい、請求項 19 に記載の方法。

## 【請求項 27】

第 1 及び第 2 複合体を、基質に抱合された第 3 抗体に結合することをさらに含む請求項 19 に記載の方法。

## 【請求項 28】

被験対象からの細胞集団におけるファンコーニ貧血経路における欠損を同定する方法であって、

(a) 細胞集団における FANCD2 タンパク質のアイソフォームである FANCD2-L の量を決定すること；  
 (b) 細胞集団における FANCA、FANCB、FANCC、FANCD1、FANCE、FANCF または FANCG タンパク質のいずれかの量を決定すること；  
 (c) (a) 及び (b) で決定された量を、野生型細胞集団における、FANCD2-L の量、及び、FANCA、FANCB、FANCC、FANCD1、FANCE、FANCF または FANCG タンパク質のいずれかの量と比較すること；

を含む方法。

【請求項 29】

アイソフォームの量の決定が、FANCD2タンパク質のアイソフォームである、FANCD2-LとFANCD2-Sとの分離に依存する請求項28に記載の方法。

【請求項 30】

分離がゲル電気泳動、又は遊走結合バンド試験ストリップにより達成される請求項29に記載の方法。

【請求項 31】

(a) 被験対象から細胞サンプルを入手すること；

(b) 細胞サンプルをDNA損傷手段に暴露すること；及び

(c) FANCD2タンパク質のアイソフォームであるFANCD2-Lを決定すること、

を含む、FA経路欠損を検出するための細胞に基くアッセイであって、ここで、(c)で決定された量が、初期の量より増えていないことがFA経路欠損を示している、アッセイ

。

【請求項 32】

核フォーカスを検出するためのイムノプロットティング；FANCD2アイソフォームの量を検出するためのウェスタンプロット、及びDNAプローブでのハイブリダイゼーションによりmRNAを定量すること、からなる群より選ばれる方法により、FANCD2-Lアイソフォームの量が測定される請求項31に記載のアッセイ。

【請求項 33】

(a) FANCD2タンパク質のアイソフォームであるFANCD2-Lが低減された量で作製される細胞集団を選択すること；

(b) 細胞集団を候補治療薬分子のライブラリーの各メンバーに暴露すること；及び

(c) 細胞集団においてFANCD2-Lの量を増加させる、個々のメンバー分子を同定すること；

からなる治療薬を同定するためのスクリーニングアッセイ。

【請求項 34】

細胞集団がインビトロ細胞集団、変異FANCD2遺伝子を有している実験動物のインビボ細胞集団、又はFAND2遺伝子がヒト変異FANCD2遺伝子により置換されているノックアウトマウスのインビボ細胞集団である、請求項33記載のスクリーニングアッセイ。

【請求項 35】

(b)の前に、FANCD2-Lが低減された量で作製される細胞集団に化学発癌物質が添加される、請求項33に記載のスクリーニングアッセイ。

【請求項 36】

(a) 高ストリンジェント条件下でFANCD2遺伝子の配列に結合する配列番号：115から186より選択されるプライマー対；及び

(b) 各々のプライマー対の容器；

からなる生物学的サンプル中の癌細胞を検出するのに使用するためのキット。

## フロントページの続き

(51)Int.Cl.		F I	テーマコード(参考)
A 6 1 P 35/02	(2006.01)	A 6 1 P 35/02	
A 6 1 P 25/00	(2006.01)	A 6 1 P 25/00	
A 6 1 P 1/18	(2006.01)	A 6 1 P 1/18	
A 6 1 P 15/00	(2006.01)	A 6 1 P 15/00	
A 6 1 P 13/12	(2006.01)	A 6 1 P 13/12	
A 6 1 P 1/04	(2006.01)	A 6 1 P 1/04	
A 6 1 K 48/00	(2006.01)	A 6 1 K 48/00	
G 0 1 N 33/53	(2006.01)	G 0 1 N 33/53	D
G 0 1 N 33/533	(2006.01)	G 0 1 N 33/533	
G 0 1 N 33/50	(2006.01)	G 0 1 N 33/53	M
G 0 1 N 33/15	(2006.01)	G 0 1 N 33/50	Z
G 0 1 N 27/447	(2006.01)	G 0 1 N 33/15	Z
		G 0 1 N 27/26	3 0 1 A

(74)代理人 100095500

弁理士 伊藤 正和

(74)代理人 100111235

弁理士 原 裕子

(72)発明者 ダンドレア、 アラン ディー .

アメリカ合衆国 マサチューセッツ州 0 1 8 9 0 ウィンチェスター カルメット ストリート  
2 4

(72)発明者 タニグチ、 トシヤス

アメリカ合衆国 マサチューセッツ州 0 2 2 1 5 ボストン ロングウッド アヴェニュー 3  
6 8

(72)発明者 ティマース、 シンシア

アメリカ合衆国 オハイオ州 4 3 2 2 8 コロンブス ハビット ドライブ 3 8 3 5

(72)発明者 グロンプ、 マーカス

アメリカ合衆国 オレゴン州 9 7 2 0 1 ポートランド エスダブリュー サーティフォース  
アヴェニュー 6 5 4 5

F ターム(参考) 2G045 AA26

4B024 AA01 AA12 CA01 CA04 CA11 HA08 HA12

4B063 QA17 QA18 QA19 QQ02 QQ08 QQ42 QQ52 QR08 QR32 QR35

QR42 QR50 QR55 QR62 QR72 QR77 QS03 QS25 QS28 QS34

QS36 QX02

4C084 AA13 NA14 ZA021 ZA661 ZA811 ZA891 ZB261 ZB271

【外国語明細書】

## TITLE OF THE INVENTION

METHODS AND COMPOSITIONS FOR THE DIAGNOSIS OF CANCER  
SUSCEPTIBILITIES AND DEFECTIVE DNA REPAIR MECHANISMS AND  
TREATMENT THEREOF

The work described herein was supported by the National Institute of Health,  
NIH Grant vNo. Health grants RO1HL52725-04, RO1 DK43889-09, 1PO1HL48546,  
and PO1HL54785-04. The US Government has certain rights to the claimed invention

### Technical Field and Background Art

The present invention relates to the diagnosis of cancer susceptibilities in subjects  
having a defect in the FANCD2 gene and the determination of suitable treatment  
protocols for those subjects who have developed cancer. Animal models with defects in  
the FANCD2 gene can be used to screen for therapeutic agents.

Fanconi Anemia (FA) is an autosomal recessive cancer susceptibility syndrome  
characterized by birth defects, bone marrow failure and cancer predisposition. Cells  
from FA patients display a characteristic hypersensitivity to agents that produce  
interstrand DNA crosslinks such as mitomycin C or diepoxybutane. FA patients develop  
several types of cancers including acute myeloid leukemias and cancers of the skin,  
gastrointestinal; and gynecological systems. The skin and gastrointestinal tumors are  
usually squamous cell carcinomas. At least 20% of patients with FA develop cancers.  
The average age of patients who develop cancer is 15 years for leukemia, 16 years for  
liver tumors and 23 years for other tumors. (D'Andrea et al., Blood, (1997) Vol. 90, pp.  
1725, Garcia-Higuera et al. Curr. Opin. Hematol, (1999) Vol 2, pp. 83-88 and Heijna et  
al. Am. J. Hum. Genet. Vol 66, pp 1540-1551)

FA is genetically heterogeneous. Somatic cell fusion studies have identified at  
least seven distinct complementation groups (Joenje et al., (1997) Am. J. Hum. Genet.,  
Vol. 61, pp. 940-944 and Joenje et al., (2000) Am. J. Hum. Genet, Vol. 67, pp. 759-762).  
This observation has resulted in the hypothesis that the FA genes define a  
multicomponent pathway involved in cellular responses to DNA cross-links. Five of the  
FA genes (FANCA, FANCC, FANCE, FANCF and FANCG) have been cloned and the

FANCA, FANCC and FANCG proteins have been shown to form a molecular complex with primarily nuclear localization. FANCC also localizes in the cytoplasm. Different FA proteins have few or no known sequence motifs with no strong homologs of the FANCA, FANCC, FANCE, FANCF, and FANCG proteins in non-vertebrate species.

5 FANCF has weak homology of unknown significance to an E. Coli RNA binding protein. The two most frequent complementation groups are FA-A and FA-C which together account for 75%-80% of FA patients. Multiple mutations have been recognized in the FANCA gene that span 80kb and consists of at least 43 exons. FANCC has been found to have 14 exons and spans approximately 80kb. A number of mutations in the  
10 FANCC gene have been identified which are correlated with FA of differing degrees of severity. FA-D has been identified as a distinct but rare complementation group. Although FA-D patients are phenotypically distinguishable from patients from other subtypes, the FA protein complex assembles normally in FA-D cells (Yamashita et al., (1998) P.N.A.S., Vol. 95, pp.13085 – 13090)

15 The cloned FA proteins encode orphan proteins with no sequence similarity to each other or to other proteins in GenBank and no functional domains are apparent in the protein sequence. Little is known regarding the cellular or biochemical function of these proteins.

Diagnosis of FA is complicated by the wide variability in FA patient phenotype.  
20 Further confounding diagnosis, approximately 33% of patients with FA have no obvious congenital abnormalities. Moreover, existing diagnostic tests do not differentiate FA carriers from the general population. The problems associated with diagnosis are described in D'Andrea et al., (1997). Many cellular phenotypes have been reported in FA cells but the most consistent is hypersensitivity to bifunctional alkylating agents such  
25 as mitomycin C or diepoxybutane. These agents produce interstrand DNA cross-links (an important class of DNA damage).

Diagnosing cancer susceptibility is complicated because of the large number of regulatory genes and biochemical pathways that have been implicated in the formation of cancers. Different cancers depending on how they arise and the genetic lesions  
30 involved may determine how a subject responds to any particular therapeutic treatments. Genetic lesions that are associated with defective repair mechanisms may give rise to defective cell division and apoptosis which in turn may increase a patient's susceptibility to cancer. FA is a disease condition in which multiple pathological outcomes are associated with defective repair mechanisms in addition to cancer susceptibility

An understanding of the molecular genetics and cell biology of Fanconi Anemia pathway can provide insights into prognosis, diagnosis and treatment of particular classes of cancers and conditions relating to defects in DNA repair mechanisms that arise in non-FA patients as well as FA patients

5

### Summary of the Invention

In a first embodiment of the invention there is provided an isolated nucleic acid molecule that includes a polynucleotide selected from (a) a nucleotide sequence encoding a polypeptide having an amino acid sequence as shown in SEQ ID NO: 4 (b) a nucleotide sequence at least 90% identical to the polynucleotide of (b); (c) a nucleotide sequence complementary to the polynucleotide of (b); (d) a nucleotide sequence at least 90% identical to the nucleotide sequence shown in SEQ ID No: 5-8, 187-188; and (e) a nucleotide sequence complementary to the nucleotide sequence of (d). The polynucleotide may be an RNA molecule or a DNA molecule, such as a cDNA.

In another embodiment of the invention, an isolated nucleic acid molecule is provided that consists essentially of a nucleotide sequence encoding a polypeptide having an amino acid sequence sufficiently similar to that of SEQ ID No: 4 to retain the biological property of conversion from a short form to a long form of FANCD2 in the nucleus of a cell for facilitating DNA repair. Alternately, the isolated nucleic acid molecule consists essentially of a polynucleotide having a nucleotide sequence at least 90% identical to SEQ ID NO: 9-191 or complementary to a nucleotide sequence that is at least 90% identical to SEQ ID NO: 9-191.

In an embodiment, a method is provided for making a recombinant vector that includes inserting any of the isolated nucleic acid molecules described above into a vector. A recombinant vector product may be made by this method and the vector may be introduced to form a recombinant host cell into a host cell.

In an embodiment of the invention, a method is provided for making an FA-D2 cell line, that includes (a) obtaining cells from a subject having a biallelic mutation in a complementation group associated with FA-D2; and (b) infecting the cells with a transforming virus to make the FA-D2 cell line where the cells may be selected from fibroblasts and lymphocytes and the transforming virus selected from Epstein Barr virus and retrovirus. The FA-D2 cell line may be characterized by determining the presence of a defective FANDC2 in the cell line for example by performing a diagnostic assay

selected from (i) a Western blot or nuclear immunofluorescence using an antibody specific for FANCD2 and (ii) a DNA hybridization assay.

In an embodiment of the invention, a recombinant method is provided for producing a polypeptide, that includes culturing a recombinant host cell wherein the host  
5 cell includes any of the isolated nucleic acid molecules described above.

In an embodiment of the invention, an isolated polypeptide, including an aminoacid sequence selected from (a) SEQ ID NO: 4; (b) an aminoacid sequence at least 90% identical to (a); (c)  
10 an aminoacid sequence which is encoded by a polynucleotide having a nucleotide sequence which is at least 90% identical to at least one of SEQ ID NO: 5-8, 187-188; (d) an aminoacid sequence which is encoded by a polynucleotide having a nucleotide sequence which is at least 90% identical to a complementary sequence to at least one of SEQ ID NO: 5-8, 187-188; and (e) a polypeptide fragment of (a) – (d) wherein the fragment is at least 50 aminoacids in length.

15 The isolated polypeptide may be encoded by a DNA having a mutation selected from nt 376 A to G, nt 3707 G to A, nt904C to T and nt 958C to T. Alternatively, the polypeptide may be characterized by a polymorphism in DNA encoding the polypeptide, the polymorphism being selected from nt 1122A to G, nt 1440T to C, nt1509C to T, nt2141C to T, nt2259T to C, nt4098T to G, nt4453G to A. Alternatively, the  
20 polypeptide may be characterized by a mutation at aminoacid 222 or aminoacid 561.

In an embodiment of the invention, an antibody preparation is described having a binding specificity for a FANCD2 protein where the antibody may be a monoclonal antibody or a polyclonal antibody and wherein the FANCD2 may be FANCD2-S or FANCD2-L.

25 In an embodiment of the invention, a diagnostic method is provided for measuring FANCD2 isoforms in a biological sample where the method includes (a) exposing the sample to a first antibody for forming a first complex with FANCD2-L and optionally a second antibody for forming a second complex with FANCD2-S; and (b) detecting with a marker, the amount of the first complex and the second complex in the  
30 sample. The sample may be intact cells or lyzed cells in a lysate. The biological sample may be from a human subject with a susceptibility to cancer or having the initial stages of cancer. The sample may be from a cancer in a human subject, wherein the cancer is selected from melanoma, leukemia, astocytoma, glioblastoma, lymphoma, glioma, Hodgkins lymphoma, chronic lymphocyte leukemia and cancer of the pancreas, breast,

thyroid, ovary, uterus, testis, pituitary, kidney, stomach, esophagus and rectum. The biological sample may be from a human fetus or from an adult human and may be derived from any of a blood sample, a biopsy sample of tissue from the subject and a cell line. The biological sample may be derived from heart, brain, placenta, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, uterus, small intestine, colon, peripheral blood or lymphocytes. The marker may be a fluorescent marker, the fluorescent marker optionally conjugated to the FANCD2-L antibody, a chemiluminescent marker optionally conjugated to the FANCD2-L antibody and may bind the first and the second complex to a third antibody conjugated to a substrate.

10 Where the sample is a lysate, it may be subjected to a separation procedure to separate FANCD2 isoforms and the separated isoforms may be identified by determining binding to the first or the second FANCD2 antibody.

In an embodiment of the invention, a diagnostic test is provided for identifying a defect in the Fanconi Anemia pathway in a cell population from a subject, that includes selecting an antibody to FANCD2 protein and determining whether the amount of an FANCD2-L isoform is reduced in the cell population compared with amounts, in a wild type cell population; such that if the amount of the FANCD2-L protein is reduced, then determining whether an amount of any of FANCA, FANCB, FANCC, FANCD1, FANCE, FANCF or FANCG protein is altered in the cell population compared with the wild type so as to identify the defect in the Fanconi Anemia pathway in the cell population. In one example, the amount of an isoform relies on a separation of the FANCD2-L and FANCD2-S isoforms where the separation may be achieved by gel electrophoresis or by a migration binding banded test strip.

15  
20

In an embodiment of the invention, a screening assay for identifying a therapeutic agent, is provided that includes selecting a cell population in which FANCD2-L is made in reduced amounts; exposing the cell population to individual members of a library of candidate therapeutic molecules; and identifying those individual member molecules that cause the amount of FANCD2-L to be increased in the cell population. In one example, the cell population is an *in vitro* cell population. In another example, the cell population is an *in vivo* cell population, the *in vivo* population being within an experimental animal, the experimental animal having a mutant FANCD2 gene. In a further example, the experimental animal is a knock-out mouse in which the mouse FANCD2 gene has been replaced by a human mutant FANCD2 gene. In another example, a chemical carcinogen is added to the cell population in which FANCD2 is made in reduced amounts, to

25  
30

determine if any member molecules can cause the amount of FANCD2-L to be increased so as to protect the cells from the harmful effects of the chemical carcinogen.

In an embodiment of the invention, an experimental animal model is provided in which the animal FANCD2 gene has been removed and optionally replaced by any of  
5 the nucleic acid molecules described above

In an embodiment of the invention, a method is provided for identifying in a cell sample from a subject, a mutant FANCD2 nucleotide sequence in a suspected mutant FANCD2 allele which comprises comparing the nucleotide sequence of the suspected mutant FANCD2 allele with the wild type FANCD2 nucleotide sequence wherein a  
10 difference between the suspected mutant and the wild type sequence identifies a mutant FANCD2 nucleotide sequence in the cell sample. In one example, the suspected mutant allele is a germline allele. In another example, identification of a mutant FANCD2 nucleotide sequence is diagnostic for a predisposition for a cancer in the subject or for an increased risk of the subject bearing an offspring with Fanconi Anemia. In another  
15 example, the suspected mutant allele is a somatic allele in a tumor type and identifying a mutant FANCD2 nucleotide sequence is diagnostic for the tumor type. In another example, the nucleotide sequence of the wild type and the suspected mutant FANCD2 nucleotide sequence is selected from a gene, a mRNA and a cDNA made from a mRNA. In another example, comparing the polynucleotide sequence of the suspected mutant  
20 FANCD2 allele with the wild type FANCD2 polynucleotide sequence, further includes selecting a FANCD2 probe which specifically hybridizes to the mutant FANCD2 nucleotide sequence, and detecting the presence of the mutant sequence by hybridization with the probe. In another example, comparing the polynucleotide sequence of the suspected mutant FANCD2 allele with the wild type FANCD2 polynucleotide sequence,  
25 further comprises amplifying all or part of the FANCD2 gene using a set of primers specific for wild type FANCD2 DNA to produce amplified FANCD2 DNA and sequencing the FANCD2 DNA so as to identify the mutant sequence. In another example, where the mutant FANCD2 nucleotide sequence is a germline alteration in the FANCD2 allele of the human subject, the alteration is selected from the alterations set  
30 forth in Table 3 and where the mutant FANCD2 nucleotide sequence is a somatic alteration in the FANCD2 allele of the human subject, the alteration is selected from the alterations set forth in Table 3.

In an embodiment of the invention, a method is provided for diagnosing a susceptibility to cancer in a subject which comprises comparing the germline sequence



comparing the sequence of the FANCD2 gene or the sequence of its mRNA in a tissue sample from a lesion of the subject with the sequence of the wild-type FANCD2 gene or the sequence of its mRNA, wherein an alteration in the sequence of the FANCD2 gene or the sequence of its mRNA of the subject indicates a mutation at the FANCD2 gene of the neoplastic lesion. A therapeutic protocol may be provided for treating the neoplastic lesion according to the mutation at the FANCD2 gene of the neoplastic lesion.

In an embodiment of the invention, a method is provided for confirming the lack of a FANCD2 mutation in a neoplastic lesion from a human subject which comprises comparing the sequence of the FANCD2 gene or the sequence of its mRNA in a tissue sample from a lesion of said subject with the sequence of the wild-type FANCD2 gene or the sequence of its RNA, wherein the presence of the wild-type sequence in the tissue sample indicates the lack of a mutation at the FANCD2 gene.

In an embodiment of the invention, a method is provided for determining a therapeutic protocol for a subject having a cancer, that includes (a) determining if a deficiency in FANCD2-L occurs in a cell sample from the subject by measuring FANCD2 isoforms using specific antibodies (b) if a deficiency is detected in (a), then determining whether the deficiency is a result of genetic defect in non-cancer cells; and (c) if (b) is positive, reducing the use of a therapeutic protocol that causes increased DNA damage so as to protect normal tissue in the subject and if (b) is negative, and the deficiency is contained within a genetic defect in cancer cells only, then increasing the use of a therapeutic protocol that causes increased DNA damage so as to adversely affect the cancer cells.

In an embodiment of the invention, a method of treating a FA pathway defect in a cell target is provided that includes: administering an effective amount of FANCD2 protein or an exogenous nucleic acid to the target. The FA pathway defect may be a defective FANCD2 gene and the exogenous nucleic acid vector may further include introducing a vector according to those described above. The vector may be selected from a mutant herpesvirus, a E1/E4 deleted recombinant adenovirus, a mutant retrovirus, the viral vector being defective in respect of a viral gene essential for production of infectious new virus particles. The vector may be contained in a lipid micelle.

In an embodiment of the invention, a method is provided for treating a patient with a defective FANCD2 gene, that includes providing a polypeptide described in SEQ ID No: 4, for functionally correcting a defect arising from a condition arising from the defective FANCD2 gene.

In an embodiment of the invention, a cell based assay for detecting a FA pathway defect is provided that includes obtaining a cell sample from a subject; exposing the cell sample to DNA damaging agents; and detecting whether FANCD2-L is upregulated, the absence of upregulation being indicative of the FA pathway defect. In the cell-based assay, amounts of FANCD2 may be measured by an analysis technique selected from:  
5 immunoblotting for detecting nuclear foci; Western blots to detect amounts of FANCD2 isoforms and quantifying mRNA by hybridising with DNA probes.

In an embodiment of the invention, a kit is provided for use in detecting a cancer cell in a biological sample, that includes (a) primer pair which binds under high  
10 stringency conditions to a sequence in the FANCD2 gene, the primer pair being selected to specifically amplify an altered nucleic acid sequence described in Table 7; and containers for each of the primers.

#### **Brief Description of the Figures**

The foregoing features of the invention will be more readily understood by  
15 reference to the following detailed description, taken with reference to the accompanying drawings, in which:

Figure 1A provides a Western blot demonstrating that the Fanconi Anemia protein complex is required for the monoubiquitination of FANCD2. Normal (WT) cells (lane 1) express two isoforms of the FANCD2 protein, a low molecular weight isoform (FANCD2-S) (155 kD) and a high molecular weight isoform (FANCD2-L) (162 kD).  
20 Lanes 3, 7, 9, 11 show that FA cell lines derived from type A, C, G, and F patients only express the FANCD2-S isoform. Lanes 4, 8, 10, 12 show the restoration of the high molecular weight isoform FANCD2-L following transfection of cell lines with corresponding FAcDNA.

Figure 1B shows a Western blot obtained after HeLa cells were transfected with a  
25 cDNA encoding HA-ubiquitin. After transfection, cells were treated with the indicated dose of mitomycin C (MMC). Cellular proteins were immunoprecipitated with a polyclonal antibody (E35) to FANCD2, as indicated. FANCD2 was immunoprecipitated, and immune complexes were blotted with anti-FANCD2 or anti-  
30 HA monoclonal antibody.

Figure 1C shows a Western blot obtained after HeLa cells were transfected with a cDNA encoding HA-ubiquitin. After transfection, cells were treated with the indicated

dose of ionizing radiation (IR). FANCD2 was immunoprecipitated, and immune complexes were blotted with anti-FANCD2 or anti-HA monoclonal antibody.

Figure 1D shows a Western blot obtained after FA-G fibroblast line (FAG326SV) or corrected cells (FAG326SV plus FANCG cDNA) were transfected with the HA-Ub cDNA, FANCD2 was immunoprecipitated, and immune complexes were  
5 blotted with anti-FANCD2 or anti-HA antisera.

Figure 1E shows a Western blot obtained after treatment of HeLa cells with 1mM hydroxyurea for 24hours. HeLa cell lysates were extracted and incubated at the indicated temperature for the indicated time period with or without 2.5 $\mu$ M ubiquitin  
10 aldehyde. The FANCD2 protein was detected by immunoblot with monoclonal anti-FANCD2 (FI17).

Figure 2 demonstrates that the Fanconi Anemia pathway is required for the formation of FANCD2 nuclear foci. Top panel shows anti-FANCD2 immunoblots of SV40 transformed fibroblasts prepared as whole cell extracts. Panels a-h show  
15 immunofluorescence with the affinity-purified anti-FANCD2 antiserum. The uncorrected (mutant, M) FA fibroblasts were FA-A (GM6914), FA-G (FAG326SV), FA-C (PD426), and FA-D (PD20F). The FA-A, FA-G, and FA-C fibroblasts were functionally complemented with the corresponding FA cDNA. The FA-D cells were complemented with neomycin-tagged human chromosome 3p (Whitney and al, 1995).

Figure 3 shows the cell cycle dependent expression of the two isoforms of the FANCD2 protein. (a) HeLa cells, SV40 transformed fibroblasts from an FA-A patient (GM6914), and GM6914 cells corrected with FANCA cDNA were synchronized by the double thymidine block method. Cells corresponding to the indicated phase of the cell cycle were lysed, and processed for FANCD2 immunoblotting (b) Synchrony by  
20 nocodazole block (c) Synchrony by mimosine block (d) HeLa cells were synchronized in the cell cycle using nocodazole or (e) mimosine, and cells corresponding to the indicated phase of the cell cycle were immunostained with the anti-FANCD2 antibody and analyzed by immunofluorescence.

Figure 4. shows the formation of activated FANCD2 nuclear foci following  
30 cellular exposure to MMC, Ionizing Radiation, or Ultraviolet Light. Exponentially-growing HeLa cells were either untreated or exposed to the indicated DNA damaging agents, (a) Mitomycin C (MMC), (b)  $\gamma$ -irradiation (IR), or (c) Ultraviolet Light (UV), and processed for FANCD2 immunoblotting or FANCD2 immunostaining. (a) Cells

were continuously exposed to 40 ng/ml MMC for 0-72 hours as indicated, or treated for 24 hours and fixed for immunofluorescence. (b) and (c) Cells were exposed to  $\gamma$ -irradiation (10 Gy, B) or UV light (60 J/m<sup>2</sup> C) and collected after the indicated time (upper panels) or irradiated with the indicated doses and harvested one hour later (lower panels). For immunofluorescence analysis cells were fixed 8 hours after treatment (B, 10 Gy, C, 60 J/m<sup>2</sup>). (d) The indicated EBV-transformed lymphoblast lines from a normal individual (PD7) or from various Fanconi Anemia patients were either treated with 40ng/ml of Mitomycin C continuously (lanes 1-21) or exposed to 15 Gy of  $\gamma$ -irradiation (lanes 22-33) and processed for FANCD2 immunoblotting. The upregulation of FANCD-L after MMC or IR treatment was seen in PD7 (lanes 2-5) and in the corrected FA-A cells (lanes 28-33), but was not observed in any of the mutant Fanconi Anemia cell lines. Similarly, IR-induced FANCD2 nuclear foci were not detected in FA fibroblasts (FA-G + IR) but were restored after functional complementation (FA-G + FANCG).

Figure 5 shows co-localization of activated FANCD2 and BRCA1 in Discrete Nuclear Foci following DNA damage. HeLa cells were untreated or exposed to Ionizing Radiation (10 Gy) as indicated, and fixed 8 hours later. (a) Cells were double-stained with the D-9 monoclonal anti-BRCA1 antibody (green, panels a, d, g, h) and the rabbit polyclonal anti-FANCD2 antibody (red, panels b, e, h, k), and stained cells were analyzed by immunofluorescence. Where green and red signals overlap (Merge, panels c, f, i, l) a yellow pattern is seen, indicating colocalization of BRCA1 and FANCD2. (b) Co-immunoprecipitation of FANCD2 and BRCA1. HeLa cells were untreated (- IR) or exposed to 15 Gy of  $\gamma\gamma$  irradiation (+ IR) and collected 12 hours later. Cell lysates were prepared, and cellular proteins were immunoprecipitated with either the monoclonal FANCD2 antibody (FI-17, lanes 9-10), or any one of three independently-derived monoclonal antibodies to human BRCA1 (lanes 3-8): D-9 (Santa Cruz), Ab-1 and Ab-3 (Oncogene Research Products). The same amount of purified mouse IgG (Sigma) was used in control samples (lanes 1-2). Immune complexes were resolved by SDS-PAGE and were immunoblotted with anti-FANCD2 or anti-BRCA1 antisera. The FANCD-L isoform preferentially coimmunoprecipitated with BRCA1.

Figure 6 shows the co-localization of activated FANCD2 and BRCA1 in discrete nuclear foci during S phase. (a) HeLa cells were synchronized in late G1 with mimosine and released into S phase. S phase cells were double-stained with the monoclonal anti-

BRCA1 antibody (green, panels a, d) and the rabbit polyclonal anti-FANCD2 antibody (red, panels b, e), and stained cells were analyzed by immunofluorescence. Where green and red signals overlap (merge, panels c, f), a yellow pattern is seen, indicating co-localization of BRCA1 and FANCD2. (b) HeLa cells synchronized in S phase were  
5 either untreated (a, b, k, l) or exposed to IR (50 Gy, panels c, d, m, n), MMC (20  $\mu\text{g}/\text{ml}$ , panels e, f, o, p), or UV (100  $\text{j}/\text{m}^2$ , panels g, h, q, r) as indicated and fixed 1 hour later. Cells were subsequently immunostained with an antibody specific for FANCD2 or BRCA1.

Figure 7 shows that FANCD2 forms foci on synaptonemal complexes that can  
10 co-localize with BRCA1 during meiosis I in mouse spermatocytes. (a) Anti-SCP3 (white) and anti-FANCD2 (red) staining of synaptonemal complexes in a late pachytene mouse nucleus. (b) SCP3 staining of late pachytene chromosomes. (c) Staining of this spread with preimmune serum for the anti-FANCD2 E35 antibody. (d) Anti-SCP3 staining of synaptonemal complexes in a mouse diplotene nucleus. (e) Costaining of this  
15 spread with E35 anti-FANCD2 antibody. Note staining of both the unpaired sex chromosomes and the telomeres of the autosomes with anti-FANCD2. (f) Costaining of this spread with anti-BRCA1 antibody. The sex chromosomes are preferentially stained. (g) Anti-FANCD2 staining of late pachytene sex chromosome synaptonemal complexes. (h) Anti-BRCA1 staining of the same complexes. (i) Anti-FANCD2 (red) and anti-  
20 BRCA1 (green) co-staining (co-localization reflected by yellow areas).

Figure 8 provides a schematic interaction of the FA proteins in a cellular pathway. The FA proteins (A, C, and G) bind in a functional nuclear complex. Upon activation of this complex, by either S phase entry or DNA damage, this complex enzymatically modifies (monoubiquitinates) the D protein. According to this model, the  
25 activated D protein is subsequently targeted to nuclear foci where it interacts with the BRCA1 protein and other proteins involved in DNA repair.

Figure 9 shows a Northern blot of cells from heart, brain, placenta, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, uterus, small intestine, colon and peripheral blood lymphocytes from a human adult and brain, lung, liver and kidney  
30 from a human fetus probed with a full-length FANCD2 cDNA and exposed for 24 hours.

Figure 10 shows allele specific assays for mutation analysis of 2 FANCD2 families where the family pedigrees (a, d) and panels b, c, e and f are vertically aligned such that the corresponding mutation analysis is below the individual in question. Panels a-c depict the PD20 and panels d-f the VU008 family. Panels b and e show the

segregation of the maternal mutations as detected by the creation of a new *MspI* site (PD20) or *DdeI* site (VU008). The paternally inherited mutations in both families were detected with allele specific oligonucleotide hybridization (panels c and f).

Figure 11 shows a Western blot analysis of the FANCD2 protein in human Fanconi Anemia cell lines. Whole cell lysates were generated from the indicated fibroblast and lymphoblast lines. Protein lysates (70  $\mu$ g) were probed directly by immunoblotting with the anti-FANCD2 antiserum. The FANCD2 proteins (155 kD and 162 kD) are indicated by arrows. Other bands in the immunoblot are non-specific. (a) Cell lines tested included wild-type cells (lanes 1,7), PD20 Fibroblasts (lane 2), PD20 lymphoblasts (lane 4), revertant MMC-resistant PD20 lymphoblasts (lane 5, 6), and chromosome 3p complemented PD20 fibroblasts (lane 3). Several other FA group D cell lines were analyzed including HSC62 (lane 8) and VU008 (lane 9). FA-A cells were HSC72 (lane 10), FA-C cells were PD4 (lane 11), and FA-G cells were EUFA316 (lane 12). (b) Identification of a third FANCD2 patient. FANCD2 protein was readily detectable in wild-type and FA group G cells but not in PD733 cells. (c) Specificity of the antibody. PD20i cells transduced with a retroviral FANCD2 expression vector displayed both isoforms of the FANCD2 protein (lane 4) in contrast to empty vector controls (lane 3) and untransfected PD20i cells (lane 2). In wild-type cells the endogenous FANCD2 protein (two isoforms) was also immunoreactive with the antibody (lane 1).

Figure 12 shows functional complementation of FA-D2 cells with the cloned FANCD2 cDNA. The SV40-transformed FA-D2 fibroblast line, PD20i, was transduced with pMMP-puro (PD20 + vector) or pMMP-FANCD2 (PD20 + FANCD2wt). Puromycin-selected cells were subjected to MMC sensitivity analysis. Cells analyzed were the parental PD20F cells ( $\Delta$ ), PD20 corrected with human chromosome 3p ( $\circ$ ), and PD20 cells transduced with either pMMP-puro ( ) or pMMP-FANCD2(wt)-puro ( $\blacklozenge$ ).

Figure 13 shows a molecular basis for the reversion of PD20 Lymphoblasts. (a) PCR primers to exons 5 and 6 were used to amplify cDNA. Control samples (right lane) yielded a single band of 114 bp, whereas PD20 cDNA (left lane) showed 2 bands, the larger reflecting the insertion of 13 bp of intronic sequence into the maternal allele. Reverted, MMC resistant lymphoblasts (middle lane) from PD20 revealed a third, in-frame splice variant of 114 + 36 bp (b) Schematic representation of splicing at the FANCD2 exon 5/intron 5 boundary. In wild-type cDNA 100% of splice events occur at

the proper exon/intron boundary, whereas the maternal A->G mutation (indicated by arrow) leads to aberrant splicing, also in 100%. In the reverted cells all cDNAs with the maternal mutation also had a second sequence change (fat arrow) and showed a mixed splicing pattern with insertion of either 13 bp (~40% of mRNA) or 36 bp (~60% of mRNA).

Figure 14 shows an FANCD2 Western blot of cancer cell lines derived from patients with ovarian cancer.

Figure 15 shows a sequence listing for amino acid sequence of human FANCD2 and alignment with fly and plant homologues using the BEAUTY algorithm (Worley, et al., 1995, Genome Res. Vol. 5, pp.173-184). (SEQ. ID. NO. 1-3) Black boxes indicate amino acid identity and gray similarity. The best alignment scores were observed with hypothetical proteins in *D melanogaster* ( $p=8.4 \times 10^{-58}$ , accession number AAF55806) and *A thaliana* ( $p=9.4 \times 10^{-45}$ , accession number B71413).

Figure 16 is the FANCD cDNA sequence -63 to 5127 nucleotides (SEQ ID NO: 5) and polypeptide encoded by this sequence from amino acid 1 to 1472 (SEQ ID NO: 4).

Figure 17 is the nucleotide sequence for FANCD-S.ORF (SEQ ID NO: 187) compared with FANCD cDNA (SEQ ID NO: 188).

Figure 18 is the nucleotide sequence for human FANCD2-L (SEQ ID NO: 6).

Figure 19 is the nucleotide sequence for human FANCD2-S (SEQ ID NO: 7).

Figure 20 is the nucleotide sequence for mouse FANCD2 (SEQ ID NO: 8).

### **Detailed Description of Specific Embodiments**

Definitions. As used in this description and the accompanying claims, the following terms shall have the meanings indicated, unless the context otherwise requires:

“FANCD2-L therapeutic agent” shall mean any of a protein isoform, and includes a peptide, a peptide derivative, analogue or isomer of the FANCD2-L protein and further include any of a small molecule derivative, analog, isomer or agonist that is functionally equivalent to FANCD2-L. Also included in the definition is a nucleic acid encoding FANCD2 which may be a full length or partial length gene sequence or cDNA or may be a gene activating nucleic acid or a nucleic acid binding molecule including an aptamer of antisense molecule which may act to modulate gene expression.

“Nucleic acid encoding FANCD-2” shall include the complete cDNA or genomic sequence of FANCD2 or portions thereof for expressing FANCD2-L protein as defined

above. The nucleic acid may further be included in a nucleic acid carrier or vector and includes nucleic acid that has been suitably modified for effective delivery to the target site.

5 “Stringent conditions of hybridization” will generally include temperatures in excess of 30°C, typically in excess of 37°C, and preferably in excess of 45°C. Stringent salt conditions will ordinarily be less than 1000 mM, typically less than 500 mM, and preferably less than 200 mM.

10 “Substantial homology or similiarity” for a nucleic acid is when a nucleic acid or fragment thereof is “substantially homologous” (“or substantially similar”) to another if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 60% of the nucleotide bases, usually at least about 70%, more usually at least about 80.

15 “Antibodies” includes polyclonal and/or monoclonal antibodies and fragments thereof including single chain antibodies and including single chain antibodies and Fab fragments, and immunologic binding equivalents thereof, which have a binding specificity sufficient to differentiate isoforms of a protein. These antibodies will be useful in assays as well as pharmaceuticals.

20 “Isolated” is used to describe a protein, polypeptide or nucleic acid which has been separated from components which accompany it in its natural state. An “isolated” protein or nucleic acid is substantially pure when at least about 60 to 75% of a sample exhibits a single aminoacid or nucleotide sequence.

25 “Regulatory sequences” refers to those sequences normally within 100 kb of the coding region of a locus, but they may also be more distant from the coding region, which affect the expression of the gene (including transcription of the gene, and translation, splicing, stability or the like of the messenger RNA).

30 “Polynucleotide” includes RNA, cDNA, genomic DNA, synthetic forms, and mixed polymers, both sense and antisense strands, and may be chemically or biochemically modified or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those skilled in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.), charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), pendent moieties

(e.g., polypeptides), and modified linkages (e.g., alpha anomeric nucleic acids, etc.). Also included are synthetic molecules that mimic nucleic acids in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions.

“Mutation” is a change in nucleotide sequence within a gene, or outside the gene in a regulatory sequence compared to wild type. The change may be a deletion, substitution, point mutation, mutation of multiple nucleotides, transposition, inversion, frame shift, nonsense mutation or other forms of aberration that differentiate the nucleic acid or protein sequence from that of a normally expressed gene in a functional cell where expression and functionality are within the normally occurring range.

“Subject” refers to an animal including mammal, including human.

“Wild type FANCD2” refers to a gene that encodes a protein or an expressed protein capable of being monoubiquitinated to form FANCD2-L from FANCD-S within a cell.

