

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2005-505272  
(P2005-505272A)

(43) 公表日 平成17年2月24日(2005.2.24)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	4 B O 2 4
C 1 2 Q 1/68	C 1 2 Q 1/68	4 B O 6 3
G O 1 N 33/53	C 1 2 Q 1/68	Z
G O 1 N 33/566	G O 1 N 33/53	M
	G O 1 N 33/566	
審査請求 未請求 予備審査請求 有 (全 71 頁)		

(21) 出願番号	特願2003-529994 (P2003-529994)	(71) 出願人	390033008
(86) (22) 出願日	平成14年9月17日 (2002. 9. 17)		ジャンセン・ファーマシューチカ・ナーム
(85) 翻訳文提出日	平成16年3月10日 (2004. 3. 10)		ローゼ・フェンノートシャツプ
(86) 国際出願番号	PCT/EP2002/010667		JANSEN PHARMACEUTICA
(87) 国際公開番号	W02003/025222		NAAMLOZE VENNOOTSCHAP
(87) 国際公開日	平成15年3月27日 (2003. 3. 27)		ベルギー・ビー-2340-ビールセ・ト
(31) 優先権主張番号	01203558.0		ウルンホウトセベーク30
(32) 優先日	平成13年9月17日 (2001. 9. 17)	(74) 代理人	100060782
(33) 優先権主張国	欧州特許庁 (EP)		弁理士 小田島 平吉

最終頁に続く

(54) 【発明の名称】 双極性障害と関連する脳で発現されるCAP-2遺伝子およびタンパク質

(57) 【要約】

我々は以前に双極性 (BP) 障害の候補領域として 18q21.33-q23 を同定し、そして酵母人工染色体 (YAC) コンティグ地図を構築した。次の段階として、我々はこの領域から全ての CAG / CTG 反復を単離しそして解析し、それらを BP 障害への関与から除外した。今回、該領域から全ての CCG / CGG 反復を同定する過程において、我々はセルピンをコードする 6 個の遺伝子のクラスターを見出した。それらのうちの一つ、18q21.3 に位置する CAP2 に関して、突然変異解析を行った。このようにして同定された一塩基多型 (全部で 6 種) の解析により BP 障害と SNP c.942G>T との統計的に有意な関連が示された。従って、個体からのサンプルにおいて、個体の CAP2 遺伝子における一塩基多型を決定すること、および CAP2 遺伝子における多型を参照することにより個体の状態を決定することを含んでなる個体における BP もしくは BP に対する感受性を診断する方法を提供することは本発明の目的である。

## 【特許請求の範囲】

## 【請求項 1】

個体からのサンプルにおいて、個体の C A P 2 遺伝子における一塩基多型を決定すること、および C A P 2 遺伝子における多型を参照することにより個体の状態を決定することを含んでなる個体における B P もしくは B P に対する感受性を診断する方法。

## 【請求項 2】

個体の一塩基多型が C A P 2 遺伝子における多型と連鎖不平衡にある請