We have found that some Fanconi Anemia has similarities with a group of syndromes including ataxia telangiectasia (AT), Xeroderma pigmentosum (XP), Cockayne syndrome (CS), Bloom’s syndrome, myelodysplastic syndrome, aplastic anemia, cancer susceptibility syndromes and HNPCC (see Table 2). These syndromes have an underlying defect in DNA repair and are associated with defects in maintenance of chromosomal integrity. Defects in pathways associated with DNA repair and maintenance of chromosomal integrity result in genomic instability, and cellular sensitivity to DNA damaging agents such as bifunctional alkylating agents that cause intrastrand crosslinking. Moreover, deficiencies in DNA repair mechanisms appear to substantially increase the probability of initiating a range of cancers through genetic rearrangements. This observation is pertinent with regard to the clinical use of DNA cross-linking drugs including mitomycin C, cisplatin, cyclophosphamide, psoralen and UVA irradiation.

Although Fanconi Anemia is a rare disease, the pleiotropic effects of FA indicate the importance of the wild type function of FA proteins in the pathway for diverse cellular processes including genome stability, apoptosis, cell cycle control and resistance to DNA crosslinks. The cellular abnormalities in FA include sensitivity to cross-linking agents, prolongation of G2 phase of cell cycle, sensitivity to oxygen including poor growth at ambient O<sub>2</sub>, overproduction of O<sub>2</sub> radicals, deficient O<sub>2</sub> radical defense, deficiency in superoxide dismutase; sensitivity to ionizing radiation (G2 specific); overproduction of tumor necrosis factor, direct defects in DNA repair including

accumulations of DNA adducts, and defects in repair of DNA cross-links, genomic instability including spontaneous chromosome breakage, and hypermutability by deletion mechanism, increased apoptosis, defective p53 induction, intrinsic stem cell defect, including decreased colony growth *in vitro*; and decreased gonadal stem cell survival.

These features are reflective of the involvement of FA in maintenance of hematopoietic and gonadal stem cells, as well as the normal embryonic development of many different structures, including the skeleton and urogenital systems. Cell samples from patients were analyzed to determine defects in the FA complementation group D. Lymphoblasts from one patient gave rise to the PD20 cell line which was found to be mutated in a different gene from HSC 62 derived from another patient with a defect in the D complementation group. Mutations from both patients mapped to the D complementation group but to different genes hence the naming of two FANCD proteins-FANCD1 (HSC 62) and FANCD2 (PD20) (Timmers et al., (2001) Molecular Cell, Vol. 7, pp. 241-248).. We have shown that FANCD2 is the endpoint of the FA pathway and is not part of the FA nuclear complex nor required for its assembly or stability and that FANCD2 exists in two isoforms, FANCD2-S and FANCD2-L. We have also shown that transformation of the protein short form (FANCD2-S) to the protein long form (FANCD2-L) occurs in response to the FA complex (Fig. 8). Defects in particular proteins associated with the FA pathway result in failure to make an important post translationally modified form of FANCD2 identified as FANCD2-L. The two isoforms of FANCD2 are identified as the short form and the long form.

Failure to make FANCD2-L correlates with errors in DNA repair and cell cycle abnormalities associated with diseases listed above.

To understand more about the role of FANCD2 in the aforementioned syndromes, we cloned the FANCD2 gene and determined the protein sequence. The FANCD2 gene has an open reading frame of 4,353 base pairs and forty four exons which encodes a novel 1451 amino acid nuclear protein, with a predicted molecular weight of 166kD. Western blot analysis revealed the existence of 2 protein isoforms of 162 and 155 kD. The sequence corresponding to the 44 Intron/Exon Junctions are provided in Table 6 (SEQ ID NO: 9-94).

Unlike previously cloned FA proteins, FANCD2 proteins from several non-vertebrate eukaryotes showed highly significant alignment scores with proteins in *D. melanogaster*, *A. thaliana*, and *C. elegans*. The drosophila homologue, has 28% amino

acid identity and 50% similarity to FANCD2 (Fig. and SEQ ID NO: 1-3) and no functional studies have been carried out in the respective species. No proteins similar to FANCD2 were found in *E. coli* or *S. cerevisiae*.

We obtained the FANCD2 DNA sequence (SEQ ID No: 5) by analyzing the  
5 chromosome 3p locus in PD20 and VU008, two FA cell lines having biallelic mutations  
in the FANCD2 gene (Fig. 10). The cell lines were assigned as complementation group  
D because lymphoblasts from the patients failed to complement HSC62, the reference  
cell line for group D. FANCD2 mutations were not detected in this group D reference  
cell line which indicates that the gene mutated in HSC62 is the gene encoding FANCD1  
10 and in PD20 and VU008 is FANCD2 (Fig. 11). Microcell mediated chromosome transfer  
was used to identify the mutations (Whitney et al. Blood (1995) Vol. 88, 49-58).  
Detailed analysis of five microcell hybrids containing small overlapping deletions  
encompassing the locus narrowed the candidate region of the FANCD2 gene to 200Kb.  
The FANCD2 gene was isolated as follows: Three candidate ESTs were localized in or  
15 near this *FANCD2* critical region. Using 5' and 3' RACE to obtain full-length cDNAs,  
the genes were sequenced, and the expression pattern of each was analyzed by northern  
blot. EST SGC34603 had ubiquitous and low level expression of a 5kb and 7 kb mRNA  
similar to previously cloned FA genes. Open reading frames were found for TIGR-  
A004X28, AA609512 and SGC34603 and were 234, 531 and 4413 bp in length  
20 respectively. All 3 were analyzed for mutations in PD20 cells by sequencing cloned RT-  
PCR products. Whereas no sequence changes were detected in TIGR-A004X28 and  
AA609512, five sequence changes were found in SGC34603. Next, we determined the  
structure of the SGC34603 gene by using cDNA sequencing primers on BAC 177N7  
from the critical region.

25 Based on the genomic sequence information, PCR primer pairs were designed  
(Table 7), the exons containing putative mutations were amplified, and allele-specific  
assays were developed to screen the PD20 family as well as 568 control chromosomes.  
Three of the alleles were common polymorphisms; however, 2 changes were not found  
in the controls and thus represented potential mutations (Table 3). The first was a  
30 maternally inherited A->G change at nt 376. In addition to changing an amino acid  
(S126G), this alteration was associated with mis-splicing and insertion of 13 bp from  
intron 5 into the mRNA. 43/43 (100%) independently cloned RT-PCR products with the  
maternal mutation contained this insertion, whereas only 3 % (1/31) of control cDNA  
clones displayed mis-spliced mRNA. The 13 bp insertion generated a frame-shift and

predicts a severely truncated protein only 180 amino acids in length. The second alteration was a paternally inherited missense change at position 1236 (R1236H). The segregation of the mutations in the PD20 core family is depicted in Figure 10. Because the SGC34603 gene of PD20 contained both a maternal and a paternal allele not present on 568 control chromosomes and because the maternal mutation was associated with mis-splicing in 100% of cDNAs analyzed, we concluded that SGC34603 is the *FANCD2* gene.

*The protein encoded by FANCD2 is absent in PD20:* To further confirm the identity of SGC34603 as *FANCD2*, an antibody was raised against the protein, and Western blot analysis was performed (Figure 11). The specificity of the antibody was shown by retroviral transduction and stable expression *FANCD2* in PD20 cells (Figure 11c). In wild-type cells this antibody detected two bands (155 and 162kD) which we call *FANCD2* -S and -L (best seen in Figure 11c). *FANCD2* protein levels were markedly diminished in all MMC-sensitive cell lines from patient PD20 (Figure 11a, lanes 2, 4) but present in all wild-type cell lines and FA cells from other complementation groups. Furthermore, PD20 cells corrected by microcell-mediated transfer of chromosome 3 also made normal amounts of protein (Fig. 11a, lane 3).

*Functional complementation of FA-D2 cells with the FANCD2 cDNA:* We next assessed the ability of the cloned *FANCD2* cDNA to complement the MMC sensitivity of FA-D2 cells (Figure 12). The full length *FANCD2* cDNA was subcloned into the retroviral expression vector, pMMP-puro, as previously described (Pulsipher, et al., (1998) Mol. Med., Vol. 4, pp. 468-479). The transduced PD-20 cells expressed both isoforms of the *FANCD2* protein, *FANCD2*-S and *FANCD2*-L (Fig. 12c). Transduction of FA-D2 (PD20) cells with pMMP-*FANCD2* corrected the MMC sensitivity of the cells. These results further show that the cloned *FANCD2* cDNA encodes the *FANCD2*-S protein, which can be post-translationally-modified to the *FANCD2*-L isoform. This important modification is discussed in greater detail below.

*Analysis of a phenotypically reverted PD20 clone:* We next generated additional evidence demonstrating that the sequence variations in PD20 cells were not functionally neutral polymorphisms. Towards this end we performed a molecular analysis of a revertant lymphoblast clone (PD20-cl.1) from patient PD20 which was no longer sensitive to MMC. Phenotypic reversion and somatic mosaicism are frequent findings in FA and have been associated with intragenic events such as mitotic recombination or compensatory frame-shifts. Indeed, ~60% of maternally derived SGC34603 cDNAs had

a novel splice variant inserting 36 bp of intron 5 sequence rather than the usually observed 13 bp (Figure 13). The appearance of this in-frame splice variant correlated with a de novo base change at position IVS5+6 from G to A (Figure 13) and restoration of the correct reading frame was confirmed by Western blot analysis. In contrast to all  
5 MMC sensitive fibroblasts and lymphoblasts from patient PD20, PD20-cl.1 produced readily detectable amounts of FANCD2 protein of slightly higher molecular weight than the normal protein.

*Analysis of cell lines from other "FANCD" patients:* The antibody was also used to screen additional FA patient cell lines, including the reference cell line for FA group  
10 D, HSC and 2 other cell lines identified as group D by the European Fanconi Anemia Registry (EUFAR). VU008 did not express the FANCD2 protein and was found to be a compound heterozygote, with a missense and nonsense mutation, both in exon 12, and not found on 370 control chromosomes (Table 3, Figure 11). The missense mutation appears to destabilize the FANCD2 protein, as there is no detectable FANCD2 protein in  
15 lysates from VU008 cells. A third patient PD733 also lacked FANCD2 protein (Figure 11b, lane 3) and a splice mutation leading to absence of exon 17 and an internal deletion of the protein was found. The correlation of the mutations with the absence of FANCD2 protein in cell lysates derived from these patients substantiates the identity of *FANCD2* as a FA gene. In contrast, readily detectable amounts of both isoforms of the FANCD2  
20 protein were found in HSC 62 (Fig.11a, lane 8) and VU423 cDNA and genomic DNA from both cell lines were extensively analyzed for mutations, and none were found. In addition, a whole cell fusion between VU423 and PD 20 fibroblasts showed complementation of the chromosome breakage phenotype (Table 5). Taken together these data show that FA group D are genetically heterogeneous and that the gene(s)  
25 defective in HSC 62 and VU423 are distinct from *FANCD2*.

The identification and sequencing of the *FANCD2* gene and protein provides a novel target for therapeutic development, diagnostic tests and screening assays for diseases associated with failure of DNA repair and cell cycle abnormalities including but not limited to those listed in Table 2.

30 The following description provides novel and useful insights into the biological role of *FANCD2* in the FA pathway which provides a basis for diagnosis and treatment of the aforementioned syndromes.

Evidence that FA cells have an underlying molecular defect in cell cycle regulation include the following: (a) FA cells display a cell cycle delay with 4N DNA

content which is enhanced by treatment with chemical crosslinking agents, (b) the cell cycle arrest and reduced proliferation of FA cells can be partially corrected by overexpression of a protein, SPHAR, a member of the cyclin family of proteins and (c) caffeine abrogates the G2 arrest of FA cells. Consistent with these results, caffeine constitutively activates cdc2 and may override a normal G2 cell cycle checkpoint in FA cells. Finally, the FANCC protein binds to the cyclin dependent kinase, cdc2. We propose that the FA complex may be a substrate or modulator of the cyclinB/cdc2 complex.

Additionally, evidence that FA cells have an underlying defect in DNA repair is suggested by (a) FA cells that are sensitive to DNA cross-linking agents and ionizing radiation (IR), suggesting a specific defect in the repair of cross-linked DNA or double strand breaks, (b) DNA damage of FA cells which results in a hyperactive p53 response, suggesting the presence of defective repair yet intact checkpoint activities; and (c) FA cells with a defect in the fidelity of non-homologous end joining and an increased rate of homologous recombination (Garcia-Higuera et al. Mol. Cell. (2001), Vol. 7, pp. 249-262), (Grompe et al., Hum.Mol.Genet. (2001) Vol. 10, pp. 1-7).

Despite these general abnormalities in cell cycle and DNA repair. The mechanism by which FA pathway regulates these activities has remained elusive. Here we show that the FANCD2 protein functions downstream of the FA protein complex. In the presence of the assembled FA protein complex, the FANCD2 protein is activated to a high molecular weight, monoubiquitinated isoform which appears to modulate an S phase specific DNA repair response. The activated FANCD2 protein accumulates in nuclear foci in response to DNA damaging agents and co-localizes and co-immunoprecipitates with a known DNA repair protein, BRCA1. These results resolve previous conflicting models of the FA pathway (D'Andrea et al., 1997) and demonstrate that the FA proteins cooperate in a cellular response to DNA damage.

The FA pathway includes the formation of the FA multisubunit nuclear complex which in addition to A/C/G, we have shown also includes FANCF as a subunit of the complex (Figure 8). The FA pathway becomes "active" during the S phase to provide S-phase specific repair response or checkpoint response. The normal activation of the FA pathway which relies on the FA multisubunit complex results in the regulated monoubiquitination of the phosphoprotein-FANCD2 via a phosphorylation step to a high molecular weight activated isoform identified as FANCD-2L (Fig. 1).

Monoubiquitination is associated with cell trafficking. FANCD2-L appears to modulate

an S-phase specific DNA repair response (Fig. 3). The failure of FA cells to activate the S phase-specific activation of FANCD2 is associated with cell cycle specific abnormalities. The activated FANCD2 protein accumulates in nuclear foci in response to the DNA damaging agents, MMC and IR, and colocalizes and co-immunoprecipitates with a known DNA repair protein, BRCA1 (Fig. 4-6). These results resolve previous conflicting models of FA protein function (D'Andrea et. al., 1997) and strongly support a role of the FA pathway in DNA repair.

We have identified for the first time, an association between FANCD2 isoforms with respect to the FA pathway and proteins that are known diagnostic molecules for various cancers. A similar pathway with respect to DNA damage for the BRCA 1 protein which is activated to a high molecular weight, post-translationally-modified isoform in S phase or in response to DNA damage suggests that activated FANCD2 protein interacts with BRCA1. More particularly, the regulated monoubiquitination of FANCD2 appears to target the FANCD2 protein to nuclear foci containing BRCA1.

FANCD2 co-immunoprecipitates with BRCA1, and may further bind with other "dot" proteins, such as RAD50, Mre11, NBS, or RAD51. Recent studies demonstrate that BRCA1 foci are composed of a large (2 Megadalton) multi-protein complex (Wang et al., (2000) Genes Dev., Vol. 14, pp. 927-939). This complex includes ATM, ATM substrates involved in DNA repair functions (BRCA1), and ATM substrates involved in checkpoint functions (NBS). It is further suggested that damage recognition and activation of the FA pathway involve kinases which respond to DNA damage including ATM, ATR, CHK1, or CHK2.

We have found that the DNA damaging reagents, IR and MMC, activate independent post-translational modifications of FANCD2 result in distinct functional consequences. IR activates the ATM --dependent phosphorylation of FANCD2 at Serine 222 resulting in an S phase checkpoint response. MMC activates the BRACA-1 dependent and FA pathway dependent monouniquitination of FANCD2 at lysine 561, resulting in the assembly of FANCD2/BRCA1 nuclear foci and MMC resistance.

FANCD2 therefore has two independent functional roles in the maintenance of chromosomal stability resulting from two discrete post-translational modifications provide a link between two additional cancer susceptibility genes (ATM and BRCA1) in a common pathway. Several additional lines of evidence support an interaction between FANCD2 and BRCA1. First, the BRCA1 (-/-) cell line, HCC1937 (Scully, et al., (1999) Mol. Cell, Vol. 4, pp. 1093-1099) has a "Fanconi Anemia-like" phenotype, with

chromosome instability and increased tri-radial and tetra-radial chromosome formations. Second, although FA cells form BRCA1 foci (and RAD51 foci) normally in response to IR, BRCA1 (-/-) cells have no detectable BRCA1 foci and a greatly decreased number of FANCD2 foci compared to normal cells. Functional complementation of BRCA1 (-/-) cells restored BRCA1 foci and FANCD2 foci to normal levels, and restored normal MMC resistance.

The amount of FANCD2-L is determined in part by the amount of FANCD2-S that is synthesized from the *fancd2* gene and in part by the availability of the FA complex to monoubiquitinated FANCD2-S to form FANCD2-L. The association of FANCD2-L with nuclear foci including BRCA and ATM and determining the role of FANCD2-L in DNA repair make this protein a powerful target for looking at potential cancer development in patients for a wide range of cancers. Such cancers include those that arise through lesions on chromosome 3p as well as cancers on other chromosomes such that mutations result in interfering with production of upstream members of the FA pathway such as FANCG, FANCC or FANCA. Cancer lines and primary cells from cancer patients including tumor biopsies are being screened for FANCD-L and abnormal levels of this protein is expected to correlate with early diagnosis of disease. Because FANCD2 protein is a final step in a pathway to DNA repair, it is envisaged that any abnormality in a protein in the one or more pathways that lead to the conversion of FANCD2 -S to FANCD-L will be readily detected by measuring levels of FANCD2. Moreover, levels of FANCD2 affect how other proteins such as BRCA and ATM functionally interact in the nucleus with consequences for the patient. Analysis of levels of FANCD2 in a patient is expected to aid a physician in a clinical decision with respect to understanding the class of cancer presented by the patient. For instance, if a cancer cell fails to generate the monoubiquitinated FANCD2-L isoform, the cell may have increased chromosome instability and perhaps increased sensitivity to irradiation or chemotherapeutic agents. This information will assist the physician in procedure improved treatment for the patient.

Fanconi Anemia is associated not only with a broad spectrum of different cancers but also with congenital abnormalities. Development of the fetus is a complex but orderly process. Certain proteins have a particularly broad spectrum of effects because they disrupt this orderly progression of development. The FA pathway plays a significant role in development and disruption of the FA pathway results in a multitude of adverse effects. Errors in the FA pathway are detectable through the analysis of the FANCD2-L

protein from fetal cells. FANCD2 represents a diagnostic marker for normal fetal development and a possible target for therapeutic intervention.

Consistent with the above, we have shown that FANCD2 plays a role in the production of viable sperm. FANCD2 forms foci on the unpaired axes of chromosomes XY bivalents in late pachytene and in diplotene murine spermatocytes. (Fig. 7) Interestingly, FANCD2 foci are also seen at the autosomal telomeres in diplonema. Taken together with the known fertility defects in FA patients and FA-C knockout mice, our observations suggest that activated FANCD2 protein is required for normal progression of spermatocytes through meiosis I. Most of the FANCD2 foci seen on the XY axes were found to colocalize with BRCA1 foci, suggesting that the two proteins may function together in meiotic cells. Like BRCA1, FANCD2 was detected on the axial (unsynapsed) elements of developing synaptonemal complexes. Since recombination occurs in synapsed regions, FANCD2 may function prior to the initiation of recombination, perhaps to help prepare chromosomes for synapsis or to regulate subsequent recombinational events. The relatively synchronous manner in which FANCD2 assembles on meiotic chromosomes, and forms dot structures in mitotic cells, suggests a role of FANCD2 in both mitotic and meiotic cell cycle control.

Embodiments of the invention are directed to the use of the post translationally modified isoform: FANCD-2L as a diagnostic target for determining the integrity of the FA pathway. Ubiquitination of FANCD2 and the formation of FANCD2 nuclear foci are downstream events in the FA pathway, requiring the function of several FA genes. We have found that biallelic mutations of any of the upstream FA genes (FANCA, FANCB, FANCC, FANCE, FANCF and FANCG) block the posttranslational modification of FANCD2 the unubiquitinated FANCD2 (FANCD2-S) form to the ubiquitinated (FANCD2-L). Any of these upstream defects can be overridden by transfecting cells with FANCD2 cDNA. (Fig. 1a)

We have demonstrated for the first time the existence of FANCD2 and its role in the FA pathway. We have shown that FANCD2 accumulates in nuclear foci in response to DNA damaging agents where it is associated with other DNA repair proteins such as BRCA 1 and ATM. We have also demonstrated that FANCD2 exists in two isoforms in cells where a reduction in one of the two isoforms, FANCD2- L is correlated with Fanconi Anemia and with increased cancer susceptibility. We have used these findings to propose a number of diagnostic tests for use in the clinic that will assist with patient care.

These tests include: (a) genetic and prenatal counseling for parents concerned about inherited Fanconi Anemia in a future offspring or in an existing pregnancy; (b) genetic counseling and immunodiagnostic tests for adult humans to determine increased susceptibility to a cancer correlated with a defective FA pathway; and (c) diagnosing an already existing cancer in a subject to provide an opportunity for developing treatment protocols that are maximally effective for the subject while minimizing side effects.

The diagnostic tests described herein rely on standard protocols known in the art for which we have provided novel reagents to test for FANCD2 proteins and nucleotide sequences. These reagents include antibodies specific for FANCD2 isoforms, nucleotide sequences from which vectors, probes and primers have been derived for detecting genetic alterations in the FANCD2 gene and cells lines and recombinant cells for preserving and testing defects in the FA pathway.

We have prepared monoclonal and polyclonal antibody preparations as described in Example 1 that are specific for FANCD2-L and FANCD2-S proteins. In addition, FANCD2 isoform specific antibody fragments and single chain antibodies may be prepared using standard techniques. We have used these antibodies in wet chemistry assays such as immunoprecipitation assays, for example Western blots, to identify FANCD2 isoforms in biological samples (Figure 1). Conventional immunoassays including enzyme linked immunosorbent assays (ELISA), radioimmune assays (RIA), immunoradiometric assays (IRMA) and immunoenzymatic assays (IEMA) and further including sandwich assays may also be used. Other immunoassays may utilize a sample of whole cells or lysed cells that are reacted with antibody in solution and optionally analyzed in a liquid state within a reservoir. Isoforms of FANCD2 can be identified *in situ* in intact cells including cell lines, tissue biopsies and blood by immunological techniques using for example fluorescent activated cell sorting, and laser or light microscopy to detect immunofluorescent cells. (Figures 1-7, 9-14). For example, biopsies of tissues or cell monolayers, prepared on a slide in a preserved state such as embedded in paraffin or as frozen tissue sections can be exposed to antibody for detecting FANCD2-L and then examined by fluorescent microscopy.

In an embodiment of the invention, patient-derived cell lines or cancer cell lines are analyzed by immunoblotting and immunofluorescence to provide a novel simple diagnostic test for detecting altered amounts of FANCD2 isoforms. The diagnostic test also provides a means to screen for upstream defects in the FA pathway and a practical alternative to the currently employed DEB/MMC chromosome breakage test for FA

because individuals with upstream defects in the FA pathway are unable to ubiquitinate FANCD2. Other assays may be used including assays that combine retroviral gene transfer to form transformed patient derived cell lines (Pulsipher et al. (1998), *Mol Med* 4, 468-79 together with FANCD2 immunoblotting to provide a rapid subtyping analysis of newly diagnosed patients with any of the syndromes described in Table 2, in particular, that of FA.

The above assays may be performed by diagnostic laboratories, or, alternatively, diagnostic kits may be manufactured and sold to health care providers or to private individuals for self-diagnosis. The results of these tests and interpretive information are useful for the healthcare provider in diagnosis and treatment of a patient's condition.

Genetic tests can provide for a subject, a rapid reliable risk analysis for a particular condition against an epidemiological baseline. Our data suggests that genetic heterogeneity occurs in patients with FA within the FANCD2 complementation group. We have found a correlation between genetic heterogeneity and disease as well as genetic heterogeneity and abnormal post-translational modifications that result in the presence or absence of FANCD2-L. This correlation provides the basis for prognostic tests as well as diagnostic tests and treatments for any of the syndromes characterized by abnormal DNA repair. For example, nucleic acid from a cell sample obtained from drawn blood or from other cells derived from a subject can be analyzed for mutations in the FANCD2 gene and the subject may be diagnosed to have an increased susceptibility to cancer

We have located the FANCD2 gene at 3p25.3 on chromosome 3p in a region which correlates to a high frequency of cancer. Cytogenetic and loss of heterozygosity (LOH) studies have demonstrated that deletions of chromosome 3p occur at a high frequency in all forms of lung cancer (Todd et al. *Cancer Res.* Vol. 57, pp. 1344-52). For example, homozygous deletions were found in three squamous cell lines within a region of 3p21. Homozygous deletions were also found in a small cell tumor at 3p12 and a 3p14.2. (Franklin et al. (1997) *Cancer Res.* 57:1344-52). The present mapping of FANCD2 is supportive of the theory that this chromosomal region contains important tumor suppressor genes. Further support for this has been provided by a recent publication of Sekine et al., *Human Molecular Genetics*, (2001) 10, pp. 1421-1429 who reported localization of a novel susceptibility gene for familial ovarian cancer to chromosome 3p22-p25. The reduction or absence of FANCD2-L is here proposed to be

diagnostic for increased risk of tumors resulting from mutations not only at the FANCD2 site (3p25.3) but also at other sites in the chromosomes possibly arising from defects in DNA repair following cell damage arising from exposure to environmental agents and normal aging processes.

5 As more individuals and families are screened for genetic defects in the FANCD2 gene, a data base will be developed in which population frequencies for different mutations will be gathered and correlations made between these mutations and health profile for the individuals so that the predictive value of genetic analysis will continually improve. An example of an allele specific pedigree analysis for FANCD2 is  
10 provided in Figure 10 for two families.

Diagnosis of a mutation in the FANCD2 gene may initially be detected by a rapid immunological assay for detecting reduced amounts of FANCD2-L proteins. Positive samples may then be screened with available probes and primers for defects in any of the genes in the FA pathway. Where a defect in the FANCD2 gene is implicated, primers or  
15 probes such as provided in Table 7 may be used to detect a mutation. In those samples, where a mutation is not detected by such primers or probes, the entire FANCD2 gene may be sequenced to determine the presence and location of the mutation in the gene.

Nucleic acid screening assays for use in identifying a genetic defect in the FANCD2 gene locus may include PCR and non PCR based assays to detect mutations.  
20 There are many approaches to analyzing cell genomes for the presence of mutations in a particular allele. Alteration of a wild-type FANCD2 allele, whether, for example, by point mutation, deletion or insertions can be detected using standard methods employing probes. (US Patent 6,033,857). Standard methods include: (a) fluorescent *in situ* hybridization (FISH) which may be used on whole intact cells.; and (b) allele specific  
25 oligonucleotides (ASO) may be used to detect mutations using hybridization techniques on isolated nucleic acid (Conner et al., 1989, Hum.Genet. 85: 55-74). Other techniques include (a) observing shifts in electrophoretic mobility of single-stranded DNA on non-denaturing polyacrylamide gels, (b) hybridizing a FANCD2 gene probe to genomic DNA isolated from the tissue sample, (c) hybridizing an allele-specific probe to genomic  
30 DNA of the tissue sample, (d) amplifying all or part of the FANCD2 gene from the tissue sample to produce an amplified sequence and sequencing the amplified sequence, (e) amplifying all or part of the FANCD2 gene from the tissue sample using primers for a specific FANCD2 mutant allele, (f) molecular cloning all or part of the FANCD2 gene from the tissue sample to produce a cloned sequence and sequencing the cloned

sequence, (g) identifying a mismatch between (i) a FANCD2 gene or a FANCD2 mRNA isolated from the tissue sample, and (ii) a nucleic acid probe complementary to the human wild-type FANCD2 gene sequence, when molecules (i) and (ii) are hybridized to each other to form a duplex, (h) amplification of FANCD2 gene sequences in the tissue sample and hybridization of the amplified sequences to nucleic acid probes which comprise wild-type FANCD2 gene sequences, (I) amplification of FANCD2 gene sequences in the tissue sample and hybridization of the amplified sequences to nucleic acid probes which comprise mutant FANCD2 gene sequences, (j) screening for a deletion mutation in the tissue sample, (k) screening for a point mutation in the tissue sample, (l) screening for an insertion mutation in the tissue sample, and (m) in situ hybridization of the FANCD2 gene of the tissue sample with nucleic acid probes which comprise the FANCD2 gene.

It is often desirable to scan a relatively short region of a gene or genome for point mutations: The large numbers of oligonucleotides needed to examine all potential sites in the sequence can be made by efficient combinatorial methods (Southern, E. M et al., (1994). Nucleic Acids Res. 22:,1368-1373). Arrays may be used in conjunction with ligase or polymerase to look for mutations at all sites in the target sequence.(US 6,307,039) Analysis of mutations by hybridization can be performed for example by means of gels, arrays or dot blots.

The entire gene may be sequenced to identify mutations. (US 6,033,857). Sequencing of the FANCD2 locus can be achieved using oligonucleotide tags from a minimally cross hybridizing set which become attached to their complements on solid phase supports when attached to target sequence (US 6,280,935).

Other approaches to detecting mutations in the FANCD2 gene include those described in US 6,297,010, US 6,287,772 and US 6,300,076). It is further contemplated that the assays may employ nucleic acid microchip technology or analysis of multiple samples using laboratories on chips.

A subject who has developed a tumor maybe screened using nucleic acid diagnostic tests or antibody based tests to detect a FANCD2 gene mutation or a deficiency in FANCD2-L protein. On the basis of such screening samples may be obtained from subjects having a wide range of cancers including melanoma, leukemia, astocytoma, glioblastoma, lymphoma, glioma, Hodgkins lymphoma, chronic lymphocyte leukemia and cancer of the pancreas, breast, thyroid, ovary, uterus, testis, pituitary, kidney, stomach, esophagus and rectum. The clinician has an improved ability to select

a suitable treatment protocol for maximizing the treatment benefit for the patient. In particular, the presence of a genetic lesion or a deficiency in FANCD2-L protein may be correlated with responsiveness to various existing chemotherapeutic drugs and radiation therapies.

5           New therapeutic treatments may be developed by screening for molecules that modulate the monoubiquitination of FANCD2-S to give rise to FANCD2-L in cell assays (Examples 11-12) and in knock-out mouse models (Example 10). Such molecules may include those that bind directly to FANCD2 or to molecules such as BRACA-2 that appears to interact with BRACA-1 which in turn appears to be activated by FANCD2.

10           In addition, to screening assays that rely on defects in the FANCD2 gene or protein, an observed failure of the ubiquitination reaction that is necessary for the formation of FANCD2-L may result from a defect in the FA pathway at any point preceding the post translational modification of FANCD2 including FANCD2-S itself. Knowing the terminal step in the reactions, enables a screening assay to be formulated in  
15           which small molecules are screened in cells containing "broken FA pathway" or *in vitro* until a molecule is found to repair the broken pathway. This molecule can then be utilized as a probe to identify the nature of the defect. It may further be used as a therapeutic agent to repair the defect. For example, we have shown that cell cycle arrest and reduced proliferation of FA cells can be partially corrected by overexpression of a  
20           protein, SPHAR, a member of the cyclin family of proteins This can form the basis of an assay which is suitable as a screen for identifying therapeutic small molecules.

          Cells which are deficient in the posttranslational modified FANCD2 are particularly sensitive to DNA damage. These cells may serve as a sensitive screen for determining whether a compound (including toxic molecules) has the capability for  
25           damaging DNA.. Conversely, these cells also serve as a sensitive screen for determining whether a compound can protect cells against DNA damage.

          FA patients and patients suffering from syndromes associated with DNA repair defects die from complications of bone marrow failure. Gene transfer is a therapeutic option to correct the defect. Multiple defects may occur throughout the FA pathway.  
30           We have shown that the terminal step is critical to proper functioning of the cell and the organism. In an embodiment of the invention, correction of defects anywhere in the FA pathway may be satisfactorily achieved by gene therapy or by therapeutic agents that target the transformation of FANCD2-S to FANCD2-L so that this transformation is successfully achieved.

Gene therapy may be carried out according to generally accepted methods, for example, as described by Friedman in "Therapy for Genetic Disease," T. Friedman, ed., Oxford University Press (1991), pp. 105-121. Targeted tissues for *ex vivo* or *in vivo* gene therapy include bone marrow for example, hematopoietic stem cells prior to onset of anemia and fetal tissues involved in developmental abnormalities. Gene therapy can provide wild-type FANCD2-L function to cells which carry mutant FANCD2 alleles. Supplying such a function should suppress neoplastic growth of the recipient cells-or ameliorate the symptoms of Fanconi Anemia.

The wild-type FANCD-2 gene or a part of the gene may be introduced into the cell in a vector such that the gene remains extrachromosomal. In such a situation, the gene may be expressed by the cell from the extrachromosomal location. If a gene portion is introduced and expressed in a cell carrying a mutant FANCD-2 allele, the gene portion may encode a part of the FANCD-2 protein which is required for non-neoplastic growth of the cell. Alternatively, the wild-type FANCD-2 gene or a part thereof may be introduced into the mutant cell in such a way that it recombines with the endogenous mutant FANCD-2 gene present in the cell.

Viral vectors are one class of vectors for achieving gene therapy. Viral-mediated gene transfer can be combined with direct *in vivo* gene transfer using liposome delivery, allowing one to direct the viral vectors to the tumor cells and not into the surrounding nondividing cells. Alternatively, a viral vector producer cell line can be injected into tumors (Culver et al., 1992). Injection of producer cells would then provide a continuous source of vector particles. This technique has been approved for use in humans with inoperable brain tumors.

The vector may be injected into the patient, either locally at the site of the tumor or systemically (in order to reach any tumor cells that may have metastasized to other sites). If the transfected gene is not permanently incorporated into the genome of each of the targeted tumor cells, the treatment may have to be repeated periodically.

Vectors for introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, (for example as disclosed in US 5,252,479 and PCT 93/07282, and US 6,303,379) and include viral vectors such as retroviruses, herpes viruses (US 6,287,557) or adenoviruses (US 6,281,010) or a plasmid vector containing the FANCD2 -L.

A vector carrying the therapeutic gene sequence or the DNA encoding the gene or piece of the gene may be injected into the patient either locally at the site of a tumor

or systemically so as to reach metastasized tumor cells. Targeting may be achieved without further manipulation of the vector or the vector may be coupled to a molecule having a specificity of binding for a tumor where such molecule may be a receptor agonist or antagonist and may further include a peptide, lipid (including liposomes) or saccharide including an oligopolysaccharide or polysaccharide) as well as synthetic targeting molecules. The DNA may be conjugated via polylysine to a binding ligand. If the transfected gene is not permanently incorporated into the genome of each of the targeted tumor cells, the treatment may have to be repeated periodically.

Methods for introducing DNA into cells prior to introduction into the patient may be accomplished using techniques such as electroporation, calcium phosphate co-precipitation and viral transduction as described in the art (US 6,033,857), and the choice of method is within the competence of the routine experimenter

Cells transformed with the wild-type FANCD 2 gene or mutant FANCD 2 gene can be used as model systems to study remission of diseases resulting from defective DNA repair and drug treatments which promote such remission.

As generally discussed above, the FANCD 2 gene or fragment, where applicable, may be employed in gene therapy methods in order to increase the amount of the expression products of such genes in abnormal cells. Such gene therapy is particularly appropriate for use in pre-cancerous cells, where the level of FANCD 2-L polypeptide may be absent or diminished compared to normal cells and where enhancing the levels of FANCD2-L may slow the accumulation of defects arising from defective DNA repair and hence postpone initiation of a cancer state. It may also be useful to increase the level of expression of the FANCD 2 gene even in those cells in which the mutant gene is expressed at a "normal" level, but there is a reduced level of the FANCD 2-L isoform.

The critical role of FANCD2-L in normal DNA repair provides an opportunity for developing therapeutic agents to correct a defect that causes a reduction in levels of FANCD2-L. One approach to developing novel therapeutic agents is through rational drug design. Rational drug design can provide structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors or enhancers) in order to fashion more active or stable forms of the polypeptide, or to design small molecules which enhance or interfere with the function of a polypeptide in vivo. Hodgson, 1991. Rational drug design may provide small molecules or modified polypeptides which have improved FANCD2-L activity or stability or which act as enhancers, inhibitors, agonists or antagonists of FANCD2-L.

activity. By virtue of the availability of cloned FANCD2 sequences, sufficient amounts of the FANCD2-L polypeptide may be made available to perform such analytical studies as x-ray crystallography. In addition, the knowledge of the FANCD2-L protein sequence provided herein will guide those employing computer modeling techniques in place of, or in addition to x-ray crystallography.

Peptides or other molecules which have FANCD2-L activity can be supplied to cells which are deficient in the protein in a therapeutic formulation. The sequence of the FANCD2-L protein is disclosed for several organisms (human, fly and plant) (SEQ ID NO: 1-3). FANCD2 could be produced by expression of the cDNA sequence in bacteria, for example, using known expression vectors with additional posttranslational modifications. Alternatively, FANCD2-L polypeptide can be extracted from FANCD2-L-producing mammalian cells. In addition, the techniques of synthetic chemistry can be employed to synthesize FANCD2-L protein. Other molecules with FANCD2-L activity (for example, peptides, drugs or organic compounds) may also be used as a therapeutic agent. Modified polypeptides having substantially similar function are also used for peptide therapy.

Similarly, cells and animals which carry a mutant FANCD2 allele or make insufficient levels of FANCD2-L can be used as model systems to study and test for substances which have potential as therapeutic agents. The cells which may be either somatic or germline can be isolated from individuals with reduced levels of FANCD2-L. Alternatively, the cell line can be engineered to have a reduced levels of FANCD2-L, as described above. After a test substance is applied to the cells, the DNA repair impaired transformed phenotype of the cell is determined.

The efficacy of novel candidate therapeutic molecules can be tested in experimental animals for efficacy and lack of toxicity. Using standard techniques, animals can be selected after mutagenesis of whole animals or after genetic engineering of germline cells or zygotes to form transgenic animals. Such treatments include insertion of mutant FANCD2 alleles, usually from a second animal species, as well as insertion of disrupted homologous genes. Alternatively, the endogenous FANCD2 gene of the animals may be disrupted by insertion or deletion mutation or other genetic alterations using conventional techniques (Capecchi, 1989 Science Vol. 244, pp 1288-1292; Valancius and Smithies, 1991). After test substances have been administered to the animals, the growth of tumors must be assessed. If the test substance prevents or

suppresses pathologies arising from defective DNA repair, then the test substance is a candidate therapeutic agent for the treatment of the diseases identified herein.

The present invention is described by reference to the following Examples, which are offered by way of illustration and are not intended to limit the invention in any manner. Standard techniques well known in the art or the techniques specifically described below were utilized.

All references cited herein are incorporated by reference.

## EXAMPLES

### 10 Example 1: Experimental protocols used in Examples 2-8

*Cell Lines and Culture Conditions.* Epstein-Barr virus (EBV) transformed lymphoblasts were maintained in RPMI media supplemented with 15% heat-inactivated fetal calf serum (FCS) and grown in a humidified 5% CO<sub>2</sub>-containing atmosphere at 37°C. A control lymphoblast line (PD7) and FA lymphoblast lines (FA-A (HSC72), FA-C (PD-4), FA-D (PD-20), FA-F (EUFA121), and FA-G (EUFA316)) have been  
15 previously described (de Winter et al., 1998, *Nat.Genet.* 20: 281-283) (Whitney et al., 1995, *Nat.Genet.* 11: 341-343) (Yamashita et al., 1994, *P.N.A.S.* 91: 6712-6716) (de Winter et al., 2000, *Am.J.Hum.Genet.* 67: 1306-1308). PD81 is a lymphoblast cell line from an FA-A patient. The SV40-transformed FA fibroblasts, GM6914, PD426,  
20 FAG326SV and PD20F, as well as HeLa cells, were grown in DMEM supplemented with 15% FCS. FA cells (both lymphoblasts and fibroblasts) were functionally complemented with pMMP retroviral vectors containing the corresponding FANC cDNAs, and functional complementation was confirmed by the MMC assay (Garcia-Higuera et al., 1999, *Mol.Cell.Biol.* 19: 4866-4873) (Kuang et al., 2000, *Blood* 96: 1625-1632).

*Cell Cycle Synchronization.* HeLa cells, GM6914 cells, and GM6914 cells corrected with the pMMP-FANCA retrovirus were synchronized by the double thymidine block method as previously described, with minor modifications (Kupfer et al., 1997, *Blood* 90: 1047-1054). Briefly, cells were treated with 2mM thymidine for 18 hours, thymidine-free media for 10 hours, and additional 2mM thymidine for 18 hours to  
30 arrest the cell cycle at the G1/S boundary. Cells were washed twice with PBS and then released in DMEM +15% FCS and analyzed at various time intervals.

Alternatively, HeLa cells were treated with 0.5 mM mimosine (Sigma) for 24 hours for synchronization in late G1 phase (Krude, 1999), washed twice with PBS, and released into DMEM +15% FCS. For synchronization in M phase, a nocodazole block

was used (Ruffner et al., 1999, Mol.Cell.Biol. 19: 4843-4854). Cells were treated with 0.1  $\mu$ g/ml nocodazole (Sigma) for 15 hours, and the non-adherent cells were washed twice with PBS and replated in DMEM +15%.

*Cell Cycle Analysis.* Trypsinized cells were resuspended in 0.5ml of PBS and fixed by adding 5ml of ice-cold ethanol. Cells were next washed twice with PBS with 1% bovine serum albumin fraction V (1%BSA/PBS) (Sigma), and resuspended in 0.24 ml of 1%BSA/PBS. After adding 30  $\mu$ l of 500  $\mu$ g/ml propidium iodide (Sigma) in 38 mM sodium citrate (pH7.0) and 30  $\mu$ l of 10mg/ml DNase free RNaseA (Sigma), samples were incubated at 37°C for 30 min. DNA content was measured by FACScan (Beckton Dickinson), and data were analyzed by the CellQuest and Modfit LT program (Becton Dickinson).

*Generation of an anti-FANCD2 antiserum.* A rabbit polyclonal antiserum against FANCD2 was generated using a GST-FANCD2 (N-terminal) fusion protein as an antigen source. A 5' fragment was amplified by polymerase chain reaction (PCR) from the full length FANCD2 cDNA with the primers (SEQ ID NO: 95) DF4EcoRI (5' AGCCTCgaattcGTTTCCAAAAGAAGACTGTCA-3') and (SEQ ID NO: 96) DR816Xh (5'-GGTATCctcgagTCAAGACGACAACCTTATCCATCA-3'). The resulting PCR product of 841 bp, encoding the amino-terminal 272 amino acids of the FANCD2 polypeptide was digested with EcoRI/XhoI and subcloned into the EcoRI/XhoI sites of the plasmid pGEX4T-1 (Pharmacia). A GST-FANCD2 (N-terminal) fusion protein of the expected size (54 kD) was expressed in E. coli strain DH5 $\gamma$ , purified over glutathione-S-sepharose, and used to immunize a New Zealand White rabbit. An FANCD2-specific immune antiserum was affinity-purified by passage over an AminoLink Plus column (Pierce) loaded with GST protein and by passage over an AminoLink Plus column loaded with the GST-FANCD2 (N-terminal) fusion protein.

*Generation of anti-FANCD2 MoAbs.* Two anti-FANCD2 monoclonal antibodies were generated as follows. Balb/c mice were immunized with a GST-FANCD2 (N-terminal) fusion protein, which was the same fusion protein used for the generation of the rabbit polyclonal antiserum (E35) against FANCD2. Animals were boosted with immunogen for the four days before fusion, splenocytes were harvested, and hybridization with myeloma cells was performed. Hybridoma supernatants were collected and assayed using standard ELISA assay as the initial screen and immunoblot analysis of FANCD2 as the secondary screen. Two anti-human FANCD2 monoclonal

antibodies (MoAbs) (FI17 and FI14) were selected for further study. Hybridoma supernatants from the two positive cell lines were clarified by centrifugation.

Supernatants were used as MoAbs for western blotting. MoAbs were purified using an affinity column for IgG. MoAbs were stored as 0.5 mg/ml stocks in phosphate buffered saline (PBS). Anti-HA antibody (HA.11) was from Babco.

*Immunoblotting.* Cells were lysed with 1X sample buffer (50 mM Tris-HCl pH6.8, 86mM 2-mercaptoethanol, 2% sodium dodecyl sulfate (SDS), boiled for 5 min, and subjected to 7.5% polyacrylamide SDS gel electrophoresis. After electrophoresis, proteins were transferred to nitrocellulose using a submerged transfer apparatus (BioRad) filled with 25 mM Tris base, 200 mM glycine, 20% methanol. After blocking with 5% non-fat dried milk in TBS-T (50mM Tris-HCl, pH 8.0, 150mM NaCl, 0.1% Tween 20) the membrane was incubated with the primary antibody diluted in TBS-T (1:1000 dilution for the affinity-purified anti-FANCD2 polyclonal antibody (E35) or anti-HA (HA.11), 1:200 dilution for the anti-FANCD2 mouse monoclonal antibody FI17), washed extensively and incubated with the appropriate horseradish peroxidase-linked secondary antibody (Amersham). Chemiluminescence was used for detection.

*Generation of DNA Damage.* Gamma irradiation was delivered using a Gamma cell 40 apparatus. UV exposure was achieved using a Stratalinker (Stratagene) after gently aspirating the culture medium. For Mitomycin C treatment cells were continuously exposed to the drug for the indicated time. Hydroxyurea (Sigma) was added to a final concentration of 1 mM for 24 hours.

*Detection of Monoubiquitinated FANCD2.* HeLa cells (or the FA-G fibroblasts, FAG326SV) were transfected using FuGENE6 (Roche), following the manufacturer's protocol. HeLa cells were plated onto 15 cm tissue culture dishes and were transfected with 15 µg of a HA-tagged ubiquitin expression vector (pMT 123) (Treier et al., 1994, Cell 78: 787-798) per dish. Twelve hours following transfection, cells were treated with the indicated concentration of MMC (0,10,40,160ng/ml) or the indicated dose of IR (0,5,10,10,20Gy). After 24 hour-incubation with MMC, or two hours after IR treatment, whole cell extracts were prepared in Lysis Buffer (50 mM TrisHCl pH 7.4, 150 mM NaCl, 1% (v/v) Triton X-100) supplemented with protease inhibitors (1 µg/ml leupeptin and pepstatin, 2 µg/ml aprotinin, 1 mM phenylmethylsulfonylfluoride) and phosphatase inhibitors (1 mM sodium orthovanadate, 10 mM sodium fluoride). Using the polyclonal antibody to FANCD2 (E35), immunoprecipitation (IP) was performed essentially as described (Kupfer et al., 1997) except that each IP was normalized to contain 4mg of

protein. As a negative control, preimmune serum from the same rabbit was used in IP reaction. Immunoblotting was done using anti-HA (HA.11), or anti-FANCD2 (FI17) monoclonal antibody.

*Ubiquitin Aldehyde Treatment.* HeLa cells were treated with 1mM hydroxyurea for 24 hours, and whole cell extracts were prepared in Lysis Buffer supplemented with protease inhibitors and phosphatase inhibitors. 200 $\mu$ g of cell lysate in 67 $\mu$ l of reaction with 6.7 $\mu$ l of 25 $\mu$ M ubiquitin aldehyde (BostonBiochem) in DMSO or with 6.7 $\mu$ l of DMSO were incubated at 30°C or at 37°C for the indicated periods. Sixty-seven microliters of 2X sample buffer was added to each sample, and the samples were boiled for 5 min, separated by 7.5% SDS-PAGE, and immunoblotted for FANCD2 using the FI17 monoclonal anti-human FANCD2 antibody.

*Immunofluorescence Microscopy.* Cells were fixed with 2% paraformaldehyde in PBS for 20 min., followed by permeabilization with 0.3% Triton-X-100 in PBS (10 min). After blocking in 10% goat serum, 0.1% NP-40 in PBS (blocking buffer), specific antibodies were added at the appropriate dilution in blocking buffer and incubated for 2-4 hours at room temperature. FANCD2 was detected using the affinity-purified E35 polyclonal antibody (1/100). For BRCA1 detection, we used a commercial monoclonal antibody (D-9, Santa Cruz) at 2  $\mu$ g/ml. Cells were subsequently washed three times in PBS + 0.1% NP-40 (10-15 min each wash) and species-specific fluorescein or Texas red-conjugated secondary antibodies (Jackson ImmunoResearch) were diluted in blocking buffer (anti-mouse 1/200, anti-rabbit 1/1000) and added. After 1 hour at room temperature three more 10-15 min washes were applied and the slides were mounted in Vectashield (Vector laboratories). Images were captured on a Nikon microscope and processed using Adobe Photoshop software.

*Meiotic Chromosome Staining.* Surface spreads of pachytene and diplotene spermatocytes from male mice between the ages of 16 and 28 days old were prepared as described by (Peters et al., 1997). A polyclonal goat antibody to the mouse SCP3 protein was used to visualize axial elements and synaptonemal complexes in the meiotic preparations. The M118 mouse monoclonal antibody against mouse BRCA1 was generated by standard techniques, by immunizing mice with murine BRCA1 protein. The affinity-purified E35 rabbit polyclonal antibody was used in 1:200 dilution to detect FANCD. Antibody incubation and detection procedures were a modification of the protocol of (Moens et al., 1987, J.Cell.Biol. 105: 93-103) as described by (Keegan et al., 1996, Genes Dev. 10: 2423-2437). Combinations of donkey-anti mouse IgG-FTTC-

conjugated, Donkey-anti rabbit IgG-TRITC-conjugated, and Donkey-anti goat IgG-Cy5-conjugated secondary antibodies were used for detection (Jackson ImmunoResearch Laboratories). All preparations were counterstained with 4', 6'-diamino-2-phenylindole (DAPI, Sigma) and mounted in a DABCO (Sigma) antifade solution. The preparations were examined on a Nikon E1000 microscope (60X CFI Plan Apochromat and 100X CFI Plan Fluor oil-immersion objectives). Each fluorochrome (FITC, TRITC, Cy5 and DAPI) image was captured separately as an 800 x 1000 pixel 12-bit source image via IPLab software (Scanalytics) controlling a cooled-CCD camera (Princeton Instruments MicroMax) and the separate 12 bit grey scale images were resampled, 24-bit pseudocolored and merged using Adobe Photoshop.

Example 2: The FA genes interact in a common cellular pathway.

Normal lymphoblasts express two isoforms of the FANCD2 protein, a short form (FANCD2-S, 155 kD) and a long form (FANCD2-L, 162 kD). Figure 1 shows what happened when whole cell extracts were prepared from a lymphoblast line and cellular proteins were immunoprecipitated with an anti-FANCD2 antiserum. Normal wild type cells expressed two isoforms of the FANCD2 protein - a low molecular weight isoform FANCD2-S (155 kD isoform) and a high molecular weight isoform (FANCD2-L) (162kD isoform). FANCD2-S is the primary translation product of the cloned FANCD2 cDNA. We next evaluated a large series of FA lymphoblasts and fibroblasts for expression of the FANCD2 isoforms (Table 5). Correction of these FA cell lines with the corresponding FA cDNA resulted in functional complementation and restoration of the high molecular weight isoform, FANCD2-L.

As previously described, FA cells are sensitive to the DNA crosslinking agent, MMC, and in some cases, to ionizing radiation (IR). Interestingly, FA cells from multiple complementation groups (A, C, G, and F) only expressed the FANCD2-S isoform (Figure 1A, lanes 3, 7, 9, 11). FA cells from complementation groups B and E also express only the FANCD2-S. Functional correction of the MMC and IR sensitivity of these FA cells with the corresponding FANCD2 cDNA restored the FA protein complex (Garcia-Higuera et al., 1999) and restored the high molecular weight isoform (FANCD2-L) (Figure 1A, lanes 4,8,10,12). Taken together, these results demonstrate that the FA protein complex, containing FANCA, FANCC, FANCF, and FANCG, directly or indirectly regulates the expression of the two isoforms of FANCD2. The six cloned FA genes therefore appear to interact in a common pathway.

Example 3: The FA protein complex is required for the monoubiquitination of FANCD2

The high molecular weight isoform of FANCD2 could result from one or more mechanisms, including alternative splicing of the FANCD2 mRNA or post-translational modification(s) of the FANCD2 protein. Treatment with phosphatase did not convert FANCD2-L to FANCD2-S, demonstrating that phosphorylation alone does not account for the observed difference in their molecular mass.

In order to identify other possible post-translational modifications of FANCD2, we initially sought cellular conditions which regulate the conversion of FANCD2-S to FANCD2-L (Figure 1 B, C). Since FA cells are sensitive to MMC and IR, we reasoned that these agents might regulate the conversion of FANCD2-S to FANCD2-L in normal cells. Interestingly, HeLa cells treated with MMC (Figure 1B, lanes 1-6) or IR (Figure 1C, lanes 1-6) demonstrated a dose-dependent increase in the expression of the FANCD2-L isoform.

To determine whether FANCD2-L is a ubiquitinated isoform of FANCD2-S, we transfected HeLa cells with a cDNA encoding HA-Ubiquitin (Treier et al., 1994). Cellular exposure to MMC (Figure 1B, lanes 7-10) or IR (Figure 1C, lanes 7-10) resulted in a dose-dependent increase in the HA-ubiquitin conjugation of FANCD2. Only the FANCD2-L isoform, and not the FANCD2-S isoform, was immunoreactive with an anti-HA antibody. Although FANCD2 was not ubiquitinated in FA cells, FANCD2 ubiquitination was restored upon functional complementation of these cells. Although FANCD2 was not ubiquitinated in FA cells, FANCD2 ubiquitination was restored upon functional complementation of these cells. Since the FANCD2-S and FANCD2-L isoforms differ by 7kD, the FANCD2-L probably contains a single ubiquitin moiety (76 amino acids) covalently bound by an amide linkage to an internal lysine residue of FANCD2.

To confirm the monoubiquitination, we isolated FANCD2-L protein from HeLa cells and analyzed its tryptic fragments by mass spectrometry (Wu et al., 2000, Science 289, 11a). Ubiquitin tryptic fragments were unambiguously identified, and a site of monoubiquitination (K561 of FANCD2) was also identified. Interestingly, this lysine residue is conserved among FANCD2 sequences from human, *Drosophila*, and *C. elegans*, suggesting that the ubiquitination of this site is critical to the FA pathway in

multiple organisms. Mutation of this lysine residue, FANCD2 (K561R), resulted in loss of FANCD2 monoubiquitination.

5 Example 4: Formation of Nuclear Foci Containing FANCD2 requires an intact FA pathway.

We examined the immunofluorescence pattern of the FANCD2 protein in uncorrected, MMC-sensitive FA fibroblasts and functionally-complemented fibroblasts (Figure 2).

10 The corrected FA cells expressed both the FANCD2-S and FANCD2-L isoforms (Figure 2A, lanes 2,4,6,8). The endogenous FANCD2 protein was observed exclusively in the nucleus of human cells, and no cytoplasmic staining was evident (Figure 2B, a-h). The PD-20 (FA-D) cells have decreased nuclear immunofluorescence (Figure 2B, d), consistent with the decreased expression of FANCD2 protein in these cells by immunoblot (Figure 2A, lane 7). In PD-20 cells functionally-corrected with the  
15 FANCD2 gene by chromosome transfer, the FANCD2 protein stained in two nuclear patterns. Most corrected cells had a diffuse nuclear pattern of staining, and a minor fraction of cells stained for nuclear foci (see dots, panel h). Both nuclear patterns were observed with three independently-derived anti-FANCD2 antisera (1 polyclonal, 2 monoclonal antisera). FA fibroblasts from subtypes A, G, and C showed only the diffuse  
20 pattern of FANCD2 nuclear immunofluorescence. Functional complementation of these cells with the FANCA, FANCG, or FANCC cDNA, respectively, restored the MMC resistance of these cells (Table 6), and restored the nuclear foci in some cells. The presence of the high molecular weight FANCD2-L isoform therefore correlates with the presence of FANCD2 nuclear foci, suggesting that only the monoubiquitinated.  
25 FANCD2-L isoform is selectively localized to these foci.

Example 5: The FANCD2 protein is localized to nuclear foci during S phase of the Cell Cycle

30 Since only a fraction of the asynchronous functionally-complemented cells contained FANCD2 nuclear foci, we reasoned that these foci might assemble at discrete times during the cell cycle. To test this hypothesis, we examined the formation of the FANCD2-L isoform and FANCD2 nuclear foci in synchronized cells (Figure 3). HeLa cells were synchronized at the G1/S boundary, released into S phase, and analyzed for formation of the FANCD2-L isoform (Figure 3A). The FANCD2-L isoform was

expressed specifically during late G1 phase and throughout S phase. Synchronized, uncomplemented FA cells (FA-A fibroblasts, GM6914) expressed normal to increased levels of FANCD2-S protein but failed to express FANCD2-L at any time during the cell cycle. Functional complementation of these FA-A cells by stable transfection with the FANCA cDNA restored S phase-specific expression of FANCD2-L. The S phase-specific expression of the FANCD2-L isoform was confirmed when HeLa cells were synchronized by other methods, such as nocodazole arrest (Figure 3B) or mimosine exposure (Figure 3B). Cells arrested in mitosis did not express FANCD2-L, suggesting that the FANCD2-L isoform is removed or degraded prior to cell division (Figure 3B, mitosis). Taken together, these results demonstrate that the monoubiquitination of the FANCD2 protein is highly regulated during the cell cycle, and that this modification requires an intact FA pathway.

The cell cycle dependent expression of the FANCD2-L isoform also correlated with the formation of FANCD2 nuclear foci (Figure 3 C). Nocodazole arrested (mitotic) cells express no FANCD2-L isoform and exhibit no FANCD2 nuclear foci (Figure 3C, 0 hour). When these synchronized cells were allowed to traverse S phase (15 to 18 hours), an increase in FANCD2 nuclear foci was observed.

Example 6: The FANCD2 protein is localized to nuclear foci in response to DNA damage.

We examined the accumulation of the FANCD2-L isoform and FANCD2 nuclear foci in response to DNA damage (Figure 4). Previous studies have shown that FA cells are sensitive to agents which cause DNA interstrand crosslinks (MMC) or double strand breaks (IR) but are relatively resistant to ultraviolet light (UV) and monofunctional alkylating agents. MMC activated the conversion of FANCD2-S to FANCD2-L in asynchronous HeLa cells (Figure 4A). Maximal conversion to FANCD2-L occurred 12-24 hours after MMC exposure, correlating with the time of maximal FANCD2 nuclear focus formation. There was an increase in FANCD2 nuclear foci corresponding to the increase in FANCD2-L. Ionizing radiation also activated a time-dependent and dose-dependent increase in FANCD2-L in HeLa cells, with a corresponding increase in FANCD2 foci (Figure 4B). Surprisingly, ultraviolet (UV) light activated a time-dependent and dose-dependent conversion of FANCD2-S to FANCD2-L, with a corresponding increase in FANCD2 foci (Figure 4C).

We tested the effect of DNA damage on FA cells (Figure 4D). FA cells from multiple complementation groups (A, C, and G) failed to activate the FANCD2-L isoform and failed to activate FANCD2 nuclear foci in response to MMC or IR exposure. These data suggest that the cellular sensitivity of FA cells results, at least in part, from their failure to activate FANCD2-L and FANCD2 nuclear foci.

Example 7: Co-localization of activated FANCD2 and BRCA1 protein.

Like FANCD2, the breast cancer susceptibility protein, BRCA1, is upregulated in proliferating cells and is activated by post-translational modifications during S phase or in response to DNA damage. BRCA has a carboxy terminus 20aminoacids which contain a highly acidic HMG-like domain suggesting a possible mechanism for chromatin repair. The BRCA1 protein colocalizes in IR-inducible foci (IRIFs) with other proteins implicated in DNA repair, such as RAD51 or the NBS/Mre11/RAD50 complex. Cells with biallelic mutations in BRCA1 have a defect in DNA repair and are sensitive to DNA damaging agents such as IR and MMC (Table 5). Taken together, these data suggest a possible functional interaction between the FANCD2 and BRCA1 proteins. BRCA foci are large (2mDa) multiprotein complexes including ATM and ATM substrates involved in DNA repair (BRCA1) and checkpoint functions (NBS)

In order to determine whether the activated FANCD2 protein colocalizes with the BRCA1 protein, we performed double immunolabelling of HeLa cells (Figure 5). In the absence of ionizing radiation, approximately 30-50% of cells contained BRCA1 nuclear foci (Figure 5A). In contrast, only rare cells traversing S phase contained FANCD2 dots (b,e). These nuclear foci were also immunoreactive with antisera to both BRCA1 and FANCD2 (c,f). Following IR exposure, there was an increase in the number of cells containing nuclear foci and the number of foci per cell. These nuclear foci were larger and more fluorescent than foci observed in the absence of IR. Again, these foci contained both BRCA1 and FANCD2 protein (i, l). An interaction of FANCD2-L and BRCA1 was further confirmed by coimmunoprecipitation of the proteins (Figure 5B) from exponentially growing HeLa cells exposed to IR.

We examined the effect of BRCA1 expression on the formation of FANCD2-L and nuclear foci (Figure 6). The BRCA1 (-/-) cell line, HCC1937, expresses a mutant form of the BRCA1 protein with a carboxy terminal truncation. Although these cells expressed a low level of FANCD2-L (Figure 6A), IR failed to activate an increase in FANCD2-L levels. Also, these cells had a decreased number of IR-inducible FANCD2

foci (Figure 6B, panels c, d). Correction of these BRCA1 (-/-) cells by stable transfection with the BRCA1 cDNA restored IR-inducible FANCD2 ubiquitination and nuclear foci (Figure 6B, panels k, l). These data suggest that the wild-type BRCA1 protein is required as an “organizer” for IR-inducible FANCD2 dot formation and further suggests a functional interaction between the proteins.

**Example 8: Co-localization of FANCD2 and BRCA1 on Meiotic Chromosomes.**

The association of FANCD2 and BRCA1 in mitotic cells suggested that these proteins might also colocalize during meiotic prophase. Previous studies have demonstrated that the BRCA1 protein is concentrated on the unsynapsed/axial elements of human synaptonemal complexes in zygotene and pachytene spermatocytes. To test for a possible colocalization of FANCD2 and BRCA1 in meiotic cells, we examined surface spreads of late pachytene and early diplotene mouse spermatocytes for the presence of FANCD2 and BRCA1 protein (Figure 7). We found that the rabbit polyclonal anti-FANCD2 antibody E35 specifically stained the unpaired axes of the X and Y chromosomes in late pachynema (Figure 7a) and in diplonema (Figures 7d, 7e and 7g). Under the same experimental conditions, preimmune serum did not stain synaptonemal complexes (Figures 7b and 7c). The M118 anti-BRCA1 antibody stained the unpaired sex chromosomes in mouse pachytene and diplotene spermatocytes (Figures 7f and 7h). FANCD2 Ab staining of the unsynapsed axes of the sex chromosomes was interrupted, giving a beads-on-a-string appearance (Figure 7g). A consecutive examination of 20 pachytene nuclei indicated that most (~65%) of these anti-FANCD2 foci colocalized with regions of intense anti-BRCA1 staining, further supporting an interaction between these proteins (Figures 7g, 7h, and 7i). These results provide the first example of a FANC protein (activated FANCD2) which binds to chromatin.

**Example 9: Experimental protocols for obtaining and analyzing the DNA and protein sequence for FANCD2.**

Northern Hybridizations. Human adult and fetal multi-tissue mRNA blots were purchased from Clontech (Palo Alto, CA). Blots were probed with <sup>32</sup>P labeled DNA from EST clone SGC34603. Standard hybridization and washing conditions were used. Equal loading was confirmed by re-hybridizing the blot with an actin cDNA probe.

Mutation Analysis. Total cellular RNA was reverse transcribed using a commercial kit (Gibco/BRL). The 5' end section of FANCD2 was amplified from the resulting patient and control cDNA with a nested PCR protocol. The first round was performed with primers (SEQ ID NO: 97) MG471 5'-AATCGAAAACACTACGGGCG-3' and (SEQ ID NO: 98) MG457 5'-GAGAACACATGAATGAACGC-3'. The PCR product from this round was diluted 1:50 for a subsequent round using primers (SEQ ID NO: 99) MG492 5'-GGCGACGGCTTCTCGG AAGTAATTTAAG-3' and (SEQ ID NO: 100) MG472 5'-AGCGGCAGGAGGTTTATG-3'. The PCR conditions were as follows: 94°C for 3 min, 25 cycles of 94°C for 45 sec, 50°C for 45 sec, 72°C for 3 min and 5min of 72°C at the end. The 3' portion of the gene was amplified as described above but with primers, (SEQ ID NO: 101) MG474 5'-TGGCGGCAGACAGAAGTG-3' and (SEQ ID NO: 102) MG475 5'-TGGCGGCAGACAGAAGTG-3'. The second round of PCR was performed with (SEQ ID NO: 103) MG491 5'-AGAGAGCCAACCTGAGCGATG-3' and (SEQ ID NO: 104) MG476 5'-GTGCCAGACTCTGGTGGG-3'. The PCR products were gel-purified, cloned into the pT-Adv vector (Clontech) and sequenced using internal primers.

Allele specific assays: Allele specific assays were performed in the PD20 family and 290 control samples (=580 chromosomes). The PD20 family is of mixed Northern European descent and VU008 is a Dutch family. Control DNA samples were from unrelated individuals in CEPH families (n=95), samples from unrelated North American families with either ectodermal dysplasia (n=95) or Fanconi Anemia (n=94). The maternal nt376a→g mutation in the PD 20 family created a novel MspI restriction site. For genomic DNA, the assay involved amplifying genomic DNA using the primers (SEQ ID NO: 105) MG792 5'-AGGA GACACCCTTC CTATCC-3' located in exon 4 and (SEQ ID NO: 106) MG803 5'-GAAG TTGGCAAACAGACTG-3' which is in intron 5. The size of the PCR product was 340 bp, yielding two fragments of 283 bp and 57 bp upon MspI digestion if the mutation was present. For analysis of the reverted cDNA clones, PCR was performed using primers (SEQ ID NO: 107) MG924 5'-TGTCTTGTGAGCGTCTGCAGG-3' and (SEQ ID NO: 108) MG753 5'-AGGTT TTGATAATGGCAGGC-3'. The paternal exon 37 mutation (R1236H) in PD20 and exon12 missense mutation (R302W) in VU008 were tested by allele specific oligonucleotide (ASO) hybridization (Wu, et al., 1989, DNA 8: 135-142). For the exon 12 assay, genomic DNA was amplified with primers (SEQ ID NO: 109) MG979 5'-

ACTGGACTGTGCCTACCCACTATG-3' and (SEQ ID NO: 110) MG984 5'-  
 CCTGTGTGAGGATGAGCTCT-3'. Primers (SEQ ID NO: 171) MG818 5'-  
 AGAGGTAGGGAAGGAAGCTAC-3' and (SEQ ID NO: 172) MG813 5'- CCAAAGT  
 CCACTTCTTGAAG-3' were used for exon 37. Wild-type (SEQ ID NO: 111) (5'-  
 5 TTCTCCCGAAGCTCAG-3' for R302W and (SEQ ID NO: 112) 5'- TTTCTTCC  
 GTGTGATGA-3' for R1236H) and mutant SEQ ID NO: 111 (5'-  
 TTCTCCCAAAGCTGAG-3' R302W and SEQ ID NO: 112 5'-  
 TTTCTTCCATGTGATGA-3' for R1236H) oligonucleotides were end-labeled with  $\gamma^{32}\text{P}$ -  
 [ATP] and hybridized to dot-blotted target PCR products as previously ss novel DdeI  
 10 site. The wild-type PCR product digests into a 117 and 71bp product, whereas the  
 mutant allele yields three fragments of 56, 61 and 71bps in length. PCR in all of the  
 above assays was performed with 50 ng of genomic DNA for 37 cycles of 94<sup>0</sup>C for 25  
 sec, 50<sup>0</sup>C for 25 sec and 72<sup>0</sup>C for 35 sec.

Generation of an anti-FANCD2 antiserum: A rabbit polyclonal antiserum against  
 15 FANCD2 was generated using a GST-FANCD2 (N-terminal) fusion protein as an  
 antigen source. A 5' fragment was amplified by polymerase chain reaction (PCR) from  
 the full length FANCD2 cDNA with the primers (SEQ ID NO: 113) DF4EcoRI (5'-  
 AGCCTCgaattcGTTTCC AAAAGAAGACTGTCA-3') and (SEQ ID NO: 114)  
 DR816Xh (5'- GGTATCctcgagTCAAGA CGACA ACTTATCCATCA-3'). The  
 20 resulting PCR product of 841 bp, encoding the amino-terminal 272 amino acids of the  
 FANCD2 polypeptide was digested with EcoRI/XhoI and subcloned into the  
 EcoRI/XhoI sites of the plasmid pGEX4T-1 (Pharmacia). A GST-FANCD2 (N-  
 terminal) fusion protein of the expected size (54 kD) was expressed in *E. coli* strain  
 DH5 $\alpha$ , purified over glutathione-S-sepharose, and used to immunize a New Zealand  
 25 White rabbit. An FANCD2-specific immune antiserum was affinity-purified over an  
 AminoLink Plus column (Pierce) loaded with GST protein and over an AminoLink Plus  
 column loaded with the GST-FANCD2 (N-terminal) fusion protein.

Immunoblotting As in Example 1.

Cell Lines and Transfections: PD20i is an immortalized and PD733 a primary  
 30 FA fibroblast cell line generated by the Oregon Health Sciences Fanconi Anemia cell  
 repository (Jakobs, et al., 1996, *Somet.Cell.Mol.Genet.* 22: 151-157). PD20  
 lymphoblasts were derived from bone marrow samples. VU008 is a lymphoblast and  
 VU423 a fibroblast line generated by the European Fanconi Anemia Registry (EUFAR).

VU423i was an immortalized line derived by transfection with SV40 T-antigen (Jakobs, et al., 1996) and telomerase (Bodnar, et al., 1998, *Science* 279: 349-352). The other FA cell lines have been previously described. Human fibroblasts were cultured in MEM and 20% fetal calf serum. Transformed lymphoblasts were cultured in RPMI 1640  
5 supplemented with 15% heat-inactivated fetal calf serum.

To generate FANCD2 expression constructs, the full-length cDNA was assembled from cloned RT-PCR products in pBluescript and the absence of PCR induced mutations was confirmed by sequencing. The expression vectors pIRES-Neo, pEGFP-N1, pRevTRE and pRevTet-off were from ClonTech (Palo Alto, CA). The  
10 FANCD2 was inserted into the appropriate multi-cloning site of these vectors. Expression constructs were electroporated into cell line PD20 and a normal control fibroblast cell line, GM639 using standard conditions (van den Hoff, et al., 1992). Neomycin selection was carried out with 400  $\mu\text{g/ml}$  active G418 (Gibco).

Whole cell fusions: For the whole cell fusion experiments, a PD20 cell line  
15 (PD20i) resistant to hygromycin B and deleted for the HPRT locus was used (Jakobs, et al., 1997, *Somat. Cell. Mol. Genet.* Vol. 23, pp. 1-7). Controls included PD 24 (primary fibroblasts from affected sibling of PD20) and PD 319i (Jakobs, et al., 1997) (immortal fibroblasts from a non-A, C, D or G FA patient).  $2.5 \times 10^5$  cells from each cell line were mixed in a T25 flask and allowed to recover for 24 hours. The cells were washed with  
20 serum-free medium and then fused with 50% PEG for 1 min. After removal of the PEG, the cells were washed 3x with serum-free medium and allowed to recover overnight in complete medium without selection. The next day, cells were split 1:10 into selective medium containing 400  $\mu\text{g/ml}$  hygromycin B (Roche Molecular) and 1X HAT. After the selection was complete, hybrids were passaged once and then analyzed as described  
25 below.

Retroviral Transduction of FA-D2 cells and complementation analysis: The full length FANCD2 cDNA was subcloned into the vector, pMMP-puro (Pulsipher, et al., 1998). Retroviral supernatants were used to transduce PD20F, and puromycin resistant cells were selected. Cells were analyzed for MMC sensitivity by the crystal violet assay  
30 (Naf, et al., 1998).

Chromosome Breakage Analysis: Chromosome breakage analysis was performed by the Cytogenetics Core Lab at OHSU (Portland, OR). For the analysis (Cohen, et al., 1982) cells were plated into T<sub>25</sub> flasks, allowed to recover and then treated with 300 ng/ml of DEB for two days. After treatment, the cells were exposed to colcemid for 3

hours and harvested using 0.075 M KCl and 3:1 methanol:acetic acid. Slides were stained with Wright's stain and 50-100 metaphases were scored for radials.

Example 10: Mouse models for FA for use in screening potential therapeutic agents.

5 Murine models of FANCD2 can be made using homologous recombination in embryonic stem cells or targeted disruption as described in D'Andrea et al. (1997) 90: 1725-1736, and Yang et al. (2001) Blood 98; 1-6. The knockout of FANCD2 locus in mice is not a lethal mutation. These knock-out animals have increased susceptibility to cancer and furthermore display other symptoms characteristic of FA. It is expected that  
10 administering certain therapeutic agents to the knock-out mice will reduce their susceptibility to cancer. Moreover, it is expected that certain established chemotherapeutic agents will be identified that are more effective for treating knock-out mice who have developed cancers as a result of the particular genetic defect and this will also be useful in treating human subjects with susceptibility to cancer or who have  
15 developed cancers as a result of a mutation in the FANCD2 locus.

We can generate experimental mice models with targeted disruptions of FANCD2 using for example the approach described by Chen et al (1996) Nat. Genet. Vol. 12, pp. 448-451, for Fancc who created a disruption in an exon of the gene, and by Whitney et al (1996) Vol. 88, pp. 49-58, who used homologous recombination to create  
20 a disruption of an exon of the gene. In both animal models, spontaneous chromosome breakage and an increase in chromosome breaks in splenic lymphocytes in response to bifunctional alkylating agents are observed. In both models, Fancd2<sup>-/-</sup> mice have germ cell defects and decreased fertility. The Fancd2 murine knockout model is useful in examining (1) the role of the Fancd2 gene in the physiologic response of hematopoietic  
25 cells to DNA damage, (2) the in vivo effects of inhibitory cytokines on FA marrow cells.

129/Sv and C57BL can be generated following standard protocols. Mouse tail genomic DNA can be prepared as previously described and used as a template for polymerase chain reaction (PCR) genotyping.

Splenocytes can be prepared from 6-week-old mice of known *Fancd2* genotype.

5 The spleen is dissected, crushed in RPMI medium into a single-cell suspension, and filtered through a 70  $\mu$ m filter. Red cells are lysed in hypotonic ammonium chloride. The remaining splenic lymphocytes are washed in phosphate-buffered saline and resuspended in RPMI/10% fetal bovine serum plus phytohemagglutinin. Cells are tested for viability by the trypan blue exclusion assay. Cells are cultured for 24 hours in media  
10 and exposed to MMC or DEB for an additional 48 hours. Alternatively, cells are cultured for 50 hours, exposed to IR (2 or 4 Gy, as indicated), and allowed to recover for 12 hours before chromosome breakage or trypan blue exclusion (viability) analysis.

Mononuclear cells can be isolated from the femurs and tibiae of 4- to 6-week-old *Fancd2*<sup>+/-</sup> or *Fancd2*<sup>-/-</sup> mice, as previously described. A total of 2 x 10<sup>4</sup> cells were  
15 cultured in 1 mL of MethoCult M343 media (StemCell Technologies, Vancouver, BC) with or without MMC treatment. Colonies are scored at day 7, when most of the colonies belong to the granulocyte-macrophage colony-forming unit or erythroid burst-forming unit lineages. Each number are averaged from duplicate plates, and the data derived from 2 independent experiments.

20 Lymphocytes isolated from thymus, spleen, and peripheral lymph nodes are stained for T- or B-lymphocyte surface molecules with fluorescein isothiocyanate-conjugated anti-CD3, CD4, and CD19 and PE-conjugated anti-CD8, CD44, CD 45B, immunoglobulin M, and B220 (BD Pharmingen, CA). Stained cells were analyzed on a Counter Epics XL flow cytometry system.

25 Mice ovaries and testes were isolated and fixed in 4% paraformaldehyde and further processed by the core facility of the Department of Pathology at Massachusetts General Hospital.

30 Example 11: Screening assays using antibody reagents for detecting increased cancer susceptibility in human subjects

Blood samples or tissue samples can be taken from subjects for testing for the relative amounts of FANCD2-S compared to FANCD2-L and the presence or absence of FANCD2-L. Using antibody reagents specific for FANCD2-S and FANCD2-L proteins  
35 (Example 1), positive samples can be identified on Western blots as shown in Figure 14.

Other antibody assays may be utilized such as, for example, one step migration binding banded assays described in 5,654,162 and 5,073,484. Enzyme linked immunosorbent assays (ELISA), sandwich assays, radioimmune assays and other immunodiagnostic assays known in the art may be used to determine relative binding concentrations of  
5 FANCD2-S and FANCD2-L.

The feasibility of this approach is illustrated by the following:

#### FANCD2 Diagnostic Western Blot for Screening Human Cancer Cell Lines

Human cancer cell lines were treated with or without ionizing radiation (as indicated in Figure 14) and total cell proteins were electrophoresed, transferred to  
10 nitrocellulose and immunoblotted with the anti-FANCD2 monoclonal antibody of Example 1. Ovarian cancer cell line (TOV21G) expressed FANCD2-S but not FANCD2-L (see lanes 9, 10). This cell line has a deletion of human chromosome 3p overlapping the FANCD2 gene and is hemizygous for FANCD2 and is predicted to have a mutation in the second FANCD2 allele which therefore fails to be monoubiquitinated  
15 by the FA complex hence no FANCD2-L (lanes 9, 10). This example demonstrates that antibody based tests are suited for determining lesions in the FANCD2 gene which lead to increased cancer susceptibility.

#### 20 Example 12: Screening assays using nucleic acid reagents for detecting increased cancer susceptibility in human subjects

Blood samples or tissue samples can be taken from subjects and screened using sequencing techniques or nucleic acid probes to determine the size and location of the genetic lesion if any in the genome of the subject. The screening method may include  
25 sequencing the entire gene or by using sets of probes or single probes to identify lesions. It is expected that a single lesion may predominant in the population but that other lesions may arise throughout the gene with low frequency as is the case for other genetic conditions such as cystic fibrosis and the P53 tumor suppressor gene.

The feasibility of this approach is illustrated by the following:

30 Peripheral blood lymphocytes are isolated from the patient using standard Ficoll-Hypaque gradients and genomic DNA is isolated from these lymphocytes. We use genomic PCR to amplify 44 exons of the human FANCD2 gene (see primer Table 7) and sequence the two FANCD2 alleles to identify mutations. Where such mutations are found, we distinguish these from benign polymorphisms by their ability to ablate the  
35 functional complementation of an FA-D2 indicator cell line.

**Table 1** Complementation groups and responsible genes of Fanconi Anemia

Subtype	Estimated percentage of patients	Responsible gene	Chromosome location	Number of exons	Protein product
A	66%	FANCA	16q24.3	43	163Kd
B	4.3%	FANCB	-	-	-
C	12.7%	FANCC	9q22.3	14	63Kd
D1	rare	FANCD1	-	-	-
D2	rare	FANCD2	3p25.3	44	155, 162kD
E	12.7%	FANCE	6p21.2-21.3	10	60kD
F	rare	FANCF	11p15	1	42kD
G	rare	FANCG (XRCC9)	9p13	14	68kD

TABLE 2

## Diseases of Genomic Instability

Disease	Damaging Agent	Neoplasm	Function
FA	Cross-linking agents leukemia, hepatic, gastrointestinal, and gynecological tumors	Acute myeloblastic	Unknown
XP	UV light	Squamous cell carcinomas	Excision repair
AT	Ionizing radiation	Lymphoma	Afferent pathway to p53
Bloom's Syndrome	Alkylating agents	Acute lymphoblastic leukemia	Cell-cycle regulation
Cockayne's Syndrome	UV light	Basal cell carcinoma repair	Transcription coupled
Hereditary non-polyposis colon cancer (HNPCC)	Unknown	Adenocarcinoma of colon, ovarian cancer	DNA mismatch repair

**Table 3**

<i>FANCD2</i> Sequence Alterations		
Mutations		
PD20	nt376a→g	S126G/splice
	nt3707g→a	R1236H
VU008	nt904c→t	R302W
	nt958c→t	Q320X
PD733	deletion of exon 17	
Polymorphisms		
	nt1122a→g	V374V
	nt1440t→c*	H480H
	nt1509c→t <sup>†</sup>	N503N
	nt2141c→t <sup>†</sup>	L714P
	nt2259t→c	D753D
	nt4098t→g <sup>*†</sup>	L1366L
	nt4453g→a <sup>†</sup>	3UTR
* PD20 is heterozygous; † VU008 is heterozygous.		

**TABLE 4**Chromosome Breakage Analysis of Whole-cell Fusions.

Cell line/hybrids	DEB (ng/ml)	MMC (ng/ml)	% of Cells with radials	Phenotype
PD20i	300		58	S
PD24p	300		na*	S
VU423p	300		na*	S
PD319i	300		52	S
PD20i/VU423p	300		6	R
PD20i/PD24p	300		30	S
PD20i/PD319i	300		0	R
PD20i		40	48	S
VU423i		40	78	S
PD20i/VU423i		40	10	R
VU423i + chr. 3, clone 1		40	74	S
VU423i + chr. 3, clone 2		40	68	S
VU423i + chr. 3, clone 3		40	88	S
PD20i + empty vector	0	0	2	
		40	24	S
	200		62	S
PD20i + FANCD2 vector	0	0	0	
		40	2	R
	200		10	R

Groups of experiments are separated by line spaces. S, cross-linker sensitive; R, cross-linker resistant; i = immortal fibroblast line; p = primary fibroblasts

\* Cell viability at this concentration was too low to score for radial formation, indicating the exquisite sensitivity of primary fibroblasts to interstrand DNA-crosslinks.

Cell line/ plasmid		FA Group	FA protein complex (1)	MMC sensitivity (2)	IR/Bleomycin sensitivity (3)
Lymphoblasts	PD7	Wt	+	R	R
	HSC72	A	-	S	
	HSC72+A	A	+	R	
	PD4	C	-	S	
	PD4+C	C	+	R	
	EUFA316	G	-	S	
	EUFA316+G	G	+	R	
	EUFA121	F	-	S	S
	EUFA121+F	F	+	R	R
	PD20	D	+	S	S
	PD20(R)	D	+	R	R
Fibroblasts	GM0637	Wt	+	R	R
	GM6914	A	-	S	S
	GM694+A	A	+	R	R
	PD426	C	-	S	
	PD426+C	C	+	R	
	FAG326SV	G	-	S	
	FAG326SV+G	G	+	R	
	PD20F	D	+	S	S
	20-3-15 (+D)	D	+	R	R
	NBS (-/-)	NBS	+	S	S
	ATM (-/-)	ATM	+	S	S
	BRCA1 (-/-)	BRCA1	+	S	S

- 1) The presence of the FA protein complex (FANCA/FANCG/FANCC) was determined as previously described (Garcia-Higuera et al., MCB 19:4866-4873, 1999)
- 2) MMC sensitivity for determined by the XTT assay for lymphoblasts or by the crystal violet assay for fibroblasts.
- 2) IR/Bleomycin sensitivity was determined by analysis of chromosome breakage (See Materials and Methods)



Table 7: PCR Primers to Amplify the 44 Exons of *FANCD*

Exon	Primer Name	SEQ ID NO	Primer Sequence(5'→3')	Product Size (bp)	Annealing Temp
1	MG914	115	F: CTAGCACAGAAGCTCTGCTGC	372	54
	MG837	116	R: CTAGCACAGAAGCTCTGCTGC		
2	MG746	117	F:	422	50
	MG747	118	CTTCAGCAACAGCGAAGTAGTCT G		
			R: GATTCTCAGCACTTGAAAAGCAG G		
3	MG773	119	F: GGACACATCAGTTTTCCTCTC	309	50
	MG789	120	R: GAAAAGCCATGATTCAGTCC		
4-5	MG816	121	F: TCATCAGGCAAGAACTTGG	467	50
	MG803	122	R: GAAGTTGGCAAAACAGACTG		
6	MG804	123	F: GAGCCATCTGCCATTTCTTG	283	50
	MG812	124	R: CCCCTATTTAGACTTGAGGC		
7	MG775	125	F: CAAAGTGTATTCCAGGAGC	343	50
	MG802	126	R: CATCAGGGTACTTTGAACATTC		
8-9	MG727	127	F: TTGACCAGAAAGGCTCAGTCC	640	50
	MG915	128	R: AGATGATGCCAGAGGGTTTATCC		
10	MG790	129	F: TGCCAGCTCTGTTCAAACC	222	50
	MG774	130	R: AGGCAATGACTGACTGACAC		
11	MG805	131	F:	392	50
	MG791	132	TGCCCGTCTATTTTGATGAAGC R: TCTCAGTTAGTCTGGGGACAG		
12	MG751	133	F:	432	50
	MG972	134	TCATGGTAGAGAGACTGGACTGT GC R: ACCCTGGAGCAAATGACAACC		
13-14	MG978	135	F: ATTTGGTCCAGGCTACATGGC	555	50
	MG974	136	R: GAAAGACAATGGGAAGCCATGC		
15	MG975	137	F:	513	50
	MG976	138	GGGAGTGTGTGGAACAAATGAGC R: AGTTTCTACAGGCTGGTCCTATTC C		
16	MG755	139	F: ACCCTGGAAATCCCATTTGATGC	679	58
	MG740	140	R: TTCTGTRGTCCCTCCCTTGC		
17	MG794	141	F: GATGGTCAAGTTACACTGGC	382	50
	MG778	142	R: CACCTCCCACCAATTATAGTATTC		
18	MG808	143	F: CTATGTTGTCTCTTTTACAGGG	234	48
	MG817	144	R: AATCTTTCCACCAATTGGC		
19	MG779	145	F: CATACTTCTTTTGCTGTGC	199	48
	MG795	146	R: CCACAGAAGTCAGAATCTCCACG		
20	MG731	147	F: TGTAACAAACCTGCACGTTG	632	56
	MG732	148	R: TGCTACCCAAGCCAGTAGTTCC		
21	MG788	149	F: GAGTTTGGGAAAGATTGGCAGC	232	50
	MG772	150	R: TGTAGTAAAGCAGCTCTCATGC		
22-23	MG733	151	F: CAAGTACACTCTGCACTGCC	652	50
	MG758	152	R: TGACTCAACTTCCCCACCAAGAG		

24-25	MG736 MG737	153 154	F: R:	732	50
			CTCCCTATGTACGTGGAGTAATAC GGGAGTCTTGTGGAACTAAG		
26	MG780 MG759	155 156	F: R:	284	50
			TTCATAGAGATCTCTCAGCTCTG GTTTTGGTATGAGGGAAAGC		
27-28	MG760 MG781	157 158	F: R:	653	50
			AGCCATGCTTGGAAATTTTGG CTCACTGGGATGTCACAAAC		
29	MG740 MG741	159 160	F: R:	447	50
			GGTCTTGATGTGTGACTTGTATEC C R CCTCACGTGCACAGTGTTCCTTGT G		
30	MG809 MG797	161 162	F: R:	281	48
			CATGAAATGACTAGGACATTCC CTACCCAGTGACCCAAACAC		
31-32	MG761 MG742	163 164	F: R:	503	50
			CGAACCCCTTAGTTCTGAGAGGC TCACTGCCCTGGTGACTGTTC		
33	MG916 MG810	165 166	F: R:	274	50
			TTGATGGTACAGACTGGAGGC AAGAAAGTTGCCAATCCTGTTCC		
34	MG762 MG743	167 168	F: R:	848	50
			AGGACCTGAAATAAGGAGGTT GCCCAAAGTTGTAACTGTGAG		
35-36	MG787 MG806	169 170	F: R:	590	50
			AGCAAGAATGAGGTCAAGTTC GGGAAAACTGGAGGAAAGAACT C		
37	MG818 MG813	171 172	F: R:	633	50
			AGAGGCTAGGGAAGGAAGCTAG CCAAAGTGCACCTTCTTGAAG		
38	MG834 MG836	173 174	F: R:	275	50
			GATGCACTGGTTGCTACATC CCAGGACACTTGGTTTCTGC		
39	MG839 MG871	175 176	F: R:	370	50
			AAGCTGCCAGTTGGAAATCAG CTTGTGGCCAAAGAAATGAG		
40	MG829 MG870	177 178	F: R:	223	50
			TGGGCTGGATGAGACTATTCC CCAAGGACATATCTTCTGAGCAA C		
41	MG820 MG811	179 180	F: R:	278	50
			TGATATCAGCATAGGCTTGC GATCCCGCAATAGGAACCTGC		
42	MG763 MG782	181 182	F: R:	227	50
			CATTCAGATTCACCAGGACAC CCTTACATGCCATCTGATGC		
43	MG754 MG835	183 184	F: R:	232	50
			AAGCTTCGCCCTATTACCC GGAAATGAGAGGCTATTAATGC		
44	MG1006 3'UTR MG1005	185 186	F: R:	234	50
			TGTATTCCAGAGGTCACCCAGAG C R: CCAGTAAGAAAGGCAAACAGCG		

## SEQUENCE LISTING

<110> Dana-Farber Cancer Institute

<120> Methods and Compositions for the Diagnosis of Cancer Susceptibilities and Defective DNA Repair Mechanisms and Treatment Thereof

<130> 07023/2058 (formerly 2486/107 WO)

<140> PCT/US01/45561

<141> 2001-11-02

<150> US 60/245,756

<151> 2000-11-03

<160> 191

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1451

<212> PRT

<213> Homo sapiens

<220>

<221> PEPTIDE

<222> (1)...(1451)

<223> Human fanod2

<400> 1

Met Val Ser Lys Arg Arg Leu Ser Lys Ser Glu Asp Lys Glu Ser Leu  
 1 5 10 15  
 Thr Glu Asp Ala Ser Lys Thr Arg Lys Gln Pro Leu Ser Lys Lys Thr  
 20 25 30  
 Lys Lys Ser His Ile Ala Asn Ala Val Glu Glu Asn Asp Ser Ile Phe  
 35 40 45  
 Val Lys Leu Leu Lys Ile Ser Gly Ile Ile Leu Lys Thr Gly Glu Ser  
 50 55 60  
 Gln Asn Gln Leu Ala Val Asp Gln Ile Ala Phe Gln Lys Lys Leu Phe  
 65 70 75 80  
 Gln Thr Leu Arg Arg His Pro Ser Tyr Pro Lys Ile Ile Glu Glu Phe  
 85 90 95  
 Val Ser Gly Leu Glu Ser Tyr Ile Glu Asp Glu Asp Ser Phe Arg Asn  
 100 105 110  
 Cys Leu Leu Ser Cys Glu Arg Leu Gln Asp Glu Glu Ala Ser Met Gly  
 115 120 125  
 Ala Ser Tyr Ser Lys Ser Leu Ile Lys Leu Leu Leu Gly Ile Asp Ile  
 130 135 140  
 Leu Gln Pro Ala Ile Ile Lys Thr Leu Phe Glu Lys Leu Pro Glu Tyr  
 145 150 155 160  
 Phe Phe Glu Asn Arg Asn Ser Asp Glu Ile Asn Ile Phe Arg Leu Ile  
 165 170 175  
 Val Ser Gln Leu Lys Trp Leu Asp Arg Val Val Asp Gly Lys Asp Leu  
 180 185 190  
 Thr Thr Lys Ile Met Gln Leu Ile Ser Ile Ala Pro Glu Asn Leu Gln  
 195 200 205  
 His Asp Ile Ile Thr Ser Lys Pro Glu Ile Leu Gly Asp Ser Gln His  
 210 215 220  
 Ala Asp Val Gly Lys Glu Leu Ser Asp Leu Leu Ile Glu Asn Thr Ser  
 225 230 235 240  
 Leu Thr Val Pro Ile Leu Asp Val Leu Ser Ser Leu Arg Leu Asp Pro  
 245 250 255  
 Asn Phe Leu Leu Lys Val Arg Gln Leu Val Met Asp Lys Leu Ser Ser  
 260 265 270  
 Ile Arg Leu Glu Asp Leu Pro Val Ile Ile Lys Phe Ile Leu His Ser  
 275 280 285  
 Val Thr Ala Met Asp Thr Leu Glu Val Ile Ser Glu Leu Arg Glu Lys  
 290 295 300  
 Leu Asp Leu Gln His Cys Val Leu Pro Ser Arg Leu Gln Ala Ser Gln  
 305 310 315 320  
 Val Lys Leu Lys Ser Lys Gly Arg Ala Ser Ser Gly Asn Gln Glu  
 325 330 335  
 Ser Ser Gly Gln Ser Cys Ile Ile Leu Leu Phe Asp Val Ile Lys Ser

Ala Ile Arg Tyr Glu Lys Thr Ile Ser Glu Ala Trp Ile Lys Ala Ile  
 340 345 350  
 Glu Asn Thr Ala Ser Val Ser Glu His Lys Val Phe Asp Leu Val Met  
 355 360 365  
 Leu Phe Ile Ile Val Ser Thr Asn Thr Gln Thr Lys Lys Tyr Ile Asp  
 370 375 380  
 385 Arg Val Leu Arg Asn Lys Ile Arg Ser Gly Cys Ile Gln Glu Gln Leu  
 390 395 400  
 Leu Gln Ser Thr Phe Ser Val His Tyr Leu Val Leu Lys Asp Met Cys  
 405 410 415  
 Ser Ser Ile Leu Ser Leu Ala Gln Ser Leu Leu His Ser Leu Asp Gln  
 420 425 430  
 Ser Ile Ile Ser Phe Gly Ser Leu Leu Tyr Lys Tyr Ala Phe Lys Phe  
 435 440 445  
 450 Phe Asp Thr Tyr Cys Gln Gln Glu Val Val Gly Ala Leu Val Thr His  
 455 460 465  
 470 Ile Cys Ser Gly Asn Glu Ala Glu Val Asp Asp Ala Leu Asp Val Leu  
 475 480  
 485 Leu Glu Leu Val Val Leu Asn Pro Ser Ala Met Met Met Asn Ala Val  
 490 495 500  
 Phe Val Gln Gly Ile Leu Asp Tyr Leu Asp Asn Ile Ser Pro Gln Gln  
 505 510 515  
 Ile Arg Lys Leu Phe Tyr Val Leu Ser Thr Leu Ala Phe Ser Lys Gln  
 520 525 530  
 535 Asn Glu Ala Ser Ser His Ile Gln Asp Asp Met His Leu Val Ile Arg  
 540 545 550  
 555 Lys Gln Leu Ser Ser Thr Val Phe Lys Tyr Lys Leu Ile Gly Ile Ile  
 560 565 570  
 Gly Ala Val Thr Met Ala Gly Ile Met Ala Ala Asp Arg Ser Glu Ser  
 575 580 585  
 Pro Ser Leu Thr Gln Glu Arg Ala Asn Leu Ser Asp Glu Gln Cys Thr  
 590 595 600  
 Gln Val Thr Ser Leu Leu Gln Leu Val His Ser Cys Ser Glu Gln Ser  
 605 610 615  
 Pro Gln Ala Ser Ala Leu Tyr Tyr Asp Glu Phe Ala Asn Leu Ile Gln  
 620 625 630  
 635 His Glu Lys Leu Asp Pro Lys Ala Leu Glu Trp Val Gly His Thr Ile  
 640 645 650  
 655 Cys Asn Asp Phe Gln Asp Ala Phe Val Val Asp Ser Cys Val Val Pro  
 660 665 670  
 Glu Gly Asp Phe Pro Phe Pro Val Lys Ala Leu Tyr Gly Leu Glu Glu  
 675 680 685  
 Tyr Asp Thr Gln Asp Gly Ile Ala Ile Asn Leu Leu Pro Leu Leu Phe  
 690 695 700  
 Ser Gln Asp Phe Ala Lys Asp Gly Gly Pro Val Thr Ser Gln Glu Ser  
 705 710 715  
 Gly Gly Lys Leu Val Ser Pro Leu Cys Leu Ala Pro Tyr Phe Arg Leu  
 720 725 730  
 735 Leu Arg Leu Cys Val Glu Arg Gln His Asn Gly Asn Leu Glu Glu Ile  
 740 745 750  
 Asp Gly Leu Leu Asp Cys Pro Ile Phe Leu Thr Asp Leu Glu Pro Gly  
 755 760 765  
 Glu Lys Leu Glu Ser Met Ser Ala Lys Glu Ala Ser Phe Met Cys Ser  
 770 775 780  
 Leu Ile Phe Leu Thr Leu Asn Trp Phe Arg Glu Ile Val Asn Ala Phe  
 785 790 795 800  
 Cys Gln Glu Thr Ser Pro Glu Asn Lys Gly Lys Val Leu Thr Arg Leu  
 805 810 815  
 Lys His Ile Val Glu Leu Gln Ile Leu Leu Glu Lys Tyr Leu Ala Val  
 820 825 830  
 Thr Pro Asp Tyr Val Pro Pro Leu Gly Asn Phe Asp Val Glu Thr Leu  
 835 840 845  
 Asp Ile Thr Pro His Thr Val Thr Ala Ile Ser Ala Lys Ile Arg Lys  
 850 855 860  
 Lys Gly Lys Ile Glu Arg Lys Gln Lys Thr Asp Gly Ser Lys Thr Ser  
 865 870 875 880  
 Ser Ser Asp Thr Leu Ser Glu Glu Lys Asn Ser Glu Cys Asp Pro Thr  
 885 890 895  
 Pro Ser His Arg Gly Gln Leu Asn Lys Glu Phe Thr Gly Lys Glu Glu  
 900 905 910  
 Lys Thr Ser Leu Leu Leu His Asn Ser His Ala Phe Phe Arg Glu Leu  
 915 920 925  
 Asp Ile Glu Val Phe Ser Ile Leu His Cys Gly Leu Val Thr Lys Phe

930 935 940  
 Ile Leu Asp Thr Glu Met His Thr Glu Ala Thr Glu Val Val Gln Leu  
 945 950 955 960  
 Gly Pro Pro Glu Leu Leu Phe Leu Leu Glu Asp Leu Ser Gln Lys Leu  
 965 970 975  
 Glu Ser Met Leu Thr Pro Pro Ile Ala Arg Arg Val Pro Phe Leu Lys  
 980 985 990  
 Asn Lys Gly Ser Arg Asn Ile Gly Phe Ser His Leu Gln Gln Arg Ser  
 995 1000 1005  
 Ala Gln Glu Ile Val His Cys Val Glu Gln Leu Leu Thr Pro Met Cys  
 1010 1015 1020  
 Asn His Leu Glu Asn Ile His Asn Tyr Ile Gln Cys Leu Ala Ala Glu  
 1025 1030 1035 1040  
 Asn His Gly Val Val Asp Gly Pro Gly Val Lys Val Gln Glu Tyr His  
 1045 1050 1055  
 Ile Met Ser Ser Cys Tyr Gln Arg Leu Leu Gln Ile Phe His Gly Leu  
 1060 1065 1070  
 Phe Ala Trp Ser Gly Phe Ser Gln Pro Glu Asn Gln Asn Leu Leu Tyr  
 1075 1080 1085  
 Ser Ala Leu His Val Leu Ser Ser Arg Leu Lys Gln Gly Glu His Ser  
 1090 1095 1100  
 Gln Pro Leu Glu Glu Leu Ser Gln Ser Val His Tyr Leu Gln Asn  
 1105 1110 1115 1120  
 Phe His Gln Ser Ile Pro Ser Phe Gln Cys Ala Leu Tyr Leu Ile Arg  
 1125 1130 1135  
 Leu Leu Met Val Ile Leu Glu Lys Ser Thr Ala Ser Ala Gln Asn Lys  
 1140 1145 1150  
 Glu Lys Ile Ala Ser Leu Ala Arg Gln Phe Leu Cys Arg Val Trp Pro  
 1155 1160 1165  
 Ser Gly Asp Lys Glu Lys Ser Asn Ile Ser Asn Asp Gln Leu His Ala  
 1170 1175 1180  
 Leu Leu Cys Ile Tyr Leu Glu His Thr Glu Ser Ile Leu Lys Ala Ile  
 1185 1190 1195 1200  
 Glu Glu Ile Ala Gln Val Gly Val Pro Glu Leu Ile Asn Ser Pro Lys  
 1205 1210 1215  
 Asp Ala Ser Ser Ser Thr Phe Pro Thr Leu Thr Arg His Thr Pro Val  
 1220 1225 1230  
 Val Phe Phe Arg Val Met Met Ala Glu Leu Glu Lys Ile Val Lys Lys  
 1235 1240 1245  
 Ile Glu Pro Gly Thr Ala Ala Asp Ser Gln Gln Ile His Glu Glu Lys  
 1250 1255 1260  
 Leu Leu Tyr Trp Asn Met Ala Val Arg Asp Phe Ser Ile Leu Ile Asn  
 1265 1270 1275 1280  
 Leu Ile Lys Val Phe Asp Ser His Pro Val Leu His Val Cys Leu Lys  
 1285 1290 1295  
 Val Gly Arg Leu Phe Val Glu Ala Phe Leu Lys Gln Cys Met Pro Leu  
 1300 1305 1310  
 Leu Asp Ile Ser Phe Arg Lys His Arg Glu Asp Val Leu Ser Leu Leu  
 1315 1320 1325  
 Glu Thr Phe Gln Leu Asp Thr Arg Leu Leu His His Leu Cys Gly His  
 1330 1335 1340  
 Ser Lys Ile His Gln Asp Thr Arg Leu Thr Gln His Val Pro Leu Leu  
 1345 1350 1355 1360  
 Lys Lys Thr Leu Glu Leu Leu Val Cys Arg Val Lys Ala Met Leu Thr  
 1365 1370 1375  
 Leu Asn Asn Cys Arg Glu Ala Phe Trp Leu Gly Asn Leu Lys Asn Arg  
 1380 1385 1390  
 Asp Leu Gln Gly Glu Glu Ile Lys Ser Gln Asn Ser Gln Glu Ser Thr  
 1395 1400 1405  
 Ala Asp Glu Ser Glu Asp Asp Met Ser Ser Gln Ala Ser Lys Ser Lys  
 1410 1415 1420  
 Ala Thr Glu Asp Gly Glu Glu Asp Glu Val Ser Ala Gly Glu Lys Glu  
 1425 1430 1435 1440  
 Gln Asp Ser Asp Glu Ser Tyr Asp Asp Ser Asp  
 1445 1450

<210> 2  
 <211> 1269  
 <212> PRT  
 <213> Drosophila melanogaster

<220>  
 <221> PEPTIDE  
 <222> (1)... (1269)

## &lt;223&gt; Drosophila fancd2

&lt;400&gt; 2

```

Met Tyr Lys Gln Phe Lys Lys Arg Ser Lys Lys Pro Leu Asn Thr Ile
 1      5      10      15
Asp Glu Asn Ala Thr Ile Lys Val Pro Arg Leu Ala Glu Thr Thr Thr
 20      25      30
Asn Ile Ser Val Glu Ser Ser Ser Gly Gly Ser Glu Glu Asn Ile Pro
 35      40      45
Ala Ser Gln Glu His Thr Gln Arg Phe Leu Ser Gln His Ser Val Ile
 50      55      60
Leu Ala Ala Thr Leu Gly Ala Thr Gly Glu Ser Ser Arg Asp Ile Ala
 65      70      75
Thr Leu Ser Arg Gln Pro Asn Asn Phe Phe Glu Leu Val Leu Val Arg
 85      90      95
Ala Gly Val Gln Leu Asp Gln Gly Asp Ser Leu Ile Leu Ala Cys Asp
100      105      110
His Val Pro Ile Val Ser Lys Leu Ala Glu Ile Phe Thr Ser Ala Ser
115      120      125
Ser Tyr Thr Asp Lys Met Glu Thr Phe Lys Thr Gly Leu Asn Ala Ala
130      135      140
Met Ala Pro Gly Ser Lys Leu Val Gln Lys Leu Thr Gly Cys Thr
145      150      155
Val Asp Ala Ala Gly Glu Glu Gln Ile Tyr Gln Ser Gln Asn Ser Met
165      170      175
Phe Met Asn Phe Leu Met Ile Asp Phe Met Arg Asp Ala Cys Val Glu
180      185      190
Val Leu Leu Asn Lys Ile Glu Glu Val Ala Lys Ser Asp Arg Val Ile
195      200      205
Met Gly Lys Ala Ala Ile Pro Leu Pro Leu Leu Pro Leu Met Leu Thr
210      215      220
Gln Leu Arg Tyr Leu Thr Ala Ser His Lys Val Glu Ile Tyr Ser Arg
225      230      235
Ile Glu Val Ile Phe Asn Arg Ala Thr Glu Ser Ala Lys Leu Asp Ile
245      250      255
Ile Ala Asn Ala Glu Leu Ile Leu Asp Ala Ser Met His Asp Glu Phe
260      265      270
Val Glu Leu Leu Asn Thr Glu Asp Leu Phe His Met Thr Thr Val Gln
275      280      285
Thr Leu Gly Asn Leu Ser Leu Ser Asp Arg Thr Gln Ala Lys Leu Arg
290      295      300
Val Arg Ile Leu Asp Phe Ala Thr Ser Gly Gln Cys Ser Asp Ala Ile
305      310      315
Leu Pro His Leu Ile Arg Leu Leu Leu Asn Val Leu Lys Ile Asp Thr
325      330      335
Asp Asp Ser Val Arg Asp Leu Arg Arg Arg Ile Lys Leu Glu His
340      345      350
Ile Thr Val Ser Ile Leu Glu Glu Ile Gln His Tyr Arg His Ile Leu
355      360      365
Glu Gln His Ile Thr Thr Leu Met Asn Ile Leu His Asp Phe Met Arg
370      375      380
Glu Lys Asn Arg Ile Val Ser Asp Phe Ala Lys Ser Ser Tyr Ser Ile
385      390      395
Leu Phe Lys Ile Phe Asn Ser Ile Gln Lys Asn Ile Leu Lys Lys Leu
405      410      415
Leu Glu Leu Thr Cys Asp Lys Ser Ser Pro His Leu Thr Thr His Ala
420      425      430
Leu Glu Leu Leu Arg Glu Leu Gln Arg Lys Ser Ala Lys Asp Val Gln
435      440      445
Asn Cys Ala Thr Leu Leu Ile Pro Met Leu Asp Arg Thr Ser Asp Leu
450      455      460
Ser Leu Thr Gln Thr Arg Val Ala Met Asp Leu Leu Cys His Val Ala
465      470      475
Phe Pro Asp Pro Asn Leu Ser Pro Cys Leu Gln Leu Gln Glu Gln Val
485      490      495
Asp Met Val Val Lys Lys Gln Leu Ile Asn Ser Ile Asp Asn Ile Lys
500      505      510
Lys Gln Gly Ile Ile Gly Cys Val Gln Leu Ile Asp Ala Met Ala Arg
515      520      525
Ile Ala Asn Asn Gly Val Asp Arg Asp Phe Phe Ile Ala Ser Val Glu
530      535      540
Asn Val Asp Ser Leu Pro Asp Gly Arg Gly Lys Met Ala Ala Asn Leu
545      550      555
Ile Ile Arg Thr Glu Ala Ser Ile Gly Asn Ser Thr Glu Ser Leu Ala
560

```

565 570 575  
 Leu Phe Phe Glu Glu Leu Ala Thr Val Phe Asn Gln Arg Asn Glu Gly  
 580 585 590  
 Thr Ser Gly Cys Glu Leu Asp Asn Gln Phe Ile Ala Trp Ala Cys Asp  
 595 600 605  
 Leu Val Thr Phe Arg Phe Gln Ala Ser Phe Val Thr Glu Asn Val Pro  
 610 615 620  
 Glu Thr Lys Ala Cys Asp Ser Ile Tyr Val Leu Ala Pro Leu Phe Asn  
 625 630 635 640  
 Tyr Val Arg Val Leu Tyr Lys His Arg His Gln Asp Ser Leu Glu Ser  
 645 650 655  
 Ile Asn Ala Leu Leu Gly Cys Ala Ile Val Leu Pro Ser Phe Phe Glu  
 660 665 670  
 Asp Asp Asn Tyr Val Ser Val Phe Glu Asn Phe Glu Ala Glu Gln Gln  
 675 680 685  
 Lys Asp Ile Leu Ser Ile Tyr Phe His Thr Val Asn Trp Met Arg Val  
 690 695 700  
 Ser Ile Ser Ala Phe Ala Ser Gln Arg Asp Pro Thr Arg Arg Arg  
 705 710 715 720  
 Val Leu Ser Arg Leu Gly Glu Leu Ile Arg Ile Glu Gln Arg Met Lys  
 725 730 735  
 Pro Leu Leu Ala Arg Ala Pro Val Asp Phe Val Ala Pro Pro Tyr Gln  
 740 745 750  
 Phe Leu Thr Asn Val Lys Leu Ser Asn Gln Asn Gln Lys Arg Pro Gly  
 755 760 765  
 Pro Lys Pro Ala Ala Lys Leu Asn Ala Thr Leu Pro Glu Pro Asp Leu  
 770 775 780  
 Thr Gly Asn Gln Pro Ser Ile Ala Asp Phe Thr Ile Lys Val Gly Gln  
 785 790 795 800  
 Cys Lys Thr Val Lys Thr Lys Thr Asp Phe Glu Gln Met Tyr Gly Pro  
 805 810 815  
 Arg Glu Arg Tyr Arg Pro Met Glu Val Glu Ile Ile Met Leu Leu Val  
 820 825 830  
 Glu Gln Lys Phe Val Leu Asn His Gln Leu Glu Glu Glu Gln Met Gly  
 835 840 845  
 Glu Phe Leu Gly Leu Leu Glu Leu Arg Phe Leu Leu Glu Asp Val Val  
 850 855 860  
 Gln Lys Leu Glu Ala Ala Val Leu Arg His His Asp Ser Tyr Asp Ala  
 865 870 875 880  
 Asp Ser Phe Arg Pro His Leu Ala Lys Pro Glu Asp Phe Ile Cys Asp  
 885 890 895  
 Leu Leu Pro Cys Leu His Glu Val Asn Asn His Leu Ile Thr Leu Gly  
 900 905 910  
 Glu Ala Ile Asp Asn Gln Leu Thr Glu Val Ser Ser His Val Tyr Ser  
 915 920 925  
 Asn Leu Asp Leu Phe Lys Asp Gln Phe Cys Tyr Ile Lys Ser Cys Phe  
 930 935 940  
 Gly Leu Cys Val Arg Leu Phe Ala Leu Tyr Phe Ala Trp Ser Glu Trp  
 945 950 955 960  
 Ser Asp Lys Ser Gln Glu Gln Leu Leu His Arg Ile Leu Cys Gly Thr  
 965 970 975  
 Leu Leu Arg Arg Lys Trp Phe His Tyr Ser Gly Thr Leu Asp Lys Gly  
 980 985 990  
 Gly Gln Cys Asn Ile Tyr Leu Asp Glu Leu Val Lys Gly Phe Leu Lys  
 995 1000 1005  
 Lys Ser Asn Ala Lys Ser Gln Thr Glu Leu Leu Thr Glu Leu Val Lys  
 1010 1015 1020  
 Gln Cys Ser Ile Leu Asn Thr Lys Asp Lys Ala Leu Thr Ser Phe Pro  
 1025 1030 1035 1040  
 Asn Phe Lys Lys Ala Asn Phe Pro Leu Leu Phe Arg Gly Leu Cys Glu  
 1045 1050 1055  
 Val Leu Ile His Ser Leu Ser Gly Gln Val Ser Val Asp Ser Arg Gly  
 1060 1065 1070  
 Asp Lys Leu Lys Leu Trp Glu Ser Ala Val Asp Leu Leu Asn Gly Leu  
 1075 1080 1085  
 Leu Ser Ile Val Gln Gln Val Glu Gln Pro Arg Asn Phe Gly Leu Phe  
 1090 1095 1100  
 Leu Lys His Ser Leu Leu Phe Leu Lys Leu Leu Gln His Gly Met  
 1105 1110 1115 1120  
 Ser Ala Leu Glu Ser Ile Val Arg Glu Asp Pro Glu Arg Leu Thr Arg  
 1125 1130 1135  
 Phe Leu His Glu Leu Gln Lys Val Thr Arg Phe Leu His Gln Leu Cys  
 1140 1145 1150  
 Cys His Ser Lys Ser Ile Lys Asn Thr Ala Ile Ile Ser Tyr Ile Pro

1155 1160 1165  
 Ser Leu Arg Glu Thr Ile Glu Thr Leu Val Phe Arg Val Lys Ala Leu  
 1170 1175 1180  
 Leu Ala Ala Asn Asn Cys His Ser Ala Phe His Met Gly Asn Met Ile  
 1185 1190 1195 1200  
 Asn Arg Asp Leu His Gly Asp Ser Ile Ile Thr Pro Arg Ser Ser Phe  
 1205 1210 1215  
 Ala Gly Glu Glu Asn Ser Asp Asp Glu Leu Pro Ala Asp Asp Thr Ser  
 1220 1225 1230  
 Val Asp Glu Thr Val Leu Gly Asp Asp Met Gly Ile Thr Ala Val Ser  
 1235 1240 1245  
 Val Ser Thr Arg Pro Ser Asp Gly Ser Arg Arg Ser Lys Ser Ser Ser  
 1250 1255 1260  
 Arg Ser Lys Cys Phe  
 1265

<210> 3  
 <211> 1286  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> PEPTIDE  
 <222> (1)...(1286)  
 <223> Arabidopsis fancd2

<400> 3  
 Met Val Phe Leu Ser Arg Lys Lys Pro Pro Pro Pro Pro Ser Ser Ser  
 1 5 10 15  
 Ser Ala Ala Pro Ser Leu Lys Ile Pro Gln Pro Gln Lys Glu Ser Val  
 20 25 30  
 Glu Phe Asp Ala Val Glu Lys Met Thr Ala Ile Leu Ala Glu Val Gly  
 35 40 45  
 Cys Thr Leu Met Asn Pro Tyr Gly Pro Pro Cys Leu Pro Ser Asp Leu  
 50 55 60  
 His Ala Phe Arg Arg Asn Leu Thr Gly Arg Leu Ser Ser Phe Ser Ala  
 65 70 75 80  
 Asn Ser Gly Glu Arg Asp Asn Val Gly Ala Leu Cys Ser Val Phe Val  
 85 90 95  
 Ala Gly Phe Ser Leu Tyr Ile Gln Ser Pro Ser Asn Leu Arg Arg Met  
 100 105 110  
 Leu Ser Ser Ser Thr Thr Lys Arg Asp Glu Ser Leu Val Arg Asn  
 115 120 125  
 Leu Leu Leu Val Ser Pro Ile Gln Leu Asp Ile Gln Glu Met Leu Leu  
 130 135 140  
 Glu Lys Leu Pro Glu Tyr Phe Asp Val Val Thr Gly Cys Ser Leu Glu  
 145 150 155 160  
 Glu Asp Val Ala Arg Leu Ile Ile Asn His Phe Arg Thr Leu Asp Phe  
 165 170 175  
 Ile Val Asn Pro His Val Phe Thr Asp Lys Leu Met Gln Val Leu Ser  
 180 185 190  
 Ile Cys Pro Leu Glu Leu Lys Lys Glu Ile Ile Gly Ser Leu Pro Glu  
 195 200 205  
 Ile Ile Gly Asp His Asn Cys Gln Ala Val Val Asp Ser Leu Glu Lys  
 210 215 220  
 Met Leu Gln Glu Asp Ser Ala Val Val Val Ala Val Leu Asp Ser Phe  
 225 230 235 240  
 Ser Asn Leu Asn Leu Asp Asp Gln Leu Gln Glu Gln Ala Ile Thr Val  
 245 250 255  
 Ala Ile Ser Cys Ile Arg Thr Ile Asp Gly Glu His Met Pro Tyr Leu  
 260 265 270  
 Leu Arg Phe Leu Leu Leu Ala Ala Thr Pro Val Asn Val Arg Arg Ile  
 275 280 285  
 Ile Ser Gln Ile Arg Glu Gln Leu Lys Phe Thr Gly Met Ser Gln Pro  
 290 295 300  
 Cys Ala Ser Gln Asn Lys Leu Lys Gly Lys Val Pro Ala Tyr Asn Ala  
 305 310 315 320  
 Glu Gly Ser Ile Leu His Ala Leu Arg Ser Ser Leu Arg Phe Lys Asn  
 325 330 335  
 Ile Leu Cys Gln Glu Ile Ile Lys Glu Leu Asn Ser Leu Glu Lys Pro  
 340 345 350  
 Arg Asp Phe Lys Val Ile Asp Val Trp Leu Leu Ile Asp Met Tyr Met  
 355 360 365  
 Asn Gly Asp Pro Val Arg Lys Ser Ile Glu Lys Ile Phe Lys Lys Lys

370 Val Val Asp Glu Cys Ile 375 Gln Glu Ala Leu Leu 380 Gln Cys Ile Gly  
 385 Gly Asn Lys Glu Phe Val Lys Ile Leu Gly Ala Leu Val Thr His Val 400  
 Gly Ser Asp Asn Lys Phe Glu Val Ser Ser Val Leu Glu Met Met Thr 415  
 Ala Leu Val Lys Lys Tyr Ala Gln Gln Leu Leu Pro Phe Ser Ser His 430  
 Ile Asn Gly Ile Ser Gly Thr Cys Ile Leu Asp Tyr Leu Glu Gly Phe 445  
 Thr Ile Asp Asn Leu His Lys Thr Tyr Ser Gln Val Tyr Glu Val Phe 460  
 Ser Leu Leu Ala Leu Ser Ala Arg Ala Ser Gly Asp Ser Phe Arg Ser 475  
 Ser Thr Ser Asn Glu Leu Met Met Ile Val Arg Lys Gln Leu Thr Pro 490  
 Ser Cys Leu Val Leu Tyr Trp Gln Val Ser His Pro Asp Leu Lys Tyr 505  
 Lys Lys Met Gly Leu Val Gly Ser Leu Arg Ile Val Ser Ser Leu Gly 520  
 Asp Ala Lys Ser Val Pro Asp Phe Ser Ser Ser Gln Val Glu Arg Leu 535  
 Thr Asn Asp Gly Ser Leu Ala Gly Val Asp Ala Leu Leu Gly Cys Pro 550  
 Leu His Leu Pro Ser Ser Lys Leu Val Gly Ser Leu Trp Gly Arg Ser 565  
 Arg Lys Lys Ser Ser Pro Ser Arg Tyr Ile Met Leu Gln Thr Gly Tyr 580  
 Glu Asn Ser Leu Val Thr Leu Pro Cys Ile Phe Cys Asp Leu Leu Asn 595  
 Ala Phe Ser Ser Gln Ile Asp Glu Lys Ile Gly Cys Ile Ser Gln Ala 610  
 Thr Val Lys Asp Val Thr Thr Lys Leu Leu Lys Arg Leu Arg Asn Leu 625  
 Val Phe Leu Glu Ser Leu Leu Ser Asn Leu Ile Thr Leu Ser Pro Gln 640  
 Ser Leu Pro Glu Leu His Pro Tyr Ser Glu Ser His Val Glu His Pro 655  
 Arg Lys Lys Asn Glu Lys Arg Lys Leu Asp Asp Asp Ala Ser Gln Arg 670  
 Lys Val Ser Met Lys Asn Asn Leu Lys Lys Ser Lys His Ser Asp Val 685  
 Asn Glu Lys Leu Arg Gln Pro Thr Ile Met Asp Ala Phe Lys Lys Ala 700  
 Gly Ala Val Met Ser His Ser Gln Thr Gln Leu Arg Gly Thr Pro Ser 715  
 Leu Pro Ser Met Asp Gly Ser Thr Ala Ala Gly Ser Met Asp Glu Asn 730  
 Cys Ser Asp Asn Glu Ser Leu Ile Val Lys Ile Pro Gln Val Ser Ser 745  
 Ala Leu Glu Ala Gln Pro Phe Lys Phe Arg Pro Leu Leu Pro Gln Cys 760  
 Leu Ser Ile Leu Asn Phe Pro Lys Val Leu Ser Gln Asp Met Gly Ser 775  
 Pro Glu Tyr Arg Ala Glu Leu Pro Leu Tyr Leu Tyr Leu Leu His Asp 790  
 Leu His Thr Lys Leu Asp Cys Leu Val Pro Pro Gly Lys Gln His Pro 805  
 Phe Lys Arg Gly Ser Ala Pro Gly Tyr Phe Gly Arg Phe Lys Leu Val 820  
 Glu Leu Leu Asn Gln Ile Lys Arg Leu Phe Pro Ser Leu Asn Ile Lys 835  
 Leu Asn Ile Ala Ile Ser Leu Leu Ile Arg Gly Asp Glu Thr Ser Gln 850  
 Thr Thr Trp Arg Asp Glu Phe Ala Leu Ser Gly Asn Pro Asn Thr Ser 865  
 Ser Ile Val Val Ser Glu Ser Leu Val Tyr Thr Met Val Cys Lys Glu 880  
 Val Leu Tyr Cys Phe Ser Lys Ile Leu Thr Leu Pro Glu Phe Glu Thr 895  
 Asp Lys Ser Leu Leu Leu Asn Leu Leu Glu Ala Phe Gln Pro Thr Glu 910  
 Ile Pro Val Ala Asn Phe Pro Asp Phe Gln Pro Phe Pro Ser Pro Gly 925

965 970 975  
 Thr Lys Glu Tyr Leu Tyr Ile Gly Val Ser Tyr Phe Phe Glu Asp Ile  
 980 985 990  
 Leu Asn Lys Gly Asn Tyr Phe Cys Ser Phe Thr Asp Asp Phe Pro Tyr  
 995 1000 1005  
 Pro Cys Ser Phe Ser Phe Asp Leu Ala Phe Glu Cys Leu Leu Thr Leu  
 1010 1015 1020  
 Gln Leu Val Val Thr Ser Val Gln Lys Tyr Leu Gly Lys Val Ser Glu  
 1025 1030 1035 1040  
 Glu Ala Asn Arg Lys Arg Asn Pro Gly His Phe His Gly Leu Val Pro  
 1045 1050 1055  
 Asn Leu His Ala Lys Leu Gly Thr Ser Ala Glu Lys Leu Leu Arg His  
 1060 1065 1070  
 Lys Trp Val Asp Glu Ser Thr Asp Asn Lys Gly Leu Lys Asn Lys Val  
 1075 1080 1085  
 Cys Pro Phe Val Ser Asn Leu Arg Ile Val Gln Phe Thr Gly Glu Met  
 1090 1095 1100  
 Val Gln Thr Ile Leu Arg Ile Tyr Leu Glu Ala Ser Gly Ser Thr Ser  
 1105 1110 1115 1120  
 Asp Leu Leu Asp Glu Leu Ala Cys Thr Ile Leu Pro Gln Ala Ser Leu  
 1125 1130 1135  
 Ser Lys Ser Thr Gly Glu Asp Asp Asp Ala Arg Asp His Glu Phe Pro  
 1140 1145 1150  
 Thr Leu Cys Ala Ala Thr Phe Arg Gly Trp Tyr Lys Thr Leu Leu Glu  
 1155 1160 1165  
 Glu Asn Leu Ala Ile Leu Asn Lys Leu Val Lys Thr Val Ser Ser Glu  
 1170 1175 1180  
 Lys Arg Gly Asn Cys Gln Pro Lys Thr Thr Glu Ala His Leu Lys Asn  
 1185 1190 1195 1200  
 Ile Gln Lys Thr Val Asn Val Val Val Ser Leu Val Asn Leu Cys Arg  
 1205 1210 1215  
 Ser His Glu Lys Val Thr Ile His Gly Met Ala Ile Lys Tyr Gly Gly  
 1220 1225 1230  
 Lys Tyr Val Asp Ser Phe Leu Lys Gly Ser Leu Lys His Lys Asp Leu  
 1235 1240 1245  
 Arg Gly Gln Ile Val Ser Ser Gln Ala Tyr Ile Asp Asn Glu Ala Asp  
 1250 1255 1260  
 Glu Val Glu Glu Thr Met Ser Gly Glu Glu Glu Pro Met Gln Glu Asp  
 1265 1270 1275 1280  
 Glu Leu Pro Leu Thr Pro  
 1285

<210> 4  
 <211> 1471  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> PEPTIDE  
 <222> (1)... (1471)  
 <223> Human fancd2

<400> 4  
 Met Val Ser Lys Arg Arg Leu Ser Lys Ser Glu Asp Lys Glu Ser Leu  
 1 5 10 15  
 Thr Glu Asp Ala Ser Lys Thr Arg Lys Gln Pro Leu Ser Lys Lys Thr  
 20 25 30  
 Lys Lys Ser His Ile Ala Asn Ala Val Glu Glu Asn Asp Ser Ile Phe  
 35 40 45  
 Val Lys Leu Leu Lys Ile Ser Gly Ile Ile Leu Lys Thr Gly Glu Ser  
 50 55 60  
 Gln Asn Gln Leu Ala Val Asp Gln Ile Ala Phe Gln Lys Lys Leu Phe  
 65 70 75 80  
 Gln Thr Leu Arg Arg His Pro Ser Tyr Pro Lys Ile Ile Glu Glu Phe  
 85 90 95  
 Val Ser Gly Leu Glu Ser Tyr Ile Glu Asp Glu Asp Ser Phe Arg Asn  
 100 105 110  
 Cys Leu Leu Ser Cys Glu Arg Leu Gln Asp Glu Glu Ala Ser Met Gly  
 115 120 125  
 Ala Ser Tyr Ser Lys Ser Leu Ile Lys Leu Leu Leu Gly Ile Asp Ile  
 130 135 140  
 Leu Gln Pro Ala Ile Ile Lys Thr Leu Phe Glu Lys Leu Pro Glu Tyr  
 145 150 155 160  
 Phe Phe Glu Asn Arg Asn Ser Asp Glu Ile Asn Ile Phe Arg Leu Ile

Val Ser Gln Leu Lys Trp Leu Asp Arg Val Val Asp Gly Lys Asp Leu  
 165 170 175  
 Thr Thr Lys Ile Met Gln Leu Ile Ser Ile Ala Pro Glu Asn Leu Gln  
 180 185 190  
 His Asp Ile Ile Thr Ser Lys Pro Glu Ile Leu Gly Asp Ser Gln His  
 195 200 205  
 Ala Asp Val Gly Lys Glu Leu Ser Asp Leu Leu Ile Glu Asn Thr Ser  
 210 215 220 225  
 Leu Thr Val Pro Ile Leu Asp Val Leu Ser Ser Leu Arg Leu Asp Pro  
 225 230 235 240  
 Asn Phe Leu Leu Lys Val Arg Gln Leu Val Met Asp Lys Leu Ser Ser  
 245 250 255 260  
 Ile Arg Leu Glu Asp Leu Pro Val Ile Ile Lys Phe Ile Leu His Ser  
 265 270 275 280  
 Val Thr Ala Met Asp Thr Leu Glu Val Ile Ser Glu Leu Arg Glu Lys  
 285 290 295 300  
 Leu Asp Leu Gln His Cys Val Leu Pro Ser Arg Leu Gln Ala Ser Gln  
 305 310 315 320  
 Val Lys Leu Lys Ser Lys Gly Arg Ala Ser Ser Ser Gly Asn Gln Glu  
 325 330 335 340  
 Ser Ser Gly Gln Ser Cys Ile Ile Leu Leu Phe Asp Val Ile Lys Ser  
 345 350 355 360  
 Ala Ile Arg Tyr Glu Lys Thr Ile Ser Glu Ala Trp Ile Lys Ala Ile  
 365 370 375 380  
 Glu Asn Thr Ala Ser Val Ser Glu His Lys Val Phe Asp Leu Val Met  
 385 390 395 400  
 Arg Val Leu Arg Asn Lys Ile Arg Ser Gly Cys Ile Gln Glu Gln Leu  
 405 410 415 420  
 Leu Gln Ser Thr Phe Ser Val His Tyr Leu Val Leu Lys Asp Met Cys  
 425 430 435 440  
 Ser Ser Ile Leu Ser Leu Ala Gln Ser Leu Leu His Ser Leu Asp Gln  
 445 450 455 460  
 Ser Ile Ile Ser Phe Gly Ser Leu Leu Tyr Lys Tyr Ala Phe Lys Phe  
 465 470 475 480  
 Phe Asp Thr Tyr Cys Gln Glu Val Val Gly Ala Leu Val Thr His  
 485 490 495 500  
 Ile Cys Ser Gly Asn Glu Ala Glu Val Asp Asp Ala Leu Asp Val Leu  
 505 510 515 520  
 Leu Glu Leu Val Val Leu Asn Pro Ser Ala Met Met Met Asn Ala Val  
 525 530 535 540  
 Phe Val Gln Gly Ile Leu Asp Tyr Leu Asp Asn Ile Ser Pro Gln Gln  
 545 550 555 560  
 Ile Arg Lys Leu Phe Tyr Val Leu Ser Thr Leu Ala Phe Ser Lys Gln  
 565 570 575 580  
 Asn Glu Ala Ser Ser His Ile Gln Asp Asp Met His Leu Val Ile Arg  
 585 590 595 600  
 Lys Gln Leu Ser Ser Thr Val Phe Lys Tyr Lys Leu Ile Gly Ile Ile  
 605 610 615 620  
 Gly Ala Val Thr Met Ala Gly Ile Met Ala Ala Asp Arg Ser Glu Ser  
 625 630 635 640  
 Pro Ser Leu Thr Gln Glu Arg Ala Asn Leu Ser Asp Glu Gln Cys Thr  
 645 650 655 660  
 Gln Val Thr Ser Leu Leu Gln Leu Val His Ser Cys Ser Glu Gln Ser  
 665 670 675 680  
 Pro Gln Ala Ser Ala Leu Tyr Tyr Asp Glu Phe Ala Asn Leu Ile Gln  
 685 690 695 700  
 His Glu Lys Leu Asp Pro Lys Ala Leu Glu Trp Val Gly His Thr Ile  
 705 710 715 720  
 Cys Asn Asp Phe Gln Asp Ala Phe Val Val Asp Ser Cys Val Val Pro  
 725 730 735 740  
 Glu Gly Asp Phe Pro Phe Pro Val Lys Ala Leu Tyr Gly Leu Glu Glu  
 745 750 755 760  
 Tyr Asp Thr Gln Asp Gly Ile Ala Ile Asn Leu Leu Pro Leu Leu Phe  
 765 770 775 780  
 Ser Gln Asp Phe Ala Lys Asp Gly Gly Pro Val Thr Ser Gln Glu Ser  
 785 790 795 800  
 Gly Gly Lys Leu Val Ser Pro Leu Cys Leu Ala Pro Tyr Phe Arg Leu  
 805 810 815 820  
 Leu Arg Leu Cys Val Glu Arg Gln His Asn Gly Asn Leu Glu Glu Ile  
 825 830 835 840  
 Asp Gly Leu Leu Asp Cys Pro Ile Phe Leu Thr Asp Leu Glu Pro Gly  
 845 850 855 860

755 760 765  
 Glu Lys Leu Glu Ser Met Ser Ala Lys Glu Ala Ser Phe Met Cys Ser  
 770 775 780  
 Leu Ile Phe Leu Thr Leu Asn Trp Phe Arg Glu Ile Val Asn Ala Phe  
 785 790 795 800  
 Cys Gln Glu Thr Ser Pro Glu Asn Lys Gly Lys Val Leu Thr Arg Leu  
 805 810 815  
 Lys His Ile Val Glu Leu Gln Ile Leu Leu Glu Lys Tyr Leu Ala Val  
 820 825 830  
 Thr Pro Asp Tyr Val Pro Pro Leu Gly Asn Phe Asp Val Glu Thr Leu  
 835 840 845  
 Asp Ile Thr Pro His Thr Val Thr Ala Ile Ser Ala Lys Ile Arg Lys  
 850 855 860  
 Lys Gly Lys Ile Glu Arg Lys Gln Lys Thr Asp Gly Ser Lys Thr Ser  
 865 870 875 880  
 Ser Ser Asp Thr Leu Ser Glu Glu Lys Asn Ser Glu Cys Asp Pro Thr  
 885 890 895  
 Pro Ser His Arg Gly Gln Leu Asn Lys Glu Phe Thr Gly Lys Glu Glu  
 900 905 910  
 Lys Thr Ser Leu Leu Leu His Asn Ser His Ala Phe Phe Arg Glu Leu  
 915 920 925  
 Asp Ile Glu Val Phe Ser Ile Leu His Cys Gly Leu Val Thr Lys Phe  
 930 935 940  
 Ile Leu Asp Thr Glu Met His Thr Glu Ala Thr Glu Val Val Gln Leu  
 945 950 955 960  
 Gly Pro Pro Glu Leu Phe Leu Leu Glu Asp Leu Ser Gln Lys Leu  
 965 970 975  
 Glu Ser Met Leu Thr Pro Pro Ile Ala Arg Arg Val Pro Phe Leu Lys  
 980 985 990  
 Asn Lys Gly Ser Arg Asn Ile Gly Phe Ser His Leu Gln Gln Arg Ser  
 995 1000 1005  
 Ala Gln Glu Ile Val His Cys Val Glu Gln Leu Leu Thr Pro Met Cys  
 1010 1015 1020  
 Asn His Leu Glu Asn Ile His Asn Tyr Ile Gln Cys Leu Ala Ala Glu  
 1025 1030 1035 1040  
 Asn His Gly Val Val Asp Gly Pro Gly Val Lys Val Gln Glu Tyr His  
 1045 1050 1055  
 Ile Met Ser Ser Cys Tyr Gln Arg Leu Leu Gln Ile Phe His Gly Leu  
 1060 1065 1070  
 Phe Ala Trp Ser Gly Phe Ser Gln Pro Glu Asn Gln Asn Leu Leu Tyr  
 1075 1080 1085  
 Ser Ala Leu His Val Leu Ser Ser Arg Leu Lys Gln Gly Glu His Ser  
 1090 1095 1100  
 Gln Pro Leu Glu Glu Leu Leu Ser Gln Ser Val His Tyr Leu Gln Asn  
 1105 1110 1115 1120  
 Phe His Gln Ser Ile Pro Ser Phe Gln Cys Ala Leu Tyr Leu Ile Arg  
 1125 1130 1135  
 Leu Leu Met Val Ile Leu Glu Lys Ser Thr Ala Ser Ala Gln Asn Lys  
 1140 1145 1150  
 Glu Lys Ile Ala Ser Leu Ala Arg Gln Phe Leu Cys Arg Val Trp Pro  
 1155 1160 1165  
 Ser Gly Asp Lys Glu Lys Ser Asn Ile Ser Asn Asp Gln Leu His Ala  
 1170 1175 1180  
 Leu Leu Cys Ile Tyr Leu Glu His Thr Glu Ser Ile Leu Lys Ala Ile  
 1185 1190 1195 1200  
 Glu Glu Ile Ala Gln Val Gly Val Pro Glu Leu Ile Asn Ser Pro Lys  
 1205 1210 1215  
 Asp Ala Ser Ser Ser Thr Phe Pro Thr Leu Thr Arg His Thr Pro Val  
 1220 1225 1230  
 Val Phe Phe Arg Val Met Met Ala Glu Leu Glu Lys Ile Val Lys Lys  
 1235 1240 1245  
 Ile Glu Pro Gly Thr Ala Ala Asp Ser Gln Gln Ile His Glu Glu Lys  
 1250 1255 1260  
 Leu Leu Tyr Trp Asn Met Ala Val Arg Asp Phe Ser Ile Leu Ile Asn  
 1265 1270 1275 1280  
 Leu Ile Lys Val Phe Asp Ser His Pro Val Leu His Val Cys Leu Lys  
 1285 1290 1295  
 Val Gly Arg Leu Phe Val Glu Ala Phe Leu Lys Gln Cys Met Pro Leu  
 1300 1305 1310  
 Leu Asp Ile Ser Phe Arg Lys His Arg Glu Asp Val Leu Ser Leu Leu  
 1315 1320 1325  
 Glu Thr Phe Gln Leu Asp Thr Arg Leu Leu His His Leu Cys Gly His  
 1330 1335 1340  
 Ser Lys Ile His Gln Asp Thr Arg Leu Thr Gln His Val Pro Leu Leu



ottgggcccc ctgagctgct tttcttgcgt gaagatctct cccagaagct ggagagtatg 3000  
 ctgacacctc ctattgccag gagagtcccc tttctcaaga acaaagggaag ccggaatatt 3060  
 ggattctcac atotccaaca gagatctgcc caagaaattg ttcattgtgt ttttcaactg 3120  
 ctgaccccaa tgtgtaacca cctggagaac attcacaact attttcagtg tttagctgct 3180  
 gagaatcacg gtgtagttga tggaccagga gtgaaagttc aggagtacca cataatgtct 3240  
 tctctgatc agaggctgct gcagattttt catgggcttt ttgcttggag tggattttct 3300  
 caacctgaaa atcagaattt actgtattca gccctccatg tccttagtag ccgactgaaa 3360  
 cagggagaac acagccagcc tttggaggaa ctactcagcc agagcgtcca ttacttgacg 3420  
 aatttccatc aaagcattcc cagtttccag tgtgctottt atctcatcag acttttgatg 3480  
 gttattttgg agaaatcaact agcttctgct cagaacaaag aaaaaattgc ttccottgoc 3540  
 agacaattcc totgtcgggt gtggccaagt ggggataaag agaagagcaa catctctaat 3600  
 gaccagctcc atgctctgct ctgtatctac ctggagcaca cagagagcat tctgaaggcc 3660  
 atagaggaga ttgctgtgtg tgggtgccca gaactgatca actotcctaa agatgcattc 3720  
 tctccacat tccctcaact gaocaggcat acttttgttg ttttctccg tgtgatgatg 3780  
 gctgaactag agaagacggt gaaaaaaatt gagcctggca cagcagcaga ctgcagcag 3840  
 attcatgaag agaaactctc ctactggaac atggctgttc gagacttcag tatctctatc 3900  
 aacttgataa aggtatttga tagtcatctt gttctgatg tatgtttgaa gtatgggctg 3960  
 ctctttgtgg aagcaatgt atgccgctcc tagacttcag ttttagaaaa 4020  
 caccgggaag atgttctgag cttactggaa accttccagt tggacacaag gctgcttcat 4080  
 cacctgtgtg ggcatlccaa gattcaccag gacacgagac tcaccaaca tgtgctctg 4140  
 ctcaaaaaga ccttggaaact tttagtttgc agagtcaaag ctatgctcac tctcaacaa 4200  
 tgtagagagg cttttgctg gggcaatcta aaaaaccggg acttgcaggg tgaagagatt 4260  
 aagtcccaaa attccaggga gagcacagca gatgagagtg aggatgacat gtcattccag 4320  
 gcctccaaga gcaaaaccac tgaggatctc ctacaaaacc caccagagtc tggcactgat 4380  
 ggttgcaatt tgttaattgt totaagttgg tggagcagaa ctttgcctac ttatgtttat 4440  
 tgcataatgc tctatgccc atttccatc cctccataac agcttctgtg cttatataat 4500  
 ttttgggacc cagaagaaac aacgacacaa tcttagaate actcctgagt atctcagatt 4560  
 gtggcatttg ttatagagtt gacaattttc tgcattatag cctctcattt tccatgaatt 4620  
 catatctgaa accattttag aagggagaag tcatcgaaat attttctgag tgttgagaag 4680  
 aatgagttaa accatttaaa cacattttaa acatacaaaa atagaaatgt gaaagcattt 4740  
 ggtgaaagcc aaagcagca gtcagaagct gccaccttag agaactgaaa taaaaataga 4800  
 agttcttacc cttttttgtg gtacagatgc tttcgacaat ttaaagaaag ctaaataaaa 4860  
 atgtagacat ggctggcgca gtggctcatg ctgttaatcc tagcactttt tgaggccaag 4920  
 gtagggagat gctttagctc cgggagctca aggcaaaagc gcacaacata acaagacct 4980  
 atctccacaa aaaaatgaa aaataaacct gggtgcggtg gctcacacct gtaatcccag 5040  
 cactttggga ggccgatgtg ggcagatcac aaggtcagga ggtcaagacc agcctggcca 5100  
 acatagtgaa accccatctc tactgaaat acaaaaatta gctgggtgtg gtggcacgtg 5160  
 cctgttatct cagctacttg ggaagctga 5189

&lt;210&gt; 6

&lt;211&gt; 5194

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

tagaatcgaa aactacgggc ggcgacggct tctcggaagt aatttaagt cacaagacat 60  
 tggtaaaaat ggtttccaaa agaagactgt caaaatctga ggataaagag agcctgacag 120  
 aagatgcctc caaaaccagg aagcaaccac ttccaaaaa gacaaagaaa tctcatattg 180  
 ctaatgaagt tgaagaaaaat gacagcatct ttgtaagct tcttaagata tcaggaatta 240  
 ttcttaaaac gggagagagt cagaatcaac tagctgtgga tcaaatagct ttccaaaaga 300  
 agctctttca gacctgagg agacaccctt cctatcccaa aataatagaa gaatttgtta 360  
 gtggcctgga gctttacatt gaggatgaag acagtttcag gaactgcctt ttgtctgtg 420  
 agcgtctgca ggtgaggaa gccagtatgg gtgcattcta ttotaagagt ctcattcaaac 480  
 tgcctctggg gattgacata ctgcagcctg ccattatcaa aacctattt gagaagtgc 540  
 cagaatattt ttttgaaac aagaacagtg atgaaatcaa catacctcga ctcatgtca 600  
 gtcaactaaa atggcttgac agagtgtgg atggcaagga cctcaccacc aagatcatgc 660  
 agctgatcag tattgtcca gaaacctgc agcatgacat catcaccagc ctacctgaga 720  
 tcttagggga ttcccagcac gotgatgtgg gaaaagaact cagtgaacta ctgatagaga 780  
 atacttcaact cactgtccca atcctggatg tctttcaag cctccgactt gacccaaact 840  
 tctattgaa ggttccgag ttggtgatgg ataagtgtc gtctattaga ttggaggatt 900  
 tacotgtgat aataaagttc attcttcat cogtaacagc catggataca cttgaggtaa 960  
 tttctgagct tggggagaag ttggatctgc agcattgtgt ttgccaatca cgtttacagg 1020  
 cttcccaagt aaagttgaaa agtaaggac gagcaagttc ctccaggaaat caagaaagca 1080  
 gcggtcagag ctgtattatt ctctctttg atgtaataaa gtcagctatt agatagaga 1140  
 aaaccatttc agaagcctgg attaaggcaa ttgaaaacac tgcctcagta tctgaacaca 1200  
 aggtgtttga cctggtgatg cttttcatca totatagcac caatactcag acaagaagt 1260  
 acattgacag ggtgctaaga aataagattc gatcaggctg cattcaagaa cagctgtccc 1320  
 agagtacatt ctctgttcat tacttagttc ttaaggatat gtgttcatcc attctgtcgc 1380  
 tggctcagag tttgcttcaac tctctagacc agagtataat ttcatttggc agtctctat 1440  
 acaaatatgc atttaagttt ttgacacgt actgccagca ggaagtgggt ggtgccttag 1500  
 tgaccatata ctgcagttgg aatgaagctg aagttgatac tgccttagat gtctcttag 1560  
 agttgtagt gttaaaccca tctctatga tgatgaatgc tgtctttgta aagggcattt 1620  
 tagattatct ggataacata tccctcagc aaatacgaaa actcttctat gttctcagca 1680  
 cactggcatt tgcataacag aatgaagcca gcagccacat ccaggatgac atgcacttgg 1740  
 tgataagaaa gcagctctct agcaacctat tcaagtacaa gctcatttgg attattggtg 1800

ctgtgacat ggctggcgc atggcggcag acagaagtga atcacctagt ttgacccaag 1860  
 agagagccaa cctgagcgat gacagtgca cacaggtgac ctcttgttg cagtgggttc 1920  
 attcctgcag tgagcagctc cctcaggcct ctgcacttta ctatgatgaa tttgccaacc 1980  
 tgatccaaca tgaagaagct gatccaaaag ccttggaaat ggttgggcat accatctgta 2040  
 atgatttcca ggatgocctc gtagtggact cctgtgtgtg tccggaaggt gactttccat 2100  
 ttctgtgtaa agcactgtac ggaactggaag aatcacgacac tcaggatggg attgocataa 2160  
 acctcctgcc gctgtgtgtt tctcaggact ttgcaaaaga tgggggtccg gtgacctcac 2220  
 agaatcagg ccaaaaattg gtgtctccgc tggcctggc tccgtatttc cggttactga 2280  
 gactttgtgt ggagagacag cataacggaa acttggagga gattgatggt ctactagatt 2340  
 gtcctatatt cctaactgtac ctggagcctg gagagaagtt ggagtcctat tctgctaaag 2400  
 agcgttcatt catgtgttct ctcattatct ttactctcaa ctggttccga gagattgtaa 2460  
 atgccttctg ccaggaaaca tcacctgaga tgaagggaag ggtgctcact cggttaaagc 2520  
 acattgtaga attgcaaaata atcctggaaa agtacttggc agtcacccca gactatgtcc 2580  
 ctctcttttg aaactttgat gtgaaaactt tagataaac acctcactat gttactgcta 2640  
 tttcagcaaa aatcagaagag aaaggaaaaa tagaaaaggaa acaaaaaaca gatggcagca 2700  
 agacatcctc ccttgacaca ctttcagaag agaaaaattc agaattgtac cctacgccaat 2760  
 ctcattagagg ccagctaaac aaggagtcca cagggaagga agaaaagaca tcattgttac 2820  
 tacataatct ccattgcttt ttccgagagc tggacatiga ggtcttctct attctacatt 2880  
 gtggacttgg gacgaagtgt atccttagata ctgaaatgca cactgaagct acagaagttg 2940  
 tgcacttgg gccccctgag ctgcttttct tgcctggaaga tctctcccag aagctggaga 3000  
 gtatgctgac acctcctatt gocaggagag tcccccttct caagaacaaa ggaagccgga 3060  
 atattggatt ctcacatctc caacagagat ctgcccaga aattgttcat tgtgttttc 3120  
 aactgctgac cccaacttgg aaacacattc agaacttca caactatitt cagtgtttag 3180  
 ctgctgagaa tcacgggtga gttgatggac caggagtga agttcaggag taccacataa 3240  
 tgtcttctct ctatcagagg ctgctgcaga ttttctatgg gctttttgct tggagtggat 3300  
 tttctcaacc tgaaaatcag aatttactgt attcagccct ccattgtcct agtagccgac 3360  
 tgaacacaggg agaacacagc cagcctttgg aggaactact cagccagagc gtcattact 3420  
 tgcagaattt ccatcaaaagc attcccagtt tccagtgtgc tctttatctc atcagacttt 3480  
 tgatggttat tttggagaaa tcaacagctt ctgctcagaa caaagaaaaa attgcttccc 3540  
 ttgcccagaca attcctctgt cgggtgtggc caagtgggga taaagagaag agcaacatct 3600  
 ctaatgacca gctcactgtc ctgctctgta tctacctgga gcacacagag agcattctga 3660  
 aggcocataga ggagattgct ggtgttgggt tcccagaact gatcaactct cctaagatg 3720  
 catcttctct cacattccct acactgacca ggcatacttt tgtgttttc ttcogtga 3780  
 tgatggctga actagagaa acggtgaaaa aaattgagcc tggcacagca gcagactcgc 3840  
 agcagattca tgaagagaaa ctcctctact ggaacatggc tgttcgagac ttcagtatcc 3900  
 tcatcaactt gataaaggta tttgatagtc atcctgttct gcatgtatgt ttgaagtatg 3960  
 ggctctcttt tgtggaagca tttctgaagc aatgtatgcc gctcctagac ttcagtttta 4020  
 gaaaaacacg ggaagatgtt ctgagcttac tggaaaacct ccagttggac acaaggctgc 4080  
 ttcacacact gigtgggcat tocaagattc accaggacac gagactcacc caacatgtgc 4140  
 ctctgctcaa aaagacctg gacttttag tttgcagagt caaagctatg ctcactctca 4200  
 acaattgtag agaggcttct tggctgggca atctaaaaaa cgggacttg cagggtgaag 4260  
 agattaagtc ccaaaattcc caggagagca cagcagatga gagtggagat gacatgtcat 4320  
 cccaggcctc caaggacaaa gccactgagg tatctctaca aaaccacca gagtctggca 4380  
 ctgatggttg cattttgta attgttctaa gttggtggag cagaactttg cctacttatg 4440  
 tttattgtoa aatgcttcta tgcccatttc cattccctcc ataacagctt ctgtcttat 4500  
 ataatttttg ggaccagaa gaaacaacga cacaattotta gaactcctcc tgagtatctc 4560  
 gagttgtgga atttgttata gagttgacaa tttctgcat tatagcctct cattttccat 4620  
 gaattcatat ctgaaacct tttagaaggg agaagtoatc gaagtatttt ctgagtgttg 4680  
 agaagaatga gttaaacct ttaaaccat ttgaaacata caaaaataga aatgtgaaag 4740  
 catttggta aagccaaagc acagagtcag aagctgccac cttagagaac tgaataaaaa 4800  
 atagaagttc ttacgctttt ttgtgttaca gatgctttcg acaatttaa gaaagctaaa 4860  
 taaaaatgta gacttggctg ggcagttggc tcatgctgtt aatcctagca ctttttgagg 4920  
 ccaaggtagg aggattgctt gactcgggga gctcaaggca aagctgcaca acataacaag 4980  
 acctatctc cacaaaaaaa atgaaaaata aaactgggtg cgggtgctca cacotgtaat 5040  
 cccagcactt tgggaggcog atgtgggagc atcacaaggt caggagtcca agaccagcct 5100  
 ggccaacata gtgaaacccc atctctactg aaaatacaaa aattagctgg gtgtgggtggc 5160  
 acgtgcctgt tatctcagct acttgggaag ctga 5194

&lt;210&gt; 7

&lt;211&gt; 4455

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 7

tgcgaaacta cgggcgggca oggcttctcg gaagtaattt aagtgcacaa gacattggtc 60  
 aaaatggttt ccaaaagaag actgtcaaaa tctgaggata aagagagcct gacagaagat 120  
 gcctocaaaa ccaggagca accactttcc aaaaagacaa agaaatctca tattgctaat 180  
 gaagtgaag aaaatgacag catctttgta aagcttctta agatctcagg aattattctt 240  
 aaaacgggag agagtgcagc tcaactagct gtggatcaaa tagctttcca aaagaagctc 300  
 tttcagacc tgaggagaca ccctcctat cccaaaaata tagaagaatt tgttagtggc 360  
 ctggagtott acattgagga tgaagacagt ttcaggaaact gccttttgc tttgtgagct 420  
 ctgaggatg aggaagccag tatgggtgca tcttttcta agagtctcat caaactgott 480  
 ctggggattg acatactgca gctgocatt atcaaaact tatttgagaa gttgocagaa 540  
 tattttttg aaaaacagaa cagtgatgaa atcaacatac ctogactcat tctcagtaaa 600  
 ctaaatggc ttgacagagt tgtggatggc aaggacctca ccaccaagat catgcagctg 660

atcagtattg ctccagagaa cctgcagcat gacatcatca ccagcctacc tgagatccta 720  
ggggattccc agcacgctga tgtggggaaa gaactcagtg acctactgat agagaatact 780  
tcactcactg tcccaatcct ggatgtcctt tcaagcctcc gacttgacc aaacttcta 840  
ttgaaggttc gccagttggt gatggataag ttgtcgtcta ttagattgga ggatttacct 900  
gtgataataa agttcattct tcattccgta acagccatgg atacacttga ggtaatttct 960  
gagcttcggg agaagttgga tctgcagcat tgtgttttgc catcacggtt acaggcttcc 1020  
caagtaaagt tgaaaagtaa aggcagagca agttcctcag gaaatcaaga aagcagcggg 1080  
cagagctgta ttattctcct cttttagtga ataaagtcag ctattagata tgagaaaacc 1140  
atttcagaag cctggattaa ggcaattgaa aacactgcct cagtatctga acacaagggt 1200  
tttgacctgg tgatgctttt catcatctat agcaccaata ctcagacaaa gaagtacatt 1260  
gacaggggtc taagaaataa gattcgatca ggctgcatto aagaacagot gctccagagt 1320  
acattctctg ttacttactt agttcttaag gataigtgtt catccattct gtcgtggct 1380  
cagagtttgc ttactctctt agaccagagt ataatttcat ttggcagttc cctatacaaa 1440  
tatgcattta agttttttga cacgtactgc cagcaggaag tggttgggtc cttagtacc 1500  
catatctgca gtgggaatga agctgaagtt gatactgcct tagatgtct totagagt 1560  
gtagtgttaa accatctcgc tatgatgatg aatgctgtct ttgtaaaggg catttagat 1620  
tatctggata acatattccc tcagcaata ogaaaactct tctatgttct cagcacactg 1680  
gcatttagca aacagaatga agccagcagc cacatccagg atgacatgca cttgtgata 1740  
agaaaagcagc tctctagcac cgtattcaag tacaagctca ttgggattat tgggtcgtg 1800  
accatggctg gcatcatggc ggcagacaga agtgaatcac ctagtgtgac ccaagagaga 1860  
gccaacctga gogatgagca gtgcacacag gtgacctcct tgttgacgtt ggttcattcc 1920  
tgcagtgagc agtctcctca ggccctctgca ctttactatg atgaatttgc caacctgatc 1980  
caacatgaaa agctggatcc aaaaagccctg gaatgggttg gccatccat ctgtaatgat 2040  
ttocaggatg cttctgtagt ggactcctgt gttgttccgg aaggtgactt tccatttct 2100  
gtgaaagcac tttaccgact ggaaataac gacactcagg atgggattgc cataaacctc 2160  
ctgccctgca tgtttctca ggactttgca aaagatgggg gtccgggtgac ctcacaggaa 2220  
tcagtcocaaa aattgtgtgc tccgctgtgc ctggctcctg atttccggtt actgagactt 2280  
tgtgtggaga gacagcataa cggaaaactg gaggagattg atggtctact agattgtcct 2340  
atattcctaa ctgacctgga gccctggagag aagttggagt ccatgtctgc taaagagcgt 2400  
tcattcatgt gttctctcat atttcttact ctcaactggt tccgagagat tgtaaatgcc 2460  
ttctgccagg aaacatcacc tgagatgaag gggaaagtgc tcaactcggtt aaagcacatt 2520  
gtagaattgc aaataatcct ggaaaagtac ttggcagtca ccccagacta tgtccctcct 2580  
cttgaaaact ttgatgtgga aacttttagat ataacacctc atactgttac tgctatttca 2640  
gcaaaaaatca gaaagaaagg aaaaatagaa aggaacaaa aaacagatgg cagcaagaca 2700  
tctctctctg acacactttc agaagagaaa aattcagaat ttgaccttac gccctctcat 2760  
agaggccagc taacaagga gttcacaggg aaggaagaaa agacatcatt gttactacat 2820  
aattccatg ctttttccg agagctggac attgaggtct tctctattct acattgtgga 2880  
cttgtgacga agttcatctt agactactgaa atgcacactg aagotacaga agttgtgcaa 2940  
cttggccccc ctgagctgct ttcttctgct gaagatctct cccagaagct ggagagtatg 3000  
ctgacacctc ctattgccc gagagtcccc ttctcaaga acaaaagga cgggaatatt 3060  
ggattctcac atctccaaca gagatctgcc caagaaattg ttcatgtgt ttttcaactg 3120  
ctgaccccaa tgtgtaacca cctggagaac attcacaact attttcagtg ttttagctgt 3180  
gagaatcacg gtgtagttag tggaccagga gtgaaagtcc agggatacca cataatgtct 3240  
tctctctatc agaggctgct gcagattttt catgggcttt ttgctggag tggattttct 3300  
caacctgaaa atcagaattt actgtattca gccctccatg tcttagtag ccgactgaaa 3360  
cagggagAAC acagccagcc tttggaggaa ctactcagcc agagcgtcca ttacttgacg 3420  
aatttccatc aaagcattcc cagtttccag tgtgctcttt atctcatcag actttttagt 3480  
gttatttttg agaaatcaac agcttctgct cagaacaaag aaaaaattgc ttcccttggc 3540  
agacaattcc tctgtcgggt gtggccaagt ggggataaag agaagagcaa catctctaat 3600  
gaccagctcc atgctctgct ctgtatctac ctggagcaca cagagagcat totgaaggcc 3660  
atagaggaga ttgctgggtg tgggtgtcca gaactgatca actctcttaa agatgcactc 3720  
tctccacat tccctcact gaccagcat actttttgtg ttttctccg tgtgatgatg 3780  
gctgaactag agaagacggg gaaaaaatt gagcctggca cagcagcaga ctgcagcag 3840  
attcatgaag agaaactcct ctactggaac atggctgttc gagacttcag tatctctatc 3900  
aacttgataa aggtatttga tagtcatcct gttctgcatg tatgtttgaa gtatggcgt 3960  
ctctttgtgg aagcatttct gaagcaatgt atgccgtcc tagacttcag ttttagaaaa 4020  
caccgggaag atgttotgag cttactggaa acctccagt tggacacaag gctgcttcat 4080  
cacctgtgtg ggcattccaa gattccaccg gacacgagac tcaccaaca tgtgctctg 4140  
ctcaaaaaga cctggaact tttagtttgc agagtcaag ctatgctcac tctcaacaat 4200  
tgtagagagg ctttctggct gggcaatcta aaaaaccggg acttgacagg tgaagagatt 4260  
aagtcocaaa attccagga ggcacagca gatgagagt aggatgacat gtcatcccag 4320  
gcctccaaga gcaagccac tgaggatggt gaagaagac aagtaagtgc tgggaaaaag 4380  
gagcaagata gtgatgagag ttatgatgac tctgattaga ccccagataa attgttgcct 4440  
gottctgtgt ctcaa 4455

&lt;210&gt; 8

&lt;211&gt; 5516

&lt;212&gt; DNA

&lt;213&gt; Mus musculus

&lt;400&gt; 8

gaaaagtcca aaacgaagg aagcaactgg cgggtcccca ggaagtaata taagtggcag 60  
aagacgttag tcaaatgat ttccaaaaga cgtcggctag attctgagga taaagaaaac 120  
ctgacagaag atgcctccaa aacctgccc ctttccaagc tggcaagaa gtctcaaat 180

tctcatgaag ttgaagaaaa tggcagtgct tttgtaaagc ttcttaaggo ttcaggactc 240  
 actcttaaaa ctggagagaa ccaaaatcag ctagggtggt atcaggtaat cttccaaagg 300  
 aagctcttct aggccttgag gaagcatcct gcttatccca aagtaataga agagtttgt 360  
 aatggcctgg agtccctacac tgaggacagt gagagtctca ggaactgcct gctgtcttgt 420  
 gagcgctgg aggatgagga agccagcatg ggcacatttt actccaagag tctgatcaag 480  
 ctactcttgg ggattgacat ttacagcct gccattatca aaatgttatt tgaaaaagt 540  
 cctcagtttc tttttgaaag tgagaacaga gatggaatca acatggccag actcattato 600  
 aatcaactaa aatggotgga tagaattgtg gatggcaagg acctcacggc ccagatgatg 660  
 cagttgatca gtgttgcctc cgtgaactta cagcatgact tcatcagcag ccttctgaa 720  
 atcctagggg atttccagca tcttaatgtg gggaaagagc ttggogagct gotggtgcag 780  
 aatacttccc tgactgttcc aattttggat gtcttttcca gtctccgact tgaccccaac 840  
 ttctgttcca agatccgcca gttgggtgat ggcaagotgt catctgtccg tctagaggat 900  
 ttccctgtga ttgtaaagtt ccttcttcat tctgtaacag acaccacttc ccttgaggtc 960  
 attgocgagc ttgggagaa cttgaacgct cagcagttta tttgcccgtc acgaattcag 1020  
 gcttcccaaa gcaaatgaa aagtaaagga ctagcaagct cttcaggaaa tcaagagaac 1080  
 agtgataaag actgtatigt tcttgtcttt gatgtaataa agtcagccat tagatatgag 1140  
 aaaaccattt cagaggcctg gtttaaggca attgaacgca ttgagtcocg ggctgaacat 1200  
 aaggctttgg agctgtctag ttctgctatc atctacagca ccagcacgca gaccaagaag 1260  
 ggcgtggaga agctgtctag aaacaagatt cagtcagact gcattcaaga acagctgctt 1320  
 gacagtgctt tctctacaca ttacctggtt ctttaaggata tttgcccact tattcttttg 1380  
 ctggctcaga ctttgtttca ctctcaagac cagaggatca ttttgtttgg cagtcttctg 1440  
 tacaaatag ctttttaagt ttigtact tactgccagc aggaagtgtt tgggtcccta 1500  
 gtcacccatg totgcaagtg gactgaggct gaagtcgaca ctgcacttga tctctctctg 1560  
 gagctgattg tgctaaacgc ctctgctatg aggotcaatg ctgcttttgt taagggcact 1620  
 ttagattatt tggaaaatat gtcccctcag caaatacgaa aaatctcttg tattctcagc 1680  
 actcttgcct tttagccaac gcccggaacc agcaaccata tccaggacga catgcacctg 1740  
 gtgacccgga agcagctctc tagcaactgtg ttaagtaaca agctcattgg gatcattggt 1800  
 gcagtcacca tggccggcat catggcggaa gacagaagtg taccatctaa ctcatcccag 1860  
 aggagcgcca atgtgagcag tgagcagcgc acacaggtga cttcttttct acaactagtt 1920  
 cattcttga ctgagcactc tccctggccc tcttctctgt attatgatga atttgccaac 1980  
 ctgatccaag aaaggagatt ggctccaaaa accottggagt gggttgggca gacctcttc 2040  
 aatgatttcc aagatgcctt tgtggtagac ttctgtgtgt ctccagaggg tgactttcca 2100  
 tttctgtga aagcgtctca tggactggaa gactacagca ctcaagacgg cattgtcact 2160  
 aacctctgct cgtgttctca tcaggaatgt gcaaaagatg ccagtcgagc gacatcaca 2220  
 gaatcgagcc agagctgagc gtcttcttgg tgctggcctt cccatttccg gctgotgaga 2280  
 ctttgotggt caagacaaca tgatggaaac ttggatgaga tcatggtctt cttagattgt 2340  
 cccctgttcc tccctgacct ggaacctgga gagaacctgg agtccatgct tgctaaagac 2400  
 cgttctgcta tgtgttctct cacattctca actttcaact ggttccgaga ggttgtgat 2460  
 gccctctgca aacaacatct tcttgatg atgggcaagg ttcttagtct gctaaaggac 2520  
 cttgtagaac ttcagggaa cctagagaag tacttggcag tcatcccaga ctatgttccg 2580  
 cctttcgcaa gctgtgactt ggacacttta gatatgatgc ctaggagcag ttctgctgtt 2640  
 gcagcaaaaa acagaaacaa gggaaagacg ggggaaaga acaaaaaagc tgatagcaac 2700  
 aaagcatcct gttcggagaa acttcaaca gaagacactt cacagtgtag catggcgcca 2760  
 tctgggagaa gccacgtaga caaggagctc acagggaaag aaggaaagac gtttgtgca 2820  
 ctgcagaatt accgcgttct ttccgagag ctggacattg aggtcttctc tattctacat 2880  
 tctggacttg tgaccaagtt catcttagac actgaaatgc aactgaagc tacagaggct 2940  
 gtacagctgg gctcagctct gctgctcttc ttgctggaag atctttcca gaagctagag 3000  
 aatatgctga ctgctctctt tgccaagaga atctgctgct ttaagaataa aggaaggcag 3060  
 aatattggct totcacatct tcatcagaga tctgtccagg acattgtgca ctgtgtggtt 3120  
 cagctgctaa ccccgatgtg taacctctg gagaacattc acaacttctt tcagtgttta 3180  
 ggtgctgagc atctcagctg agatgacaag gcgagagcga cagctcagga gcagcacacc 3240  
 atggcctgct ctacattccc tactgtgacg gtcttgcacg cgtcttttgc gtggaaggga 3300  
 tttactcacc aatcaaaagca ccgcctctg cactcagccc ttgaggtctt ctggaaccga 3360  
 ctaaaagcaga tggaaacagga ccagcccttg gaggaaactg tcagccagag cttcagttac 3420  
 ttgagaactc tccaccatag tgttcccagt ttccagtggt gctctacctt tctcagactt 3480  
 ctgatggccc tttcggagaa gctctgagta cctaaccaga agaaagaaaa acttgcctct 3540  
 ctggccaaac agctgotttg ccgagcatgg cctcatgggg aaaaagagaa gaacccact 3600  
 tttaatgacc acctgcatga tgtgctttac atctacttgg agcacacaga caatgttctg 3660  
 aaggccatag aggagatyac tgggtttggt gtcccagaac ttgctcagtc tccgaaagac 3720  
 gccgctctct ctacattccc tactgtgacc grgcacacct ttgtcatatt cttccgtgtg 3780  
 atgatggctg aactcgagaa gacgggtgaag ggtctycagg ctggcacagc agcagattcg 3840  
 cagcaggttc acgaagagaa gctcctctat tkgaacatgg ctgtccgaga tttcagyatc 3900  
 ctttcaatc tgatgaaagt atttgacagt tatctgttct tgcagtgtgt tttaaagtat 3960  
 ggccgtctgt ttgtggagc atttctgaag caatgtatgc cactctcga cttcagcttt 4020  
 agaaagcact gggaaagtgt totgagcttg ctgcaaaccc ttcagttgaa cacgaggota 4080  
 cttcatcacc tttgtggaca ctccaagatt ccgacggaca caagactcac caagcaytg 4140  
 cctttactca aaaagtcact ggaactgtta gtttgacag tcaaaagccat gcttgtctc 4200  
 aacaactgta gagaggcttt ctggttgggt actctcaaaa accgagactt acagggtgaa 4260  
 gaaattttt cccaggatoc ctcttctca gagagcaatg cagaggacag tgaggatggc 4320  
 gtgacatctc acgtctccag gaacagagca acagaggatg ggaagatga agcaagtgat 4380  
 gaacagaagg accaggacag tgatgaaagt gacgacagct ccagttagag ccgagtggca 4440  
 tggctccctg gctcaccctg gacagactct catctctttg gggtttgaag tcagatgtct 4500  
 gttttctag tcagaagcat cctgtttgtc catcaagaag ggggttttat ttaattcccc 4560  
 agtgggttcc acaggttctc taacctccag gtccctggtt caggagtcca gtgtagcatc 4620

```

catcgttgac taggaygaac atggctgggc tgcagtgcag tkcagtgcag gtgccctagc 4680
tgggcottgg ggttttgaaa ctaaaattta ggottataat agctttgtaa ataaatctgt 4740
ttcagagttt tgcctcagct acotttttcc tcactttaga tgtgattatt caaggatctc 4800
attatccaag gattaggtaa tattgagttg aggtttgtgc aatcgtactg gtggcctaaa 4860
agtatgttcc gtaactgttat cttcctggag gaatgaccca actttcttat caatgatcaa 4920
gtgtttggtt tggctctggt cagggctctc ttacatagtc ctggctggtg tgttattaga 4980
tatgttgacc aggagggtct tgaacattac ttttgaattt taaacatttt tgtacatatg 5040
tgtatgggca tatatgtgcc actgtgcata tgtgtaggtc agaggatagc ttatgggagt 5100
gagctctctc cttccaccat gtgggttcca gggttcaaac tctagacctt cacctgctca 5160
ggcaccttac ctttttaaaa tgtttgggta ttaatatata aaaggaagga agacaacatc 5220
aaacatgtgc tggctttgta tgtatatata gtttttattt ccacattaat ttgaattatg 5280
cctataatat atttgaata atcatacaaa ataattgtaa tttattagaa atagaacatc 5340
aggagttaaa ataggggatt cttctgtctt ctgccaggaa gccagctctc agagatgctg 5400
ccaggctctt cctcgtgtg ccattaagat tatttaattt ttgttaatat ttttactcat 5460
accggtatta aagttatgtt ttgttggaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 5516

```

<210> 9

<211> 14

<212> DNA

<213> Homo sapiens

<220>

<221> MISC\_FEATURE

<222> (1).. (14)

<223> Intron/Exon Junction of FANCD

<400> 9

tgggtgagta agtg

14

<210> 10

<211> 14

<212> DNA

<213> Homo sapiens

<220>

<221> MISC\_FEATURE

<222> (1).. (14)

<223> Intron/Exon Junction of FANCD

<400> 10

ccagtaagta tata

14

<210> 11

<211> 14

<212> DNA

<213> Homo sapiens

<220>

<221> MISC\_FEATURE

<222> (1).. (14)

<223> Intron/Exon Junction of FANCD

<400> 11

taggtaatat tttt

14

<210> 12

<211> 14

<212> DNA

<213> Homo sapiens

<220>

<221> MISC\_FEATURE

<222> (1).. (14)

<223> Intron/Exon Junction of FANCD

<400> 12

aaagtatgta tttt

14

<210> 13

<211> 14

<212> DNA

<213> Homo sapiens

<220>

<221> MISC\_FEATURE

<222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 13  
 caggtgtgga gagg 14  
  
 <210> 14  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 14  
 caggtaagac tgtc 14  
  
 <210> 15  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 15  
 aaagtaagtg gcgt 14  
  
 <210> 16  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 16  
 aaggtaggct tatg 14  
  
 <210> 17  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 17  
 caggtggata aacc 14  
  
 <210> 18  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 18  
 aagtagaaa agac 14  
  
 <210> 19  
 <211> 14  
 <212> DNA

<213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 19  
 gaggtatgct cttta 14  
  
 <210> 20  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 20  
 aaggtaaaga gctc 14  
  
 <210> 21  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 21  
 aagtgagat cttt 14  
  
 <210> 22  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 22  
 aagtaagtgc tcat 14  
  
 <210> 23  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 23  
 ttagtaagtgc tcag 14  
  
 <210> 24  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 24  
 caggtatggt gaaa 14

<210> 25  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 25  
 aaggtatcctt attg 14  
  
 <210> 26  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 26  
 caggttagag gcaa 14  
  
 <210> 27  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 27  
 caggtacacg tgga 14  
  
 <210> 28  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 28  
 caggtgagtt cttt 14  
  
 <210> 29  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 29  
 ctggtaaagc caat 14  
  
 <210> 30  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)

<223> Intron/Exon Junction of FANCD  
 <400> 30  
 aggttagta ttgt 14  
 <210> 31  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
 <400> 31  
 aaagtcagta tagt 14  
 <210> 32  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
 <400> 32  
 taggtatggg atga 14  
 <210> 33  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
 <400> 33  
 gaggtgagca gagt 14  
 <210> 34  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
 <400> 34  
 caggtaaagag aagt 14  
 <210> 35  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
 <400> 35  
 taggtaagta tggt 14  
 <210> 36  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 36  
 aaggattgg aatg 14  
  
 <210> 37  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 37  
 gaagtaagtg acag 14  
  
 <210> 38  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 38  
 aaggtagtg tagg 14  
  
 <210> 39  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 39  
 caggtcagaa gcct 14  
  
 <210> 40  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 40  
 ttgtaagta tgtg 14  
  
 <210> 41  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 41  
 caggtagtc ataa 14

<210> 42  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 42  
 ttggtgatgg gcct 14  
  
 <210> 43  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 43  
 ctggtgagat gttt 14  
  
 <210> 44  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 44  
 caggtgaggg agtt 14  
  
 <210> 45  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 45  
 caggtgagta agat 14  
  
 <210> 46  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 46  
 aaggtgagta tgga 14  
  
 <210> 47  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD

<400> 47  
 aaggtgagag attt 14  
  
 <210> 48  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 48  
 cgggtaagag ctaa 14  
  
 <210> 49  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 49  
 aaggtaagaa gggg 14  
  
 <210> 50  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 50  
 caggtaagcc ttgg 14  
  
 <210> 51  
 <211> 14  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(14)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 51  
 gaggtatctc taca 14  
  
 <210> 52  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 52  
 gtttccgat ttgctctag gaa 23  
  
 <210> 53  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 53  
 gaaaattttt ctatttcag aaa 23  
  
 <210> 54  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 54  
 ctcttttttt ttctgcatag ctg 23  
  
 <210> 55  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 55  
 attttttaaa totcottaag ata 23  
  
 <210> 56  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 56  
 gatttctttt ttttttacag tat 23  
  
 <210> 57  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 57  
 ccctatgtct tcttttttag cct 23  
  
 <210> 58  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 58  
 ttctcttcct aacattttag caa 23  
  
 <210> 59

<211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 59  
 aatagtgtct tctactgcag gac 23  
  
 <210> 60  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 60  
 ttttttcta ccattcacag tga 23  
  
 <210> 61  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 61  
 tctgtgcttt taatttttag gtt 23  
  
 <210> 62  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 62  
 ctaatattta ctttctgcag gta 23  
  
 <210> 63  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 63  
 ttctctctg ctactttag ttc 23  
  
 <210> 64  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD

<400> 64  
 actctctcct gtttttcag gca 23  
 <210> 65  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
 <400> 65  
 tgcataatta ttgacaatag gtc 23  
 <210> 66  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
 <400> 66  
 tetactotctt cccactcaag gtt 23  
 <210> 67  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
 <400> 67  
 gttgactctc cctgtatag gaa 23  
 <210> 68  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
 <400> 68  
 tggcatcatt tttccacag ggc 23  
 <210> 69  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
 <400> 69  
 ttttcatcat ctcatgag gat 23  
 <210> 70  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <220>

<221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 70  
 aaaaaattct ttgttttag aag 23  
  
 <210> 71  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 71  
 attcttctctc ttgctccag gtg 23  
  
 <210> 72  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 72  
 tgtttgttg cttcctgaag gaa 23  
  
 <210> 73  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 73  
 attctggttt ttctccgag tga 23  
  
 <210> 74  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 74  
 aattatttc tccttctcag att 23  
  
 <210> 75  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 75  
 aaatgtttgt tctctcag att 23  
  
 <210> 76  
 <211> 23

<212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 76  
 atgtaatttg tactttgcag att 23  
  
 <210> 77  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 77  
 cagcotgctg ttgtttcag tca 23  
  
 <210> 78  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 78  
 ttctttttt aatataaaag aaa 23  
  
 <210> 79  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 79  
 ttgtgtgac ttcccatag gag 23  
  
 <210> 80  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 80  
 tcctttctc catgtgacag got 23  
  
 <210> 81  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD

<400> 81  
 taactctgca tttattatag aac 23  
  
 <210> 82  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 82  
 aaaatcattt ttatttttag tgt 23  
  
 <210> 83  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 83  
 ttttaccttg acttccttag gag 23  
  
 <210> 84  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 84  
 ttttttctgt ctccctacag cca 23  
  
 <210> 85  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 85  
 ttgtcttct tttctaacag ctt 23  
  
 <210> 86  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 86  
 atatttgact ctcaatgcag tat 23  
  
 <210> 87  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>

<221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 87  
 atgcttttcc cgtcttctag gca 23  
  
 <210> 88  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 88  
 catatatttg gctgccccag att 23  
  
 <210> 89  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 89  
 cttgttcttc acctctccag gta 23  
  
 <210> 90  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 90  
 agtgtgtctc tcttcttcag tat 23  
  
 <210> 91  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 91  
 tataaactta ttggttatag gaa 23  
  
 <210> 92  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1).. (23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 92  
 tgttatttat ttccattcag att 23  
  
 <210> 93  
 <211> 23

<212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 93  
 cttggtccat tcacatttag ggt 23  
  
 <210> 94  
 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(23)  
 <223> Intron/Exon Junction of FANCD  
  
 <400> 94  
 atttattctt tgcccottag gat 23  
  
 <210> 95  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> DF4EcoRI  
  
 <400> 95  
 agcctcgaat togtttccaa aagaagactg tca 33  
  
 <210> 96  
 <211> 35  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> DR816Xh  
  
 <400> 96  
 ggtatcctcg agtcaagacg acaacttata catca 35  
  
 <210> 97  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG471  
  
 <400> 97  
 aatcgaaaac tacgggcg 18  
  
 <210> 98  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG457  
  
 <400> 98  
 gagaacacat gaatgaacgc 20  
  
 <210> 99  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>

<223> MG492  
 <400> 99  
 ggcgacggct tctcggaagt aatttaag 28  
 <210> 100  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG472  
 <400> 100  
 agcggcagga ggtttatg 18  
 <210> 101  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG474  
 <400> 101  
 tggcggcaga cagaagtg 18  
 <210> 102  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG475  
 <400> 102  
 tggcggcaga cagaagtg 18  
 <210> 103  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG491  
 <400> 103  
 agagagccaa cctgagcgat g 21  
 <210> 104  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG476  
 <400> 104  
 gtgccagact ctggtggg 18  
 <210> 105  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG792  
 <400> 105  
 aggagacacc cttcctatcc 20  
 <210> 106  
 <211> 20  
 <212> DNA

<213> Artificial Sequence	
<220>	
<223> MG803	
<400> 106	
gaagttggca aaacagactg	20
<210> 107	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG924	
<400> 107	
tgtcttgtga gogtctgcag g	21
<210> 108	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG753	
<400> 108	
aggttttgat aatggcaggc	20
<210> 109	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG979	
<400> 109	
actggactgt gcctaccac tatg	24
<210> 110	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG984	
<400> 110	
cctgtgtgag gatgagctct	20
<210> 111	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> R302W	
<400> 111	
ttctccgaa gctcag	16
<210> 112	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> R1236H	
<400> 112	
tttcttcogt gtgatga	17

<210> 113  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> DF4EcoRI  
  
 <400> 113  
 agcctcgaat tcgtttccaa aagaagactg tca 33  
  
 <210> 114  
 <211> 35  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> DR816Xh  
  
 <400> 114  
 ggtatcctcg agtcaagacg acaacttato catca 35  
  
 <210> 115  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG914  
  
 <400> 115  
 ctagcacaga actotgctgc 20  
  
 <210> 116  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG837  
  
 <400> 116  
 ctagcacaga actotgctgc 20  
  
 <210> 117  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG746  
  
 <400> 117  
 cttcagcaac agcgaagtag tctg 24  
  
 <210> 118  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG747  
  
 <400> 118  
 gattctcagc acttgaaaag cagg 24  
  
 <210> 119  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG773

<400> 119 ggacacatca gttttcctct c	21
<210> 120 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> MG789	
<400> 120 gaaaaccat gattcagtcc	20
<210> 121 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> MG816	
<400> 121 tcacaggca agaaacttgg	20
<210> 122 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> MG803	
<400> 122 gaagttggca aaacagactg	20
<210> 123 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> MG804	
<400> 123 gagccatctg ctcatttotg	20
<210> 124 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> MG812	
<400> 124 cccgtattt agacttgagc	20
<210> 125 <211> 21 <212> DNA <213> Artificial Sequence	
<220> <223> MG775	
<400> 125 caaagtgttt attccaggag c	21
<210> 126 <211> 22 <212> DNA <213> Artificial Sequence	

<220>  
 <223> MG802  
 <400> 126  
 catcaggga ctttgaacat tc 22  
 <210> 127  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG727  
 <400> 127  
 ttgaccagaa aggctoagtt cc 22  
 <210> 128  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG915  
 <400> 128  
 agatgatgoc agagggttta tcc 23  
 <210> 129  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG790  
 <400> 129  
 tgcccagctc tgttcaaacc 20  
 <210> 130  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG774  
 <400> 130  
 aggcaatgac tgactgacac 20  
 <210> 131  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG805  
 <400> 131  
 tgcccgctcta tttttgatga agc 23  
 <210> 132  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG791  
 <400> 132  
 tctcagttag tctggggaca g 21  
 <210> 133  
 <211> 25

<212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG751  
  
 <400> 133  
 tcatggtaga gagactggac tgtgc 25  
  
 <210> 134  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG972  
  
 <400> 134  
 accctggagc aaatgacaac c 21  
  
 <210> 135  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG973  
  
 <400> 135  
 atttgctcca gggtacatgg c 21  
  
 <210> 136  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG974  
  
 <400> 136  
 gaaagacagt gggaaggcaa gc 22  
  
 <210> 137  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG975  
  
 <400> 137  
 gggagtgtgt ggaacaaatg agc 23  
  
 <210> 138  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG976  
  
 <400> 138  
 agtttctaca ggctggctct attcc 25  
  
 <210> 139  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG755  
  
 <400> 139  
 aacgtggaat cccattgatg c 21

<210> 140  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG730  
  
 <400> 140  
 tttctgtgtt ccctccttgc 20  
  
 <210> 141  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG794  
  
 <400> 141  
 gatggtcaag ttacactggc 20  
  
 <210> 142  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG778  
  
 <400> 142  
 cacctcccac caattatagt attc 24  
  
 <210> 143  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG808  
  
 <400> 143  
 ctatgtgtgt ctcttttaca ggg 23  
  
 <210> 144  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG817  
  
 <400> 144  
 aatctttccc accatattgc 20  
  
 <210> 145  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG779  
  
 <400> 145  
 cataccttet tttgctgtgc 20  
  
 <210> 146  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG795

<400> 146 ccacagaagt cagaatctcc acg	23
<210> 147 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> MG731	
<400> 147 tgtaacaaac ctgcacgttg	20
<210> 148 <211> 23 <212> DNA <213> Artificial Sequence	
<220> <223> MG732	
<400> 148 tgctacccaa gccagtagtt tcc	23
<210> 149 <211> 22 <212> DNA <213> Artificial Sequence	
<220> <223> MG788	
<400> 149 gagtttggga aagattggca gc	22
<210> 150 <211> 22 <212> DNA <213> Artificial Sequence	
<220> <223> MG772	
<400> 150 tgtagtaaag cagctctcat gc	22
<210> 151 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> MG733	
<400> 151 caagtacact ctgcactgcc	20
<210> 152 <211> 23 <212> DNA <213> Artificial Sequence	
<220> <223> MG758	
<400> 152 tgactcaact tccccaccaa gag	23
<210> 153 <211> 24 <212> DNA <213> Artificial Sequence	

<220>  
 <223> MG736  
  
 <400> 153  
 ctccctatgt acgtggagta atac 24  
  
 <210> 154  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG737  
  
 <400> 154  
 gggagtcttg tggaactaa g 21  
  
 <210> 155  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG780  
  
 <400> 155  
 ttcatagaca tctctcagct ctg 23  
  
 <210> 156  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG759  
  
 <400> 156  
 gttttggtat cagggaaagc 20  
  
 <210> 157  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG760  
  
 <400> 157  
 agccatgctt ggaattttgg 20  
  
 <210> 158  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG781  
  
 <400> 158  
 ctcaactggga tgcacaaac 20  
  
 <210> 159  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG740  
  
 <400> 159  
 ggtcttgatg tgtgacttgt atccc 25  
  
 <210> 160

<211> 25  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG741  
  
 <400> 160  
 cctcagtgtc acagtgttct ttgtg 25  
  
 <210> 161  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG809  
  
 <400> 161  
 catgaaatga ctaggacatt cc 22  
  
 <210> 162  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG797  
  
 <400> 162  
 ctaccacagtg acccaaacac 20  
  
 <210> 163  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG761  
  
 <400> 163  
 cgaaccctta gtttctgaga cgc 23  
  
 <210> 164  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG742  
  
 <400> 164  
 tcagtgccctt ggtgactgtc 20  
  
 <210> 165  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG916  
  
 <400> 165  
 ttgatggtac agactggagg c 21  
  
 <210> 166  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> MG810  
  
 <400> 166

aagaaagttg ccaatcctgt tcc	23
<210> 167	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG762	
<400> 167	
agcacctgaa aataaggagg	20
<210> 168	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG743	
<400> 168	
gccccaaagt tgaagtgtg ag	22
<210> 169	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG787	
<400> 169	
agcaagaatg aggtcaagtt c	21
<210> 170	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG806	
<400> 170	
gggaaaaact ggaggaaaga actc	24
<210> 171	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG818	
<400> 171	
agaggtaggg aaggaagcta c	21
<210> 172	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> MG813	
<400> 172	
ocaaagtcca cttcttgaag	20
<210> 173	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	

<223> MG834  
 <400> 173  
 gatgcactgg ttgotacatc 20  
 <210> 174  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG836  
 <400> 174  
 ccaggacact tggtttctgc 20  
 <210> 175  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG839  
 <400> 175  
 acaactcccag ttggaatcag 20  
 <210> 176  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG871  
 <400> 176  
 cttgtgggca agaaattgag 20  
 <210> 177  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG829  
 <400> 177  
 tgggctggat gagactatc 20  
 <210> 178  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG870  
 <400> 178  
 ccaaggacat atcttctgag caac 24  
 <210> 179  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG820  
 <400> 179  
 tgattatcag cataggctgg 20  
 <210> 180  
 <211> 20  
 <212> DNA

<213> Artificial Sequence  
 <220>  
 <223> MG811  
 <400> 180  
 gatcccccaa tagcaactgc 20  
 <210> 181  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG763  
 <400> 181  
 cattcagatt caccaggaca c 21  
 <210> 182  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG782  
 <400> 182  
 ccttacatgc catctgatgc 20  
 <210> 183  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG764  
 <400> 183  
 aaccttctcc cctattacc 20  
 <210> 184  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG835  
 <400> 184  
 ggaaaatgag aggctataat gc 22  
 <210> 185  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG1006  
 <400> 185  
 tgtattccag aggtcaccca gagc 24  
 <210> 186  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> MG1005  
 <400> 186  
 ccagtaagaa aggcaaacag og 22

<210> 187  
 <211> 1094  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> FANCD-S. ORF. s

<400> 187  
 ggagaacaca gccagccttt ggaggaacta ctcagccaga gcgtccatta cttgcagaat 60  
 ttccatcaaa gcattcccag ttccagtggt gctctttatc tcatcagact tttgatgggt 120  
 attttggaga aatcaacagc ttctgctcag aacaaagaaa aaattgcttc ccttgccaga 180  
 caattcctct gtcgggtgtg gccaaagtgg gataaagaga agagcaacat ctctaatac 240  
 cagctccatg ctctgctctg tatctacctg gagcacacag agagcattct gaaggccata 300  
 gaggagattg ctgggtgttg tgtcccagaa ctgatcaact ctctaaaga tgcattctcc 360  
 tccacattcc ctacactgac caggcatact ttgtgtgtt tcttccgtgt gatgatggct 420  
 gaactagaga agacgggtgaa aaaaattgag cctggcacag cagcagactc gcagcagatt 480  
 catgaagaga aactcctcta ctggaacatg gctgttcgag acttcagtat cctcatcaac 540  
 ttgataaagg tatttgatag tcatcctgtt ctgcatgtat gtttgaagta tggcgtctc 600  
 tttgtggaag cattctgtaa gcaatgtat ccgctcctag acttcagttt tagaaaacac 660  
 cgggaagatg ttctgagctt actggaacc ttccagttgg acacaagggt cttcatcacc 720  
 tgtgtgggca ttccaagatt caccaggaca cgagactcac ccaacatgtg cctctgctca 780  
 aaaagaccct ggaactttta gtttcagag tcaaagctat gctcactctc aacaacaati 840  
 gtagagaggc tttctggctg ggcaatctaa aaaaccggga cttgcagggt gaagagatta 900  
 agtcccaaaa ttcccaggag agcacagcag atgagagtga ggatgacatg tcatcccagg 960  
 cctccaagag caaagccact gaggatgggt aagaagacga agtaagtgtt ggagaaaagg 1020  
 agcaagatag tgtgagagt tatgatgact ctgattagac cccagataaa ttgttgcctg 1080  
 cttctgtgtc tcaa 1094

<210> 188  
 <211> 1115  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> FANCD cDNA OR

<400> 188  
 ggagaacaca gccagccttt ggaggaacta ctcagccaga gcgtccatta cttgcagaat 60  
 ttccatcaaa gcattcccag ttccagtggt gctctttatc tcatcagact tttgatgggt 120  
 attttggaga aatcaacagc ttctgctcag aacaaagaaa aaattgcttc ccttgccaga 180  
 caattcctct gtcgggtgtg gccaaagtgg gataaagaga agagcaacat ctctaatac 240  
 cagctccatg ctctgctctg tatctacctg gagcacacag agagcattct gaaggccata 300  
 gaggagattg ctgggtgttg tgtcccagaa ctgatcaact ctctaaaga tgcattctcc 360  
 tccacattcc ctacactgac caggcatact ttgtgtgtt tcttccgtgt gatgatggct 420  
 gaactagaga agacgggtgaa aaaaattgag cctggcacag cagcagactc gcagcagatt 480  
 catgaagaga aactcctcta ctggaacatg gctgttcgag acttcagtat cctcatcaac 540  
 ttgataaagg tatttgatag tcatcctgtt ctgcatgtat gtttgaagta tggcgtctc 600  
 tttgtggaag cattctgtaa gcaatgtat ccgctcctag acttcagttt tagaaaacac 660  
 cgggaagatg ttctgagctt actggaacc ttccagttgg acacaagggt cttcatcacc 720  
 tgtgtgggca ttccaagatt caccaggaca cgagactcac ccaacatgtg cctctgctca 780  
 aaaagaccct ggaactttta gtttcagag tcaaagctat gctcactctc aacaattgta 840  
 gagaggcttt ctggctgggc aatctaaaa accgggactt gcagggtgaa gagattaagt 900  
 cccaaaattc ccaggagagc acagcagatg agagtgagga tgacatgtca tcccaggcct 960  
 ccaagagcaa agccactgag gtatctctac aaaaccacc agagtctggc actgatgggt 1020  
 gcattttgtt aattgttcta agttggtgga gcagaacttt gctacttat gtttattgtc 1080  
 aatgtctct atgccattt ccattccctc cataa 1115

<210> 189  
 <211> 52  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Wild-type

<400> 189  
 ggaagccagg tgtggagagg aggcattgaa tcttctgaa attcagctctg tc 52

<210> 190  
 <211> 52  
 <212> DNA  
 <213> Artificial Sequence

<220>

<223> Maternal Mutation

<400> 190

ggaagccggg tgtggagagg aggcattgaa tcttgctgaa attcagtctg tc 52

<210> 191

<211> 52

<212> DNA

<213> Artificial Sequence

<220>

<223> Maternal Mutation and Reversion

<400> 191

ggaagccggg tgtgaagagg aggcattgaa tcttgctgaa attcagtctg tc 52

We claim:

1. An isolated nucleic acid molecule, comprising: a polynucleotide selected from
  - (a) a nucleotide sequence encoding a polypeptide having an amino acid sequence as shown in SEQ ID NO: 4
  - (b) a nucleotide sequence at least 90% identical to the nucleotide sequence of (a);
  - (c) a nucleotide sequence complementary to the nucleotide sequence of (b);
  - (d) a nucleotide sequence at least 90% identical to the nucleotide sequence shown in SEQ ID NO: 5-8, 187-188; and
  - (e) a nucleotide sequence complementary to the nucleotide sequence of (d).
2. An isolated nucleic acid molecule according to claim 1, wherein the polynucleotide is a DNA molecule
3. An isolated nucleic acid molecule of claim 2, wherein the polynucleotide is cDNA.
4. An isolated nucleic acid molecule according to claim 1, wherein the polynucleotide is an RNA molecule.
5. An isolated nucleic acid molecule consisting essentially of a nucleotide sequence encoding a polypeptide having an amino acid sequence sufficiently similar to that of SEQ ID NO: 4 to retain the biological property of conversion from a short form to a long form of FANCD2 in the nucleus of a cell for facilitating DNA repair.
6. An isolated nucleic acid molecule consisting essentially of a polynucleotide having a nucleotide sequence at least 90% identical to SEQ ID NO: 9-191 or complementary to a nucleotide sequence that is at least 90% identical to SEQ ID NO: 9-191.
7. An isolated nucleic acid molecule according to claim 6, wherein the sequence is an intron/exon sequence selected from SEQ ID NO: 9-94 disclosed in Table 6.
8. An isolated nucleic acid molecule according to claim 6, wherein the sequence is a PCR primer selected from SEQ ID NO: 115-186 disclosed in Table 7.
9. A method for making a recombinant vector comprising inserting the

isolated nucleic acid molecule of claim 1 into a vector.

10. A recombinant vector produced by the method of claim 9.

11. A method of making a recombinant host cell comprising:  
introducing the recombinant vector of claim 10 into a host cell.

5 12. A recombinant host cell produced by the method of claim 11.

13. A method of making an FA-D2 cell line, comprising:  
(a) obtaining cells from a subject having a biallelic mutation in a  
complementation group associated with FA-D2; and  
(b) infecting the cells with a transforming virus to make the FA-D2  
10 cell line.

14. A method according to claim 13, wherein the cells are selected from  
fibroblasts and lymphocytes.

15. A method according to claim 13, wherein the transforming virus is  
selected from Epstein Barr virus and retrovirus.

15 16. A method according to claim 13, further comprising:  
characterizing the FA-D2 cell line by determining the presence of a defective  
FANCD2.

17. A method according to claim 16, wherein characterizing the FA-D2  
cell line further comprises:

20 performing a diagnostic assay on the cell line, the diagnostic assay  
selected from (i) a Western blot or nuclear immunofluorescence using an antibody  
specific for FANCD2 and (ii) a DNA hybridization assay.

18. A recombinant method for producing a polypeptide, comprising:  
culturing a recombinant host cell wherein the host cell comprises the isolated nucleic  
25 acid molecule of claim 1.

19. An isolated polypeptide, comprising an aminoacid sequence selected  
from

(a) SEQ ID NO: 4;

(b) an aminoacid sequence at least 90% identical to (a);

30 (c) an aminoacid sequence which is encoded by a polynucleotide  
having a nucleotide sequence which is at least 90% identical to at least one of SEQ ID  
NO: 5-8, 187-188; and

(d) a polypeptide fragment of (a) – (d) wherein the fragment is at least  
50 aminoacids in length.

20. An isolated polypeptide according to claim 19, encoded by a DNA having a mutation selected from nt 376 A to G, nt 3707 G to A, nt904C to T and nt 958C to T.
21. An isolated polypeptide according to claim 19, the polypeptide  
5 characterized by a polymorphism in DNA encoding the polypeptide, the polymorphism being selected from nt 1122A to G, nt 1440T to C, nt1509C to T, nt2141C to T, nt2259T to C, nt4098T to G, nt4453G to A.
22. An isolated polypeptide according to claim 19, the polypeptide characterized by a mutation at aminoacid 222 or aminoacid 561.
- 10 23. An antibody preparation having a binding specificity for a FANCD2 protein.
24. An antibody preparation according to claim 23, further comprising: monoclonal antibodies.
25. An antibody preparation according to claim 23, further comprising:  
15 polyclonal antibodies.
26. An antibody preparation according to claim 23, wherein the FANCD2 protein is FANCD2-S
27. An antibody preparation according to claim 23, wherein the FANCD2 protein is FANCD2-L.
- 20 28. A diagnostic method for measuring FANCD2 isoforms in a biological sample, the method comprising:
- (a) exposing the sample to a first antibody for forming a first complex with FANCD2-L and optionally a second antibody for forming a second complex with FANCD2-S ; and
- 25 (b) detecting with a marker, the amount of the first complex and the second complex in the sample.
29. A diagnostic method according to claim 28, wherein the sample comprises intact cells.
30. A diagnostic method according to claim 28, wherein the sample  
30 comprises lysed cells in a lysate.
31. A diagnostic method according to claim 28, wherein the biological sample is from a human subject with a susceptibility to cancer or having the initial stages of cancer.

32. A diagnostic method according to claim 31, wherein the biological sample is from a cancer in a human subject, wherein the cancer is selected from melanoma, leukemia, astocytoma, glioblastoma, lymphoma, glioma, Hodgkins lymphoma, chronic lymphocyte leukemia and cancer of the pancreas, breast, thyroid, 5 ovary, uterus, testis, pituitary, kidney, stomach, esophagus and rectum.

33. A diagnostic method according to claim 28, wherein the biological sample is from a human fetus.

34. A diagnostic method according to claim 28, wherein the biological sample is from an adult human.

10 35. A diagnostic method according to claim 28, wherein the biological sample is selected from: a blood sample, a biopsy sample of tissue from the subject and a cell line.

36. A diagnostic method according to claim 28, wherein the biological sample is derived from heart, brain, placenta, liver, skeletal muscle, kidney, pancreas, 15 spleen, thymus, prostate, testis, uterus, small intestine, colon, peripheral blood and lymphocytes.

37. A diagnostic method according to claim 28, wherein the marker is a fluorescent marker, the fluorescent marker optionally conjugated to the FANCD2-L antibody.

20 38. A diagnostic method according to claim 28, wherein the marker is a chemiluminescent marker, the chemiluminescent marker optionally conjugated to the FANCD2-L antibody.

39. A diagnostic method according to claim 28, further comprising: binding the first and the second complex to a third antibody conjugated to a substrate.

25 40. A diagnostic method according to claim 30, wherein the lysate is subjected to a separation procedure to separate FANCD2 isoforms and the seperated isoforms are identified by determining binding to the first or the second FANCD2 antibody.

30 41. A diagnostic test for identifying a defect in the Fanconi Anemia pathway in a cell population from a subject, comprising:

selecting an antibody to FANCD2 protein and determining whether the amount of an FAND2-L isoform is reduced in the cell population compared with amounts, in a wild type cell population; such that if the amount of the FANCD2-L protein is reduced, then determining whether an amount of any of FANCA, FANCB,

FANCC, FANCD1, FANCE, FANCF or FANCG protein is altered in the cell population compared with the wild type so as to identify the defect in the Fanconi Anemia pathway in the cell population.

42. A diagnostic test according to claim 41, wherein determining the amount of an isoform relies on a separation of the FANCD2-L and FANCD2-S isoforms.

43. A diagnostic test according to claim 41, wherein the separation is achieved by gel electrophoresis.

44. A diagnostic test according to claim 41, wherein the separation is achieved by a migration binding banded test strip.

45. A screening assay for identifying a therapeutic agent, comprising: selecting a cell population in which FAND2-L is made in reduced amounts;

exposing the cell population to individual members of a library of candidate therapeutic molecules; and

identifying those individual member molecules that cause the amount of FANCD2-L to be increased in the cell population.

46. A screening assay according to claim 45, wherein the cell population is an in vitro cell population.

47. A screening assay according to claim 45, wherein the cell population is an in vivo cell population, the in vivo population being within an experimental animal, the experimental animal having a mutant FANCD2 gene.

48. A screening assay according to claim 45, wherein the experimental animal is a knock-out mouse in which the mouse FAND2 gene has been replaced by a human mutant FANCD2 gene.

49. A screening assay according to claim 45, wherein a chemical carcinogen is added to the cell population in which FANCD2 is made in reduced amounts, to determine if any member molecules can cause the amount of FANCD2-L to be increased so as to protect the cells from the harmful effects of the chemical carcinogen.

50. An experimental animal model in which the animal FANCD2 gene has been removed and optionally replaced by a nucleic acid molecule of claim 1.

51. A method for identifying in a cell sample from a subject, a mutant FANCD2 nucleotide sequence in a suspected mutant FANCD2 allele which

comprises comparing the nucleotide sequence of the suspected mutant FANCD2 allele with the wild type FANCD2 nucleotide sequence wherein a difference between the suspected mutant and the wild type sequence identifies a mutant FANCD2 nucleotide sequence in the cell sample.

5           52.    A method according to claim 51, wherein the suspected mutant allele is a germline allele.

          53.    A method according to claim 51, wherein identification of a mutant FANCD2 nucleotide sequence is diagnostic for a predisposition for a cancer in the subject.

10           54.    A method according to claim 51, wherein identification of a mutant FANCD2 nucleotide sequence is diagnostic for an increased risk of the subject bearing an offspring with Fanconi Anemia.

          55.    A method according to claim 51, wherein the suspected mutant allele is a somatic allele in a tumor type and identifying a mutant FANCD2 nucleotide  
15    sequence is diagnostic for the tumor type.

          56.    A method according to claim 51, wherein the nucleotide sequence of the wild type and the suspected mutant FANCD2 nucleotide sequence is selected from a gene, a mRNA and a cDNA made from a mRNA.

          57.    A method according to claim 51, wherein comparing the  
20    polynucleotide sequence of the suspected mutant FANCD2 allele with the wild type FANCD2 polynucleotide sequence, further comprises: selecting a FANCD2 probe which specifically hybridizes to the mutant FANCD2 nucleotide sequence, and detecting the presence of the mutant sequence by hybridization with the probe.

          58.    A method according to claim 51, wherein comparing the  
25    polynucleotide sequence of the suspected mutant FANCD2 allele with the wild type FANCD2 polynucleotide sequence, further comprises amplifying all or part of the FANCD2 gene using a set of primers specific for wild type FANCD2 DNA to produce amplified FANCD2 DNA and sequencing the FANCD2 DNA so as to identify the mutant sequence.

30           59.    A method according to claim 51, wherein the mutant FANCD2 nucleotide sequence is a germline alteration in the FANCD2 allele of the human subject, the alteration selected from the alterations set forth in Table 3.

60. A method according to claim 51, wherein the mutant FANCD2 nucleotide sequence is a somatic alteration in the FANCD2 allele of the human subject, the alteration selected from the alterations set forth in Table 3.

61. A method for diagnosing a susceptibility to cancer in a subject which  
5 comprises comparing the germline sequence of the FANCD2 gene or the sequence of its mRNA in a tissue sample from the subject with the germline sequence of the FANCD2 gene or the sequence of its mRNA wherein an alteration in the germline sequence of the FANCD2 gene or the sequence of its mRNA of the subject indicates the susceptibility to the cancer.

10 62. A method according to claim 61, wherein an alteration is detected in a regulatory region of the FANCD2 gene.

63. A method according to claim 61, wherein the detection in the alteration in the germline sequence is determined by an assay selected from the group consisting of (a) observing shifts in electrophoretic mobility of single-stranded DNA on non-  
15 denaturing polyacrylamide gels, (b) hybridizing a FANCD2 gene probe to genomic DNA isolated from the tissue sample, (c) hybridizing an allele-specific probe to genomic DNA of the tissue sample, (d) amplifying all or part of the FANCD2 gene from the tissue sample to produce an amplified sequence and sequencing the amplified sequence, (e) amplifying all or part of the FANCD2 gene from the tissue  
20 sample using primers for a specific FANCD2 mutant allele, (f) molecular cloning all or part of the FANCD2 gene from the tissue sample to produce a cloned sequence and sequencing the cloned sequence, (g) identifying a mismatch between (i) a FANCD2 gene or a FANCD2 mRNA isolated from the tissue sample, and (ii) a nucleic acid probe complementary to the human wild-type FANCD2 gene sequence, when  
25 molecules (i) and (ii) are hybridized to each other to form a duplex, (h) amplification of FANCD2 gene sequences in the tissue sample and hybridization of the amplified sequences to nucleic acid probes which comprise wild-type FANCD2 gene sequences, (I) amplification of FANCD2 gene sequences in the tissue sample and hybridization of the amplified sequences to nucleic acid probes which comprise mutant FANCD2  
30 gene sequences, (j) screening for a deletion mutation in the tissue sample, (k) screening for a point mutation in the tissue sample, (l) screening for an insertion mutation in the tissue sample, (m) in situ hybridization of the FANCD2 gene of the tissue sample with nucleic acid probes which comprise the FANCD2 gene.

64. A method of diagnosing a susceptibility for cancer in a subject, comprising:

(a) accessing genetic material from the subject so as to determine defective DNA repair;

5 (b) determining the presence of mutations in a set of genes, the set comprising FAND2 and at least one of FANCA, FANCB, FANCC, FANCD1, FANCD2, FANDE, FANDF, FANDG, BRACA1 and ATM; and

(c) diagnosing susceptibility for cancer from the presence of mutations in the set of genes.

10 65. A method for detecting a mutation in a neoplastic lesion at the FANCD2 gene in a human subject which comprises:

comparing the sequence of the FANCD2 gene or the sequence of its mRNA in a tissue sample from a lesion of the subject with the sequence of the wild-type FANCD2 gene or the sequence of its mRNA, wherein an alteration in the  
15 sequence of the FANCD2 gene or the sequence of its mRNA of the subject indicates a mutation at the FANCD2 gene of the neoplastic lesion.

66. A method according to claim 65, further comprising :  
determining a therapeutic protocol for treating the neoplastic lesion according to the mutation at the FANCD2 gene of the neoplastic lesion.

20 67. A method for confirming the lack of a FANCD2 mutation in a neoplastic lesion from a human subject which comprises comparing the sequence of the FANCD2 gene or the sequence of its mRNA in a tissue sample from a lesion of said subject with the sequence of the wild-type FANCD2 gene or the sequence of its RNA, wherein the presence of the wild-type sequence in the tissue sample indicates  
25 the lack of a mutation at the FANCD2 gene.

68. A method for determining a therapeutic protocol for a subject having a cancer, comprising:

(a) determining if a deficiency in FANCD2-L occurs in a cell sample from the subject by measuring FANCD2 isoforms according to claim 25;

30 (b) if a deficiency is detected in (a) , then determining whether the deficiency is a result of genetic defect in non-cancer cells; and

(c) if (b) is positive, reducing the use of a therapeutic protocol that causes increased DNA damage so as to protect normal tissue in the subject and if (b) is negative, and the deficiency is contained within a genetic defect in cancer cells

only, then increasing the use of a therapeutic protocol that causes increased DNA damage so as to adversely affect the cancer cells.

69. A method of treating a FA pathway defect in a cell target, comprising: administering an effective amount of FANCD2 protein or an exogenous nucleic acid  
5 to the target.

70. A method according to claim 69, wherein the FA pathway defect is a defective FANCD2 gene and the exogenous nucleic acid vector further comprises introducing a vector according to claim 10.

71. A method according to claim 69, wherein the vector is selected from a  
10 mutant herpesvirus, a E1/E4 deleted recombinant adenovirus, a mutant retrovirus, the viral vector being defective in respect of a viral gene essential for production of infectious new virus particles.

72. A method according to claim 69, wherein the vector is contained in a lipid micelle.

73. A method for treating a patient with a defective FANCD2 gene,  
15 comprising:  
providing a polypeptide described in SEQ ID No: 4, for functionally correcting a defect arising from a condition arising from the defective FANCD2 gene.

74. A cell based assay for detecting a FA pathway defect, comprising:  
20 (a) obtaining a cell sample from a subject;  
(b) exposing the cell sample to DNA damaging agents; and  
(c) detecting whether FANCD2-L is upregulated, the absence of upregulation being indicative of the FA pathway defect

75. A cell based assay according to claim 74, wherein amounts of  
25 FANCD2 are measured by an analysis technique selected from: immunoblotting for detecting nuclear foci; Western blots to detect amounts of FANCD2 isoforms and quantifying mRNA by hybridising with DNA probes.

76. A kit for use in detecting a cancer cell in a biological sample,  
comprising:  
30 (a) a primer pair which binds under high stringency conditions to a sequence in the FANCD2 gene, the primer pair being selected to specifically amplify an altered nucleic acid sequence described in Table 7; and  
(b) containers for each of the primers.

**(57) Abstract:** Methods and compositions for the diagnosis of cancer susceptibilities, defective DNA repair mechanisms and treatments thereof are provided. Among sequences provided here, the FANCD2 gene has been identified, mapped on the 3p chromosome, cloned into recombinant vectors, used to prepare recombinant cells and sequenced. The FANCD2 gene sequence provides probes and primers for screening patients in genetic based test and for diagnosing Fanconi anemia and cancer. It has also been possible to target the FANCD2 gene *in vivo* for preparing experimental mouse models for use in screening new therapeutic agents for treating conditions involving defective DNA repair. Vectors are described for use in gene therapy. The FANCD2 polypeptide has been sequenced and has been shown to exist in two isoforms identified as FANCD2-S and the mono-ubiquitinated FANCD-L form. Antibodies including polyclonal and monoclonal antibodies have been prepared that distinguish the two isoforms and have been used in diagnostic tests to determine whether a subject has an intact FA pathway. The FANCD2 has been localized to the nucleus and is associated with BRCA 1 foci.

Representative Drawing  
Fig.8

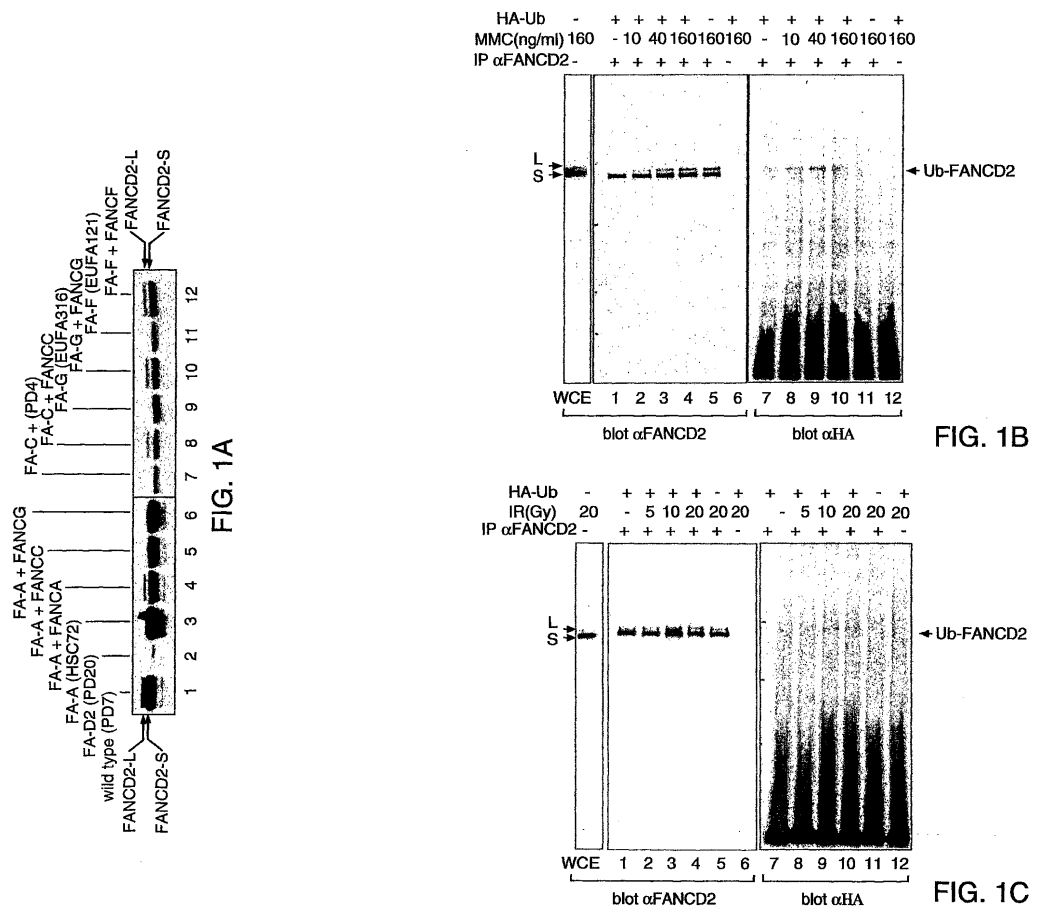


FIG. 1A

FIG. 1B

FIG. 1C

Activation of Ubiquitinated FANCD2 in Corrected FA cells

4/22

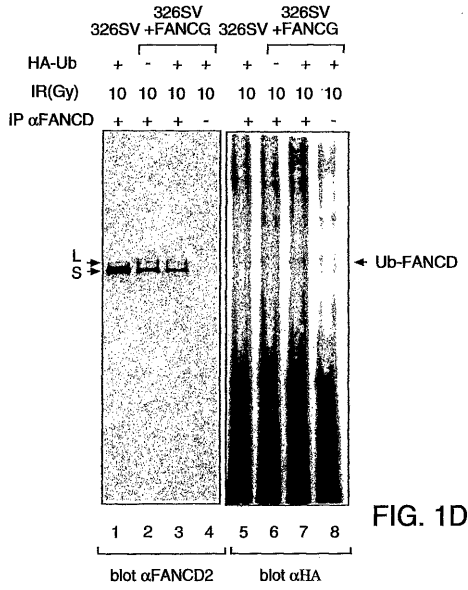


FIG. 1D

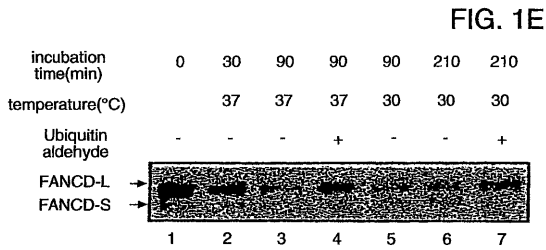


FIG. 1E

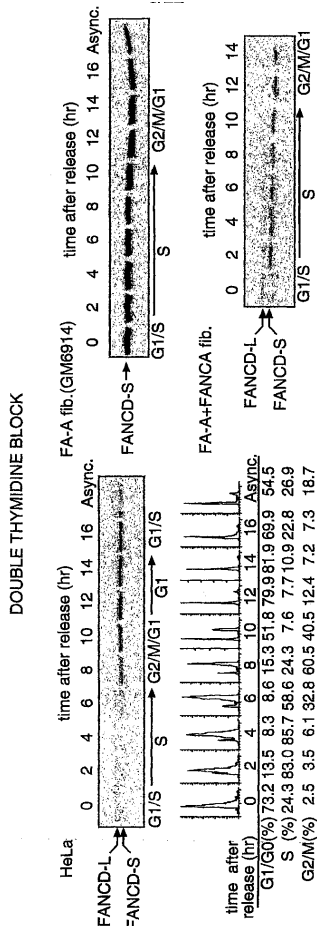


FIG. 3A

**NOCODAZOLE BLOCK**

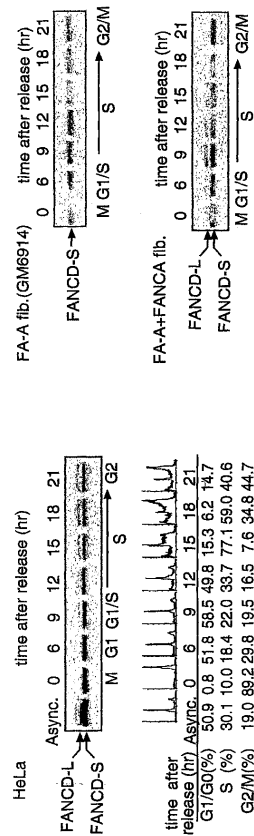


FIG. 3B

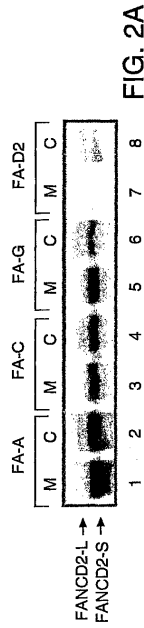


FIG. 2A

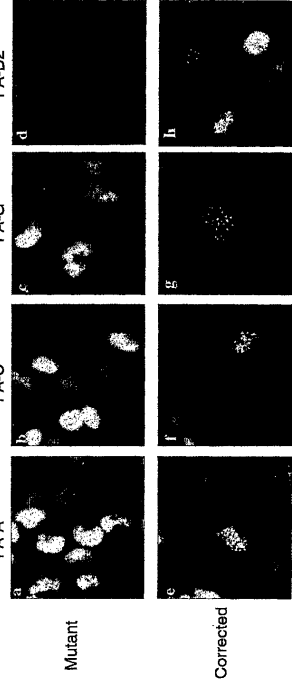


FIG. 2B

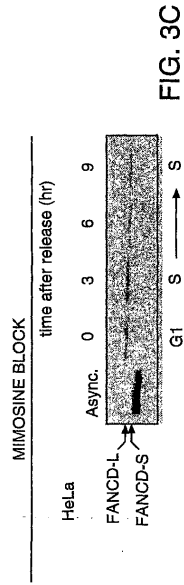


FIG. 3C

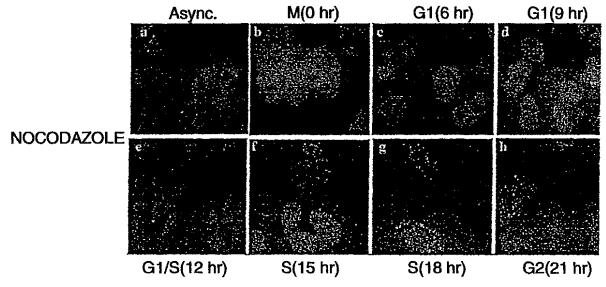


FIG. 3D

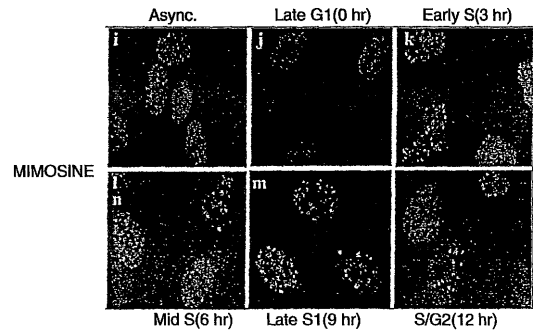


FIG. 3E

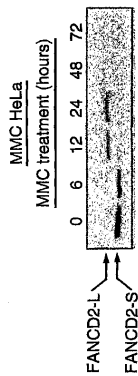


FIG. 4A

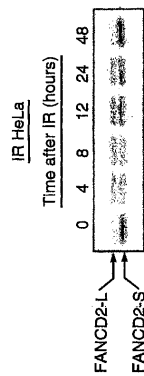


FIG. 4B

8/22

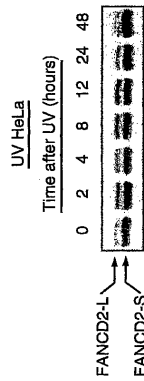
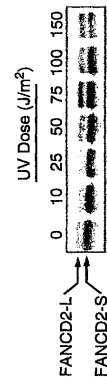


FIG. 4C



9/22

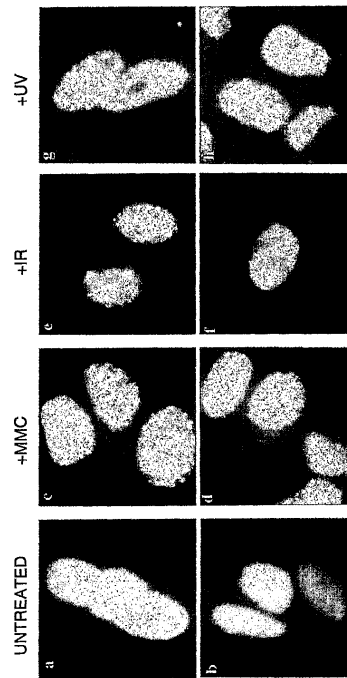


FIG. 4D

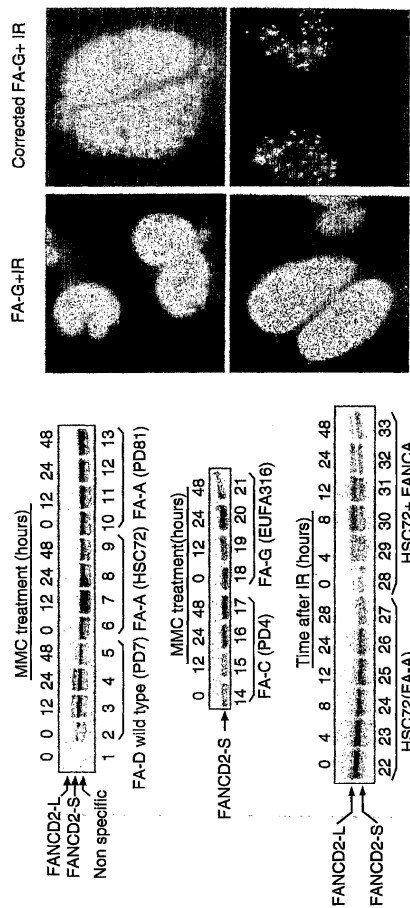


FIG. 4E

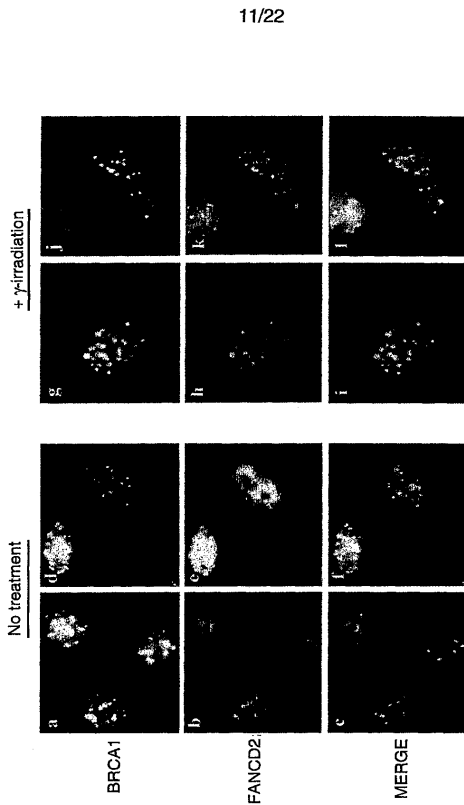


FIG. 5A

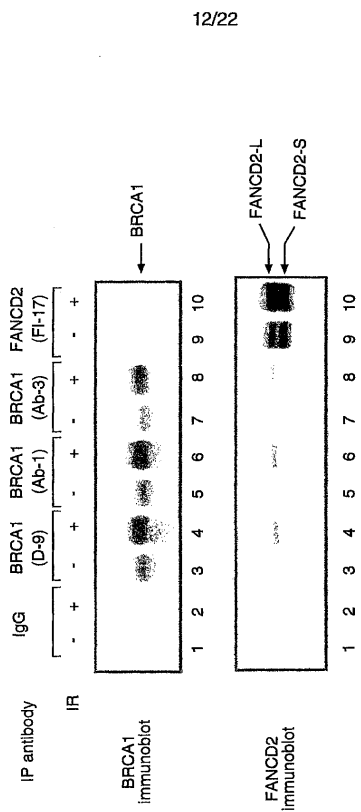


FIG. 5B

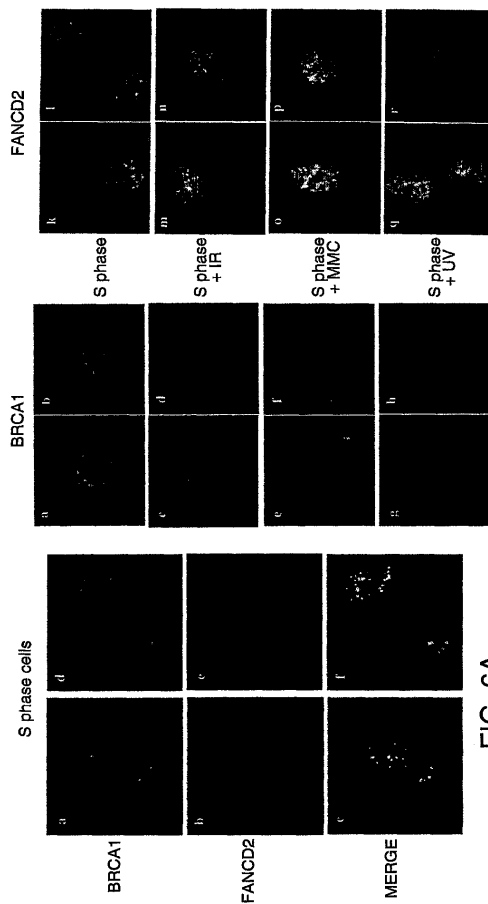


FIG. 6B

FIG. 6A

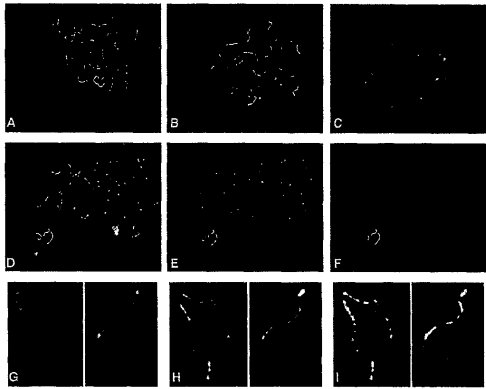


FIG. 7

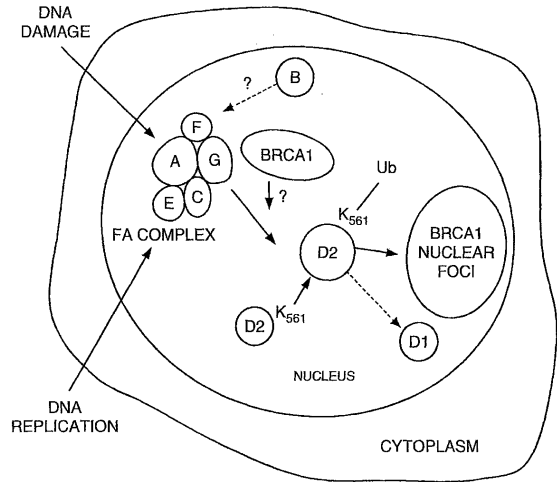


FIG. 8

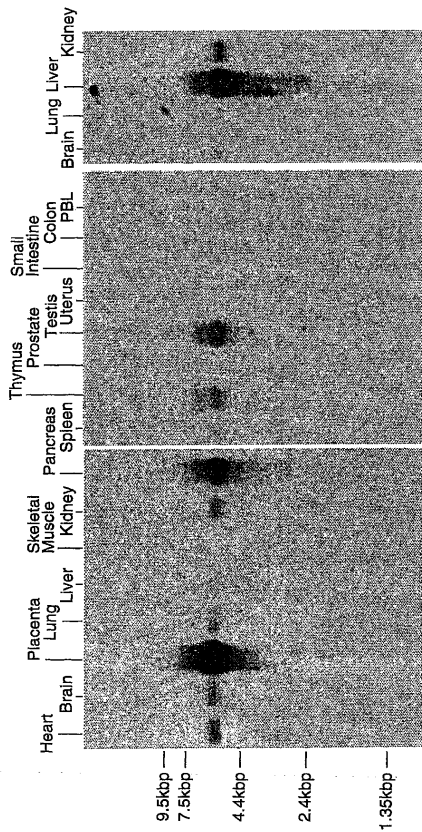


FIG. 9

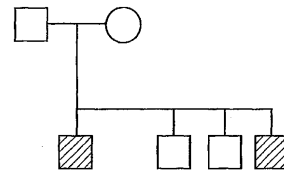


FIG. 10A

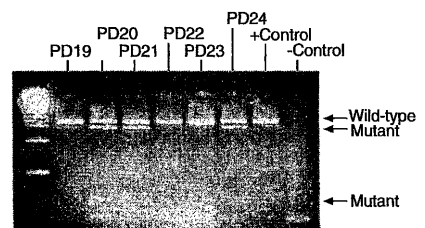


FIG. 10B

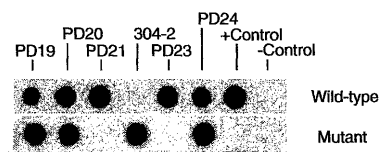


FIG. 10C

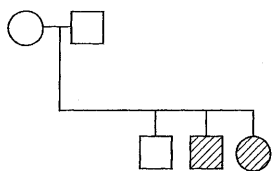


FIG. 10D

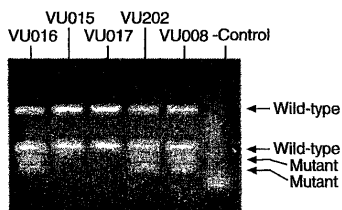


FIG. 10E

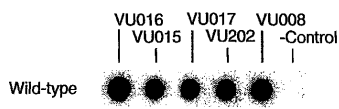


FIG. 10F

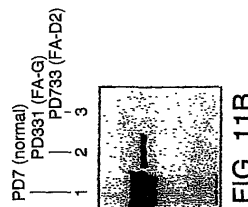


FIG. 11B



FIG. 11C

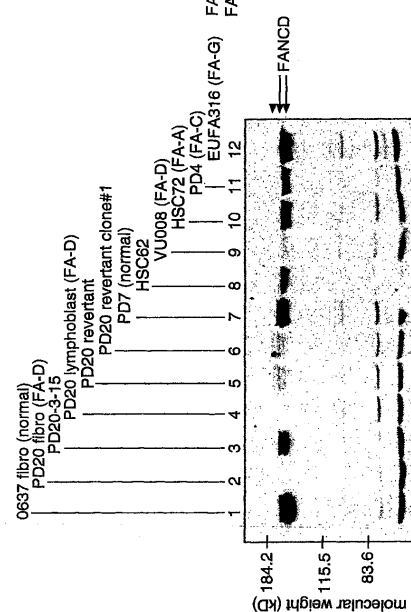


FIG. 11A



FIG. 12A

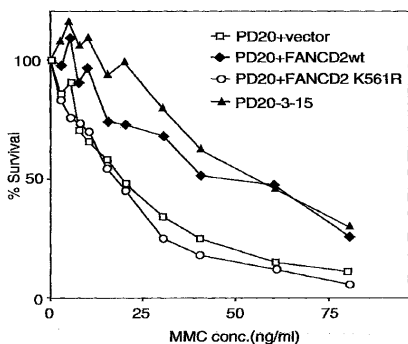


FIG. 12B

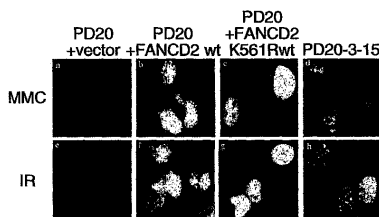


FIG. 12C

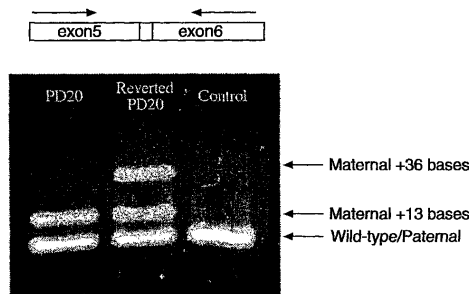


FIG. 13A

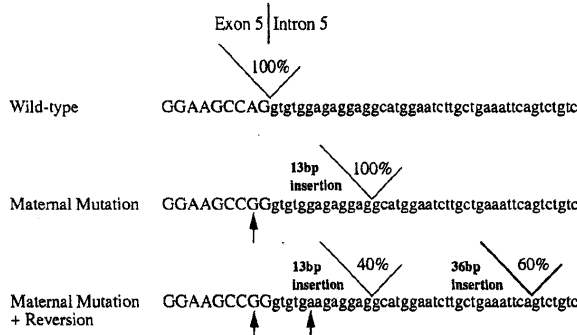
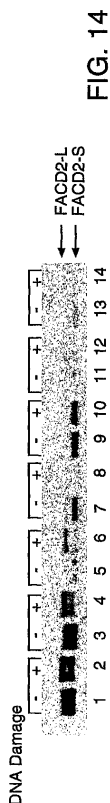


FIG. 13B



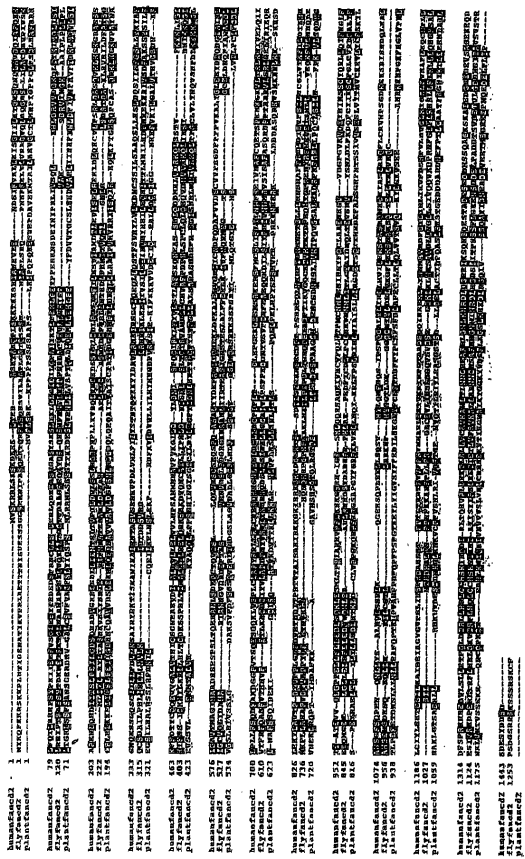
SEQ ID NOS: 4-5

FANCD2 cDNA Sequence

1 ATG GTC TCC AAA AGA GCG TCA AAA CTC GAG GAT AAA GAG ACC CTG ACA GAA GAT GGC 60  
 2 M V S K R R K L D S K I F V K L L D K I S G S I L K 20  
 61 TCC AAC ACC GAG GAG GAA GCA GCA CTC TCC AAG TCG AAA GAT GAT GAT GAT GAT 120  
 21 S K F K R K Q F L S R K K T R K K S R I A N E 40  
 121 GTC GAA GAA AAT GAG ACC ATC TTC GGA AAG CTT TTC AAG ABA GCA GCA ATT ATT ATT AAA 180  
 41 V E E N F D S I F V K L L D K I S G S I L K 60  
 181 ACC GGA GAG AAT GAG ACC ATC TTC GGA AAG CTT TTC AAG ABA GCA GCA ATT ATT ATT AAA 240  
 61 T E E N F D S I F V K L L D K I S G S I L K 80  
 241 GAG ACC CTG ACC GAG ACC CTC GAT GCG AAA ABA ABA GAA GAA TTC GTC GAT GAT GAT 300  
 81 Q P L E R K E F S Y P R I I E N E F V S G D L 100  
 301 GAG GTC TAC AAT GAG GAG GAA GAG AAT TTC ACC ACC GCG CTC TTC GTC GAT GAT GAT 360  
 101 E Y F E D R E D R K H C L D E C S R L 120  
 361 GAG GAT GAG GAA GCG ACC ACC GCG GCA GCG TTC GAG ACC ACC ACC ACC ACC ACC ACC 420  
 121 C D B E A M G A H G A I I K S L F E H L P E B Y 140  
 421 GTC GAT GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 480  
 141 Q Z D I L G F A I I K S L F E H L P E B Y 160  
 481 TTT TTA AAC AAG AAC AAG AAC GAA AAC AAC AAC AAC AAC AAC AAC AAC AAC AAC 540  
 161 F T E N N K N S D E I N I Z P K L I V S Q L 180  
 541 AAA TGG CTT GAC ACA GAT GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 600  
 181 K W L D C R V Y D G K D L P Q K I H G L Z 200  
 601 AAT TTT GTC GCA GAG AAG CTT GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 660  
 201 S I A F E M L D Q K D I T T S S L P S Z L G C 220  
 661 GAT TCC GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 720  
 221 D F G S A P Y G K E L S D L D E I E H T S 240  
 721 GTC ACC GTC GCA GAT GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 780  
 241 L D V F R L D V D F S E R K L D F K P L L 260  
 781 AAG GTC GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 840  
 261 K V R Q L H N G H L E B E I D G L D L C F I 280  
 841 ATA ATA AAG AAT GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC 900  
 281 I K P I L H S V P A H D T L S V I S E 300  
 901 GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC 960  
 301 L R B K L D L Q H G T V T P CCA TCA GTC TCA GTC TCA GTC TCA GTC TCA GTC TCA 320  
 961 GTC AAG TCC AAA ACC AAA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA 1020  
 321 V K L X S K G R A S S S G N Q E S T P H G T 340  
 1021 ACC TCC ACT ATT CTC CTC TTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC 1080  
 341 S C I L L D V I K A I R Y E K T I 360  
 1081 TCA GAA GCG TGG AAT GAG AAT GAG AAG ACC GCG TCA GTC GTC GTC GTC GTC GTC GTC 1140  
 361 S E A W I K A I E N T A S V S E H K V I 380  
 1141 GAG CTT GTC AAT GTC TTC TTC ATC GAG ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC 1200  
 381 D L V H L F I I Z Y S T M T Q T K K Y I D 400  
 1201 ACC GTC GCA AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG 1260  
 401 R V L Z N K I R S O C I Z Q L L Q L S T 420  
 1261 TTC TCC GAT GAT TTA GTC AAT GAG ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC 1320  
 421 F S V K Y L V L K D H C S S I L S L A Q 440

Figure 16

SEQ ID NOS: 1-3



1321 AGT TTT CTT CAC CTC CTA GAC CAG ACC ACA ATC TCA TTT GGC AAT CTC GCA GAC AAA TWT 1380  
 441 S L L H S L D Q S K I Z S L L V K V 460  
 1381 GCA TTT AAT TTT TTT GAC ACC TAC CTC GAG GAG GAA GAA GTC GTC GTC TTA GTC ACC GAT 1440  
 461 A F K F F D T Y C Q G A L V G A L V G A L V G A L V G A L V G A L V G A L V G A L V 480  
 1441 ACC TCC AAT GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 1500  
 481 I C S G M B A R Y D T A L D D Y L 500  
 1501 GTC TTA ACC CCA TCC ATG ATG ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC 1560  
 501 V L N F S A H M M N A C V TTT TTA AAG GCG AAT TTA GAT TAT 1580  
 501 GTC GAT AAC ABA TCC CCG CAG GCA AFA GAA GCA CTC TTC TTT TTT TTT TTT TTT TTT 1640  
 521 L D M I S F Q Q I R K L L G T Y Y V G A L S G L L A 540  
 1621 TTT ACC AAA CAG AAT GAA GCG ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC 1680  
 541 T S K Q H E A S S H I Q C G A G A C G A C G A C G A C G A C G A C G A C G A C G A C 560  
 1681 AAG CAG CTC TCT ACC ACC GTA TTC AAG TAC AAG CTC ATG GAG TAC ACC GTC GTC GTC ACC 1740  
 561 K Q L S B T Y P K Y K L L G I Z G A V G 580  
 1741 ATG GTC GTC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC 1800  
 581 M A G Z H A A D R S B S B S L T Q B R A 600  
 1801 AAG CTC AAT GAT GAG CAG TCC ACA CAG ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC 1860  
 601 M L S D E Q Q Q V S C V W F R G D F F F F F 620  
 1861 AAT GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 1920  
 621 S E Q S F Q A S A L Y Y D B F A M L I Q 640  
 1921 GAT GAA AAG GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT 1980  
 641 M E K D P K A L G H V G G K E I C H F P 660  
 1981 GAG GAT GCG TTC GCA GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC 2040  
 661 Q S F F V D S C V W F R G D F F F F F 680  
 2041 AAA GCA GTC TAC GCA GTC GCA GAA GAC GAC GAT GAT GAT GAT GAT GAT GAT GAT GAT 2100  
 681 K A L V G L S E S Y D G D G S A Z T T L 700  
 2101 GTC CTT CTT TTT TTT GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG 2160  
 701 P L L F S Q D P F A K D G G O P V T S G S S 720  
 2161 GTC GCA AAA TTT GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC 2220  
 721 G Q K L V S P L C L A S Y F R L L L L C 740  
 2221 GTC GAG AAG CAG CAG AAG GCA AAC TGG GAG GAG ATT GAT GAT GAT GAT GAT GAT GAT 2280  
 741 V E R Q H N G H L E B E I D G L D L C F I 760  
 2281 TTC GCA ACT GAT CTT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT 2340  
 761 P L T D L E F G R K L E S M S A K B R S 780  
 2341 TCC AAT TTT CTT CTT CTT CTT CTT CTT CTT CTT CTT CTT CTT CTT CTT CTT CTT CTT 2400  
 781 F H C S L I F L T L H W F R R I V H A F 800  
 2401 TGG CAG GAA ACA TCA GTC GAT AAG AAG GTC GTC GTC GTC GTC GTC GTC GTC GTC GTC 2460  
 801 C E S S F E M R C K V L S R L R H I V 820  
 2461 GAA TCC GAA ABA ACC CTT GAA AAG TAC TCA GTC ACC ACC ACC ACC ACC ACC ACC ACC 2520  
 821 R L G I L B K Y L A V T P D G S F F L 840  
 2521 GAA AAC TTT GAT GAA ACT TTA GAT AFA ACA CCG GAT ACT GTC ACT GTC ACT GTC ACT 2580  
 841 G N F D V B T L D I E T S P H T V A I S A 860  
 2581 AAA AAT AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG 2640  
 861 K I R K K G X I E R A G A A C A A C A G T G G C A C A C A C A C A C A C A C A C A 2660  
 2641 TCC TCT GAG ACA CTT TCA GAG AAG AAT TCA GAA TGG GAT ACC ACC ACC ACC ACC ACC 2700  
 861 S D T S B K W S O C D F T F S H K 900  
 2701 GGC CAG CTA AAC AAG GTC ACA GAG AAG GAA GAA AAG ACA TCA TGG TTA CCA CAT AAT 2760  
 901 G L N K E F P Q K T S L K T S L N H F 920  
 2761 TCC GAT GTC TCA GAG CTT GAG GTC TCC ATC ATC ATC ATC ATC ATC ATC ATC ATC ATC 2820  
 921 S H A F P R E L D I E V F S I G L C T G L 940

Figure 15



8/14

9/14

GGTTACTGAGACTTTGTGTGGAGAGACAGCATAACGGAAACTTGGAGGAGATTGA  
 TGGTCTACTAGATTGTCTATATTCTCACTGACCTGGAGCCTGGAGGAGAGATTGG  
 AGTCCATGTCTGCTAAAGAGCGTTCATTCATGTGTTCTCTCATATTCTTACITCA  
 ACTGGTCCGAGAGATTGTAATGCCTTCTGCCAGGAAACATCACCTGAGATGAA  
 GGGGAAGGTGCTCACTCGGTAAAGCAGCATTTGTAATGCAAAATTAATCTGGAA  
 AAGTACTTGGCAGTACCCCAAGACTATGTCCCTTCTTGGAAACTTTGATGTGGA  
 AACTTTAGATATAACACCTCATACTGTTACTGCTATTTCAGCAAAAATCAGAAA  
 AAAGGAAAAATGAAAAGGAAAAACAAAACAGATGGCAGCAAGACATCTCTCTC  
 TGACACACTTTCAGAAAGAGAAAATTCAGAAATGTGACCTTCCGATCTCATAGA  
 GGGCAGCTAAACAAGGAGTTCACAGGGAAGGAAAGAAAAGACATCTTGTACTA  
 CATAATTCCCATGCTTTTTCCGAGAGCTGGACATTTGAGGTTCTCTATTCTACAT  
 TGTGGACTTGTGACGAAATTCATCTTAGATACTGAAATGCACACTGAAGCTACAG  
 AAGTGTGCAACTTGGGCCCTGAGCTGCTTTCTTGGTGGAAAGATCTCTCCAG  
 AAGCTGGAGATGTGTGACACTCTATTGCCAGGAGAGTCCCTTTTCAAGA  
 ACAAAAGGAAAGCCGGAATTTGGATTCTCACATCTCCAACAGAGATCTGCCAAG  
 AAAATGTTCAITTTGTTTTCAACTGCTGACCCAAATGTGTAACCACTGGAGAAC  
 ATTCACAACATTTTCAAGTGTAGCTGCTGAGAATCAGGTTGATGTTGATGGAC  
 AGGAGTGAAGTTACAGGAGTACCACATAATGTCTCTGCTATCAGAGGCTGCTG  
 CAGATTTTTCATGGGCTTTTTGCTTGGAGTGGATTTTCTCAACTGAAAAACAGAA  
 TTAATGTTTACAGCCCTCCATGTCTTAGTAGCCGACTGAAACAGGGAAGACACA  
 GCCAGCTTTGGAGAACTACTACAGCAGAGCTCCATTAATGTCAGAAATTTCCAT  
 CAAAAGCATTTCCAGTTTCCAGTGTGCTTTTATCTCATCAGACTTTTGTGGTATT  
 TTGGAGAAATCAACAGCTTCTGCTCAGAACAAAGAAAAATTTGCTTCCCTGGCA  
 GACAAATCTCTGCTGGGTGTGGCCAAAGTGGGATAAAGAGAAAGAGCAACATCTC  
 TAATGACCAAGCTCCATGCTCTGCTGTATCTACCTGGAGCACACAGAGCATTTC  
 TGAAGGCCATAGAGGAGATTGCTGTGTGGTGTCCAGAACTGATCAACTCTCC  
 TAAAGATGCATCTTCTCCACATTCCTACACTGACCCAGGCATACTTTGTGTTTT  
 CTTCGTTGATGATGGTGAAGTACAGAAAGAGCGGTGAAAAAATTTGAGCCTGGC  
 ACAGCCAGACTCGCACAGCAGATTCAAGAGAAACTCTCTACTGGAACATG  
 GCTGTTGAGACTTCAATCTCATCACTGATAAAGGTTATTGATAGTCACTC  
 TGTCTGATGATGTTGAAAGTATGGCGTCTCTTTGGAAGCATTTCGAAAGCA  
 ATGTATGCGCTCCTAGACTTCACTTTTGAAGAAACCCGGAAAGTGTCTGAGCT  
 TACTGGAAACCTTCCAGTTGGACCAAGGCTGCTCATCACTGTGTGGCATTCC  
 AAGATTCCACAGGACACGAGACTCACCCAAACATGTGCTCTGCTCAAAAGACC  
 CTGGAACCTTATGTTGACAGTCAAAGCTATGCTCACTCAACAATTTGATAGA  
 GGCTTTCTGGCTGGCAACTTAAAAAACCAGGACTTGCAGGTTGAAGAGATTAA  
 TCCAAAATTTCCAGGAGACACAGCAGATGAGAGTGGAGTACATGTCATCC  
 CAGGCTTCCAGAGCAAGCCACTGAGGTATCTCTCAAAAACCCAGAGTCT  
 GGCACTGATGGTTGCAATTTGTTAAATGTTCTAAGTTGGTGGAGCAGAACTTTCCT  
 ACITATGTTTATGCTCAAATGCTTCTATGCCCAATTTCCATCTCCCTCAATAACAGCT  
 CTGTGCTATATAATTTTGGGACCCAGAAAGAAACACGACACAATTTTGAATC  
 ACTCTGATATCTCGAGTTGTGGCATTGTTATAGAGTTGACAAATTTCTGCAATTA  
 TAGCCCTCACTTTTCCATGAATTCATATCTGAAACCAATTTTGAAGGGAAGTCA  
 TCGAAGTATTTTCTGAGTGTGAGAAGATGAGTTAAACCAATTTAAACACATTTGA

AAACATACAAAAATAGAAATGTGAAAGCAATTTGGTGAAGCCAAAGCACAGAGTC  
 AGAAGCTGCCACCTTAGAGAACTGAAATAAAAAATAGAAGTCTTACGCTTTTTTGT  
 GGTACAGATGCTTTTCGCAATTTAAAGAAAGCTAAATAAAAAATGAGACATGGCT  
 GGCCGAGTGGCTCATGCTTGAATCCTAGCACTTTTGGAGCCAAAGGTAGGAGGA  
 TTGCTTGAAGTCCGGGAGCTCAAGGCAAGCTGCAACATAAACAAGACCCCTATCT  
 CCACAAAAAATGAAAAATAAACTGGGTGGCTGCTCACCTCTGATATCC  
 AGCACTTTGGAGGCCGATGTGGCAGATCAAGAGTTCAGAGTTCAAGACCAG  
 CTTGGCCAAACATAGTGAACCCCTCTCTACTGAAATAAAAAATTTAGCTGGGT  
 GTGGTGGCAGTGCCTTATCTCAGTACTTGGAAAGCTGA

2

SEQ ID NO: 7

11/14

Human FANCD2 short form

TCGAAAACTACGGCGCGCAGCGCTTCTCGGAAGTAATTAAGTGCACAAGACAT  
 TGGTCAAATGGTTTCCAAAAGAGACTGTCAAATCTGAGGATAAAGAGAGCC  
 TGACAGAAATGCCCTCCAAAACCGAAACCACTTCCAAAAGACAAAGA  
 AATCTCATATTGCTAATGAAGTTGAAGAAAATGACAGCATCTTTGTAAGGCTTCT  
 AAGATATCAGGAATTTCTTAAAACCGGAGAGAGTCAAGTCACTAGCTGTGGT  
 ATCAAATAGCTTTCCAAAAGAGCTCTTTAGACCTGAGGAGACACCTTCTTA  
 TCCAAAATATAAGAAATTTGTTAGTGGCTGGAGTCTTACATGAGGATGAA  
 GACAGTTTCAGGAACCTGCTTTTGTCTTGTGAGCGTCTGAGGATGAGAAAGCA  
 GTATGGGTGCACTTATTCTAAGAGTCTCATCAAATGCTTCTGGGAAITGACATA  
 CTGACGCTGCCATTAACAAAACCTTATTGGAAGTGTCCAGAAATATTTTTGGA  
 AAAACAAGAACAGTATGAAATCAACATACCTCGACTATTGTCAGTCAACTAAA  
 ATGGCTTGACAGAGTTGTGGATGGCAAGGACCTCACCAAGATCATGACGCTG  
 ATCAGATTGCTCCAGAGAACCTGCAGCATACATCACAGCCTACCTGAGA  
 TCTTAGGGGATTTCCAGCAGCTGATGTGGGAAAGAACTCAGTACCTACTGAT  
 AGAAATCTTCACTCACTGCTCCAACTCGAATGCTTCTTCAAGCTCCGACTTG  
 ACCAAAATCTCTATGAAGTTTCGCGAGTGTGTGATAAGTTGTGCTTATT  
 AGATTGGAGGATTAACCTGTGATAATAAAGTTCATTTCTTCCGTAACAGCCAT  
 GGATACATITGAGTAAATTTTCTGAGCTTCCGGGAGAAAGTTGGATCTGCAAGTTCTG  
 TTTTCCATCAGGTTACAGGCTTCCCAAGTAAAGTTGAAAGTAAAGGACGAGC  
 AAGTCTCAGGAAATCAAGAAAGCAGCGCTCAGAGCTGATTAATTTCTCTTTG  
 ATGATAAAGTCAGCTATTAGATGATGAAACCAATTCAGAGCTTGGATTAA  
 GCAATTTGAAACATGCTCAGTATCTGAAACAAAGGTTGTTGACCTGGTGAITG  
 CTTTTCATCTATAGCAACCAATACTCAGACAAAGAGTACATTGACAGGCTG  
 TAAAGAAATAAGATTTCGATCAGCTGCAATTAAGAACAGCTGCTCCAGATACATT  
 CTCTGTCTACTTACTAGTTCTTAAGGATATGTGTTATCCATTTCTGCTGGCTCA  
 GAGTTGTCTCACTCTAGACAGAGATATAATTTCAATTTGGCAGTCTCTATACA  
 AATATGCAATTAAGTTTTTTCAGCAGTACTGCGCAGGAAAGTGGTTGGTGCCTTA  
 GTGACCCATATCTGAGTGGGAAAGTGAAGTGAAGTTGATCTGCTTATGATGCTCT  
 TCTAGAGTTGGTGAAGTTAAACCCATCTGCTATGATGATGAATGCTGCTTTGTA  
 AGGCAATTTAGATTATCTGGATAACATATCTCCCTCAGCAAATACGAAAACCTCTT  
 TATGTTCTCAGCACACTGGCATTTAGAAAACAGAAATGAAGCCAGGACCCATCC  
 AGGATGACATGCACTTGGTGTATAAGAAAGCAGCTCTCTAGCACCTGATTTCAAGTA  
 CAAGCTCATTTGGATATTGTGCTGTGACCACTGGCTGACATGAGCCGCGAC  
 AGAATGAAATCACTAGTTTGAACCAAGAGAGCAACCTGAGCAGTACAGCAG  
 TGCACACAGGAGCTCTTGTGCAAGTGGTTCATCTGCACTGAGCAGCTCTC  
 TCAGGCTCTGCACTTACTATGATGAATTTGCCAACCTGATCCAACATGAAAAGC  
 TGGATCAAAAAGCCCTGAAATGGGTTGGGCAATCCATCTGTAATGATTTCAAGGA  
 TGCTTCTGATGCACTCTGTTGTTCCGGAAGGTCATTTCCATTTCTGTGAA  
 AGCACTGACGACTGGAAGAATAACAGCACTCAGGATGGGATGGCAATAACCT  
 CCTGCCGCTGCTTTTTCTCAGACTTTGCAAAGATGGGGTCCGGTACCTCAC  
 AGGAATCAGCCAAAATTTGGTGTCTCCGCTGTGCTGGCTCCGATTTCCGGTTA

CTGAGACTTTGTGTGGAGAGACAGCATAACGGAAACTTGGAGGAGATTGATGGTC  
 TACTAGATTGCTCTATATTCCTAAGTACCTGGAGCCTGGAGAGAAAGTTGGAGTCC  
 ATGTCTGCTAAAAGAGCCTTCACTCATGTGTTCTCTCATATTTCTACTTCAACTGG  
 TCCGAGAGATTGTAATGCCTTCTGCCAGGAAACATCACTGAGATGAAGGGGA  
 AGGTCTCACTCGGTTAAAGCACAATGTGAAATGCAAAATACTCTGAAAAAGTA  
 CTTGGCAGTCAACCCAGACTTCTCTCTGTTACTGCTATTTTCAGCAAAAATCAGAAAGAAAGG  
 AAAAAATAGAAAGGAAACAAAAACAGATGTCAGCAAGACATCTCTCTGACAC  
 CACTTTCCAGAAAGAAAAAATTCAGAAATGTCACCTACCCCTCATATAGAGGCCA  
 GCTAAACAGGAGTTACAGGGAAGGAAAGAAAGACATCATTTGTTACTACATAA  
 TTCCATGCTTTTTTCCGAGAGCTGGACATTTGAGGCTTCTCTATTCTACATTTGG  
 ACTTGTGACGAAGTTCATCTTAGATACTGAAATGCACACTGAAGCTACAGAAAT  
 GTGCAACTTGGGCCCTGAGCTGCTTTTCTGCTGGAAGATCTCTCCAGAAAGCT  
 GGAGAGATGCTGACACTCTTATTGCCAGGAGAGTCCCTTTCTCAAGAAACAAA  
 GGAAGCCGGAATATTGGAATCTCACATCTCCAACAGAGATCTGCCCAAGAAATG  
 TTCATTTGTTTTTCAACTGCTGACCCAAATGTGAACCACTGGAGAAACATTCA  
 AACTATTTGAGTGTGAGTGTGAGAAATCAGGTTAGTTGATGGACAGGAGT  
 GAAAGTTCAGGAGTACCAATAATGCTCTGCTATCAGAGGCTGCTGCAATTT  
 TTCATGGCTTTTTGCTTGGAGTGGATTTTCTCAACTGAAAATCAGAATTTACTGT  
 ATTCAGCCTTCAATGCTTATGAGCCGACTGAAACAGGGAAGACACAGCCAGCC  
 TTTGGAGAACTACTGACCCAGAGCTGCAATTTGCAAAATTTCCATCAAAAGC  
 ATTTCCAGTTTCCAGTGTGCTTTTATCTCATCAGACTTTTGTGTTATTTGGAG  
 AAATCAACAGCTTCTGCTCAGAAACAAAGAAAAAATGCTTCCCTGGCAGACAAT  
 TCTCTGCTGGGTGTGGCCAAAGTGGGATAAAGAGCAAGAGCAACATCTCTAATGA  
 CCAGCTTCTGCTCTGCTGTAICTACTCTGGAACACAGAGACTTCTGAAG  
 GCCATAGAGGAGATTGCTGTTGGTGTCCAGAACTGATCAACTCTCTAAAG  
 ATGATCTTCTCCACATTTCCCTACACTGACCCAGGCACTTTGTTGTTTCTCC  
 GTGTGATGATGGCTGAACTAGAGAAAGCAGGTAAGAAAAATGAGCCTGGCCAG  
 CAGCAGACTCGCAGAGATTCATGAAGAGAAACTCTCTACTGGAACATGGCTGT  
 TCGAGACTTCAGATCTCTCATCACTTGTATAAAGGTTATTGATGATCATCTGTTCT  
 GCATGCTTTGTAAGTATGGGCTGCTTTTGGTGAAGCAATTTCTGAGCAATGTA  
 TCCGCTCTGAGCTTCACTTGTATGAAACACCGGAAAGTGTCTGAGCTTACTG  
 GAAACCTTCAAGTTGGACAAAGGCTGCTTCACTCACTGTGTGGCATTCCAAGA  
 TCCACAGGACAGGACTACCCCAACATGTGCTCTGCTCAAAAAGACCCCTGGA  
 ACTTTTGTGTTGACAGTCAAGTACTGATCTCACTTCAACAATTTGATGAGAGCTT  
 TCTGGCTGGCAATCTAAAAAACCAGGACTTCCAGGTTGAAGATTAAGTCCCA  
 AAATTTCCAGGAGACAGCAGAGATGAGAGTGAAGGATGATGATCTCCAGGC  
 CTTCAAGGCAAGCCACTGAGGATGTTGAAGAGACGAAGTAAAGTCTGGAGA  
 AAAGGACAGATAGTGTGAGGATGATGACTCTGATTAGACCCAGATAA  
 ATTTGTGCTGCTCTGTCTCA

Figure 19

2

Mouse FANCD2

SEQ ID NO: 8

GGAAAGTCGAAAAACGAAGGGAAGCAACTGGCCGGTCCCGAGGAAGTAATATAA  
 GTGGCAGAAAGACGTAGTCAAATGATTTCCAAAAGACGTGGCTAGATTCTGAG  
 GATAAAGAAAAACCTGACAGAAAGATGCCCTCCAAAACCAATGCCCTTTCCAAAGCTG  
 GCAAAGAAAGTCTCAAAATCTCATGAAGTTGAAGAAAAATGGCAGTGTCTTTGTA  
 AGCTTCTAAGGCTTCAGGACTCACTCTTAAAACCTGGAGAGAAACCAAAATCAGCT  
 AGGTGGATCAGGTAATCTTCCAAAGGAACTCTTTCAGGCCCTTGAGGAAACAT  
 CCTGCTTATCCCAAAGTAATAGAAGAGTTTGTAAATGGCCTGGAGTCTACACTGA  
 GGACAGTGAAGTCTCAGGAACTGCCTGCTGCTTTGTGAGCGCTGCAGGATGAG  
 GAAGCCAGCATGGGCACATTTTATCCAAAGAGTCTGATCAAGCTACTCTGCGGGA  
 TTGACATTTTACAGCCTGCCAATTAACAAAATGTTATTTGAAAAAGTGCCTCAGTTTC  
 TTTTGAAGAGTGAAGACAGAGATGGAATCAACATGGCCAGACTCATTATCAATCA  
 ACTAAAATGGCTGGATAGAAATTTGGATGGCAAGGACCTCACGGCCAGATGAT  
 CGAGTTGATCAGTGTGCTCCCTGAACTTACAGCATGACTTATCAGCAGCCCTTC  
 CTGAAAATCTAGGGGATTTCCAGCATGCTAATTTGGGAAAGAGCTTGGCAGCT  
 GCTGGTGCAGAAATCTCCCTGACTGTTCCAATTTGGATGCTTTTCCAGTCTCCG  
 ACTTGACCCCAACTCTCTGTCAGATCCGCCAGTTGGTGTATGGGCAAGCTGTCA  
 CTGTCCGCTAGAGGATTTCCCTGTGATTTGAAAGTTCCTTCTCAATTCGTAACAG  
 ACACCACTTCCCTTGAGGTCATTTGCCGAGCTTCCGGGAGAACTTGAACCTCAGCA  
 GTTTATTTGCCGTCACGAAATCAGGCTTCCAAAAGCAAAATGAAAAGTAAAGGA  
 CTAGCAAGCTTTCAGGAAATCAAGAGAACAGCTGATAAAGACTGTATTTGTTCTG  
 TCTTTGATGTAATAAAGTCAGCCATTAGATAATGAGAAAAACCAATTTACAGAGCC  
 GTTTAAGGCAATGAACGATTGAGTCCGCGCTGAACATAAGGCTTTGGACGTG  
 GTATGCTGCTCATCTACAGCACCAGCAGCAGACCAAGAAAGGCGTGGAG  
 AAGCTCTGAGAAAACAAAGATTTCAGTCAGACTGCAATTCAGAAACAGCTGTGAC  
 GTGCGTCTCTACACATTAAGCTGTTCTTAAAGGATAATTTGCCATCTATCTTTGCT  
 GGCTCAGACTTTGTTTCACTCTCAAGACCAGAGGATCATTGTTTGGCAGTCTTCT  
 GTACAAATATGCTTTTAAAGTTTGTGATCTTACTGCCAGCAGGAAAGTGGTGGT  
 CCTAGTCAACCATGTCTGAGTGGGACTGAGGCTGAAGTCCAGACTGCACTGGAG  
 TGTCTCTGAGGCTGATTGTCTAAACCCCTCTGCTATGAGGCTCAATGCTGCTTT  
 TGTAAAGGCAATTCATGATTTTGGAAAAATATGTCCTCTCAGCAAAATCAGAAAA  
 ATCTCTGATTTCTCAGACTCTTGCATTTAGCCAAACAGCCCGGAACCCAGCAACCA  
 ATCCAGGACGACATGCACTGGTGTATCCCGGAAGCAGCTCTCTAGCAGTGTGTT  
 AAGTACAAGCTCAATGGGATCAATGGTGCAGTCCCATGGCCGGCATCATGGCGG  
 AAGCAGAAAGTGTACCATTAACCTATCCAGAGGAGCCCAATGTGAGCAGG  
 AGCAGCCGACACAGGGGACTTCTTTGCTACTAAGTGTCTTCTTGTGACTGACGAC  
 TCTCTTGGCCCTCTCTCTGATTAATGATGAATTTGCCAACTGATCAAGAAAG  
 GAAGTGGCTCCAAAACCTTGGAGTGGTGGGAGCAGCATCTCAATGATTT  
 CAAGATGCCCTTTGGTGTAGACTCTGTGCTGCTCCAGAGGGTGAATTTCCATTTCT  
 GTGAAAGCCGCTCTATGGAAGTGAAGAGTACAGCACTCAAGCCGATTTGTCATCA  
 ACTCTCGCCCTGTCTATCAGGAAATGTGCAAAAGATGCCAGTCGAGGCAATC  
 ACAAGAATCGAGCCAGAGATCAATGCTTCTTTGTGCTGGCTTCCCAATTCGGCC

TGCTGAGACTTTGGCTGGCAAGCAACATGATGGAACTTGGATGAGATCGATGG  
 TCTCTAGATTGTCCCTGTCTCTCCCTGACCTGGAACTTGGAGAACTTGGAGT  
 CCATGTCTGCTAAAAGACCGTTCGCTTATGTGTTCCCTCACATTCCTAACTTCACT  
 GGTTCCGAGAGGTTTGAATGCCCTTCCCAACAAACATCTCTGAGATGAAGGG  
 CAAGGTTCTTAGTCCGCTAAAAGACCTTGTAGAACTTCAAGGAACTCTAGAGAA  
 TACTTGGCAGTCACTCCAGACTATGTTCCGCTTTCGCAAGCGTTCAGTGGACAC  
 TTTAGATATGATGCCTAGGAGCAGTCTGCTGTGTCAGCAAAAAACAGAAACAA  
 GGAAAGACGGGGGAAAGAAACAAAAAGCTGATAGCAACAAAGCATCTCTGTT  
 GGACACACTTCAACAGAAAGACACTTCAGAGTGTGACATGGCCCATCTGGGAG  
 AAGCCACGTAGACAAGGAGTCCACAGGGAAGGAAAGGAAAGACGTTTGTGCTACT  
 CAGAAATACCGGCTTTTCCGAGAGCTGGACATTGAGGTTCTCTTATTCTAC  
 ATTCTGGACTTGTGACCAAGTTCATCTTAGACACTGAAATGCACACTGAAAGTAC  
 AGAGGTCTACAGCTGGGGCTGCTGAGCTGCTCTCTTCTGCTGGAAGATCTTTCC  
 AGAAGCTAGAGAAATGCTGACTGCTCTTTTCCAAAGAGAATCTGTGCTTTAAG  
 AATAAAGGAAGGCAGAAATATGGCTTCTCACATCTTCATCAGAGATCTCTCCAG  
 ACATGTGCACTGTGGTTCAGCTGCTAAACCCGATGTGTAACCATCTGGGAAAC  
 ATTCACAACTTTTTCAGTGTAGTGGTGTGAGCATCTCAGTGCAGATGACAAGGC  
 GAGAGCGACAGCTCAGGAGCAGCACACCTGCTGCTTCCAGAACTCCAGAAAGCT  
 GCAGGCTTTCAGCAGCTTTTGGTGGAAAGGATTACTACCAATCAAAGCAC  
 CGCTCTCTGCACTCAGCCCTTGAAGTCTCTGCAACCCGATAAAGCAGATGGAA  
 AGGACAGCCCTTGGGAACTGGTCAAGCCAGAGCTTCAGTACTTGCAGAACTT  
 CCACATAGTGTCCAGTTCAGTGTGGTCTACCTTCTCAGACTCTGATGGC  
 CTTCTGGAGAAAGTCTGAGTACTTAAACAGAAAGAAAGAAAACTTGCCTCTCTG  
 GCCAAACAGCTGCTTTGCCGAGCATGGCCCTATGGGGAAGAAAGAAAGAAACCC  
 ACTTTAATGACCACTGATGATGTGCTTTTACATCTACTTGGAGCACACAGACAA  
 TGTCTGAAAGCCATAGAGGATYACTGGTGTGGTGTCCAGAACTGGTCACT  
 GCTCCGAAAGCAGCCGCTCTCTTACATCTTACAGTTCAGTTCAGCCGACACCTTGT  
 CATATTCTTCCGTGTGATGATGGCTGAATCGGAAAGAGCGGTGAAGGGTCTYAC  
 GCTGGCAGCAGCAGATTCGAGCAGGTTTCAAGAGAAAGAGCTCTCTATTGKA  
 ACATGGCTGTCCGAGATTTTCAGYATCTTCTTCAATCTGATGAAAGTATTGACAGT  
 TATCTTCTGCTGATGTGTTTAAAGATTTGGCCGCTGCTTTGTGGAGCATTTCTG  
 AAGCAATGTATGCCACTCCTGACTTACGCTTAAAGAAAGCATCCGGGAAAGTGT  
 TGAGCTTGTGCAACCCCTCAGTGAACAGGAGGCTACTTATCACCCTTTTGGGA  
 CACTCCAAGATTCGCCAGGACACAAGACTCAAGCAAGAYGTGCTTACTCAAAA  
 AGTCACTGAACTGTTAGTTTGCAGAGTCAAAGCCATGCTTCTCTCAACAACGT  
 AGAGAGGCTTCTGGTGGGTACTCTCAAAAAACCCAGACTTACAGGGTGAAGAAA  
 TTATTTCCAGGATCCCTCTTCTCAGAGAGCAATGCAAGAGGACAGTGAAGGATGG  
 CGTGCATCTCAGCTTCCAGGAAACAGAGCAACAGAGGATGGGGAAGATGAAGC  
 AAGTGTGAACAAGAGGACAGGACAGTGTGAAAGTGAAGGACAGCTCCAGTTA  
 GAGCCGATGGCATGGTGCCTGCTCACCTGACAGACTCTCATCTTTTGGGG  
 TTTGAAGTCAAGTGTCTGTTTTTCTAGTCAGAAGCATCTTGTGTCATCAAGAA  
 GGGGTGTTATTTAATCCCAAGTGGGTTTACAGGTTGTCTAACCTCCAGGTCCT  
 GGTTCCAGGATCCAGTGTAGCATCCATCTGTTACTAGGAYGAACATGGCTGGCT  
 CGATGCACTKAGTGCAGGTGCCCTAGCTGGCCCTTGGGTTTTGAACTAAAA

Figure 20

TTTAGGCTTATAATAGCTTTGTAATAAATCTGTTTCCAGAGTTCCTCAGCTACC  
 TTTTCTCCTCACTTATGATGATTAATCAAGGATCTCAITATCAAGGATTAGGTA  
 TATGAGTGTAGGTTTGTGCAATCTGACTGGTGGCTAAAAGTATGTTCCGACTG  
 TTATCTTCCGGAGGAATGACCCAACTTCTTATCAATGATCAAGTGTITGGTTGG  
 TCTGTGTCAGGCTCTTTACATAGTCTGGCTGGTGTGTTATGATATGTTGACC  
 AGGAGGCTTGAACATTAATTTGAATTTTAAACATTTTGTACATAITGTGATGG  
 GCATATATGTCCTACTGTGCATATGTTAGGTCAGAGGATAGCTTATGGGATGA  
 GCTCTCTCTCCACCATGTTGGTTCAGGGTTCAAACCTTAGACCTTACCTGCTC  
 AGCCACCTTACCCTTTTAAAATGTTGGTTATTAATATATAAAGGAAGGAAGAC  
 AACATCAAACATGTGCTGGCTTGTATGATATATAGTTTATTTCCACATTAAT  
 TGAATATGCCATAATATTTGTAATAATCATACAAAATAATGTAATTTATTA  
 GAAATAGAACAATCAGGAGTAAAATAGGGGATCTTCTGCTTCTGCCAGGAAGC  
 CCAGTCTCAGAGATGCTGCCAGGCTTCTCTGCTGCTGCTTAAAGATTTAAT  
 TTTTGTAAATTTTACTCATACCGGATTAAGGTTATGTTTGTGGAAAAAAA  
 AAAAAAATAAAAAAATAAAAA

专利名称(译)	用于诊断癌症易感性和缺陷DNA修复机制及其治疗的方法和组合物		
公开(公告)号	<a href="#">JP2010246550A</a>	公开(公告)日	2010-11-04
申请号	JP2010119683	申请日	2010-05-25
[标]申请(专利权)人(译)	达那-法伯癌症研究所		
申请(专利权)人(译)	Dana Farber癌症研究所有限公司 俄勒冈健康与科学大学		
[标]发明人	ダンドレアアランディー タニグチトシヤス ティマースシンシア グロンプマーカス		
发明人	ダンドレア、アランディー、 タニグチ、トシヤス ティマース、シンシア グロンプ、マーカス		
IPC分类号	C12N15/09 C12Q1/68 C12Q1/02 A61P35/00 A61P17/00 A61P35/02 A61P25/00 A61P1/18 A61P15/00 A61P13/12 A61P1/04 A61K48/00 G01N33/53 G01N33/533 G01N33/50 G01N33/15 G01N27/447 A61K38/00 A61P7/06 A61P43/00 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12P21/02 C12P21/08 C12Q1/6886 G01N33/543 G01N33/566 G01N33/574		
CPC分类号	A01K2217/05 A01K2217/075 A61P1/04 A61P1/18 A61P13/12 A61P15/00 A61P17/00 A61P25/00 C07K14/47 C07K16/18 C12Q1/6886 C12Q2600/156 C12Q2600/158 G01N33/5091 G01N33/574 G01N33/57484 G01N2500/00		
FI分类号	C12N15/00.ZNA.A C12Q1/68.A C12Q1/02 A61P35/00 A61P17/00 A61P35/02 A61P25/00 A61P1/18 A61P15/00 A61P13/12 A61P1/04 A61K48/00 G01N33/53.D G01N33/533 G01N33/53.M G01N33/50.Z G01N33/15.Z G01N27/26.301.A C12N15/00.A C12N15/00.AZN.A C12Q1/6886.C C12Q1/6886.Z G01N27/447.301.A		
F-TERM分类号	2G045/AA26 4B024/AA01 4B024/AA12 4B024/CA01 4B024/CA04 4B024/CA11 4B024/HA08 4B024 /HA12 4B063/QA17 4B063/QA18 4B063/QA19 4B063/QQ02 4B063/QQ08 4B063/QQ42 4B063/QQ52 4B063/QR08 4B063/QR32 4B063/QR35 4B063/QR42 4B063/QR50 4B063/QR55 4B063/QR62 4B063 /QR72 4B063/QR77 4B063/QS03 4B063/QS25 4B063/QS28 4B063/QS34 4B063/QS36 4B063/QX02 4C084/AA13 4C084/NA14 4C084/ZA021 4C084/ZA661 4C084/ZA811 4C084/ZA891 4C084/ZB261 4C084/ZB271		
代理人(译)	三好秀 伊藤雅一 原 裕子		
优先权	60/245756 2000-11-03 US		
外部链接	<a href="#">Espacenet</a>		

#### 摘要(译)

要解决的问题：提供用于诊断癌症易感性，缺陷的DNA修复机制及其治疗的方法和组合物。 解决方案：FANCS2基因序列作为探针和引物用于诊断范可尼贫血和癌症，并筛选新的治疗药物。 制备实验小鼠模型，一种基因治疗阳性载体，其使用新型治疗剂进行筛选以治疗涉及缺陷DNA修复的疾病。 鉴定FANCD2亚型，并制备可区分这两种亚型及其在诊断测试中的用途的多克隆和单克隆抗体。 [选择图]图8

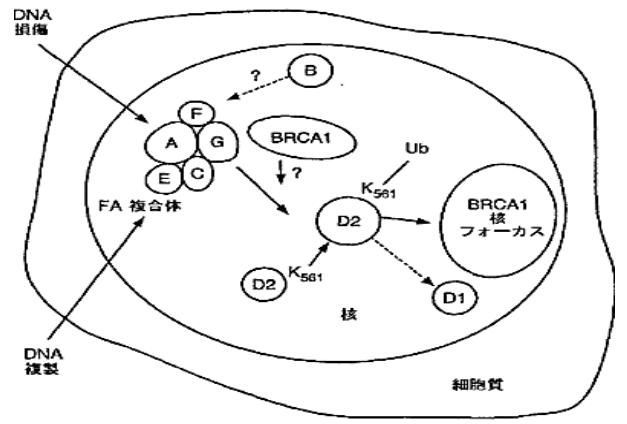


FIG. 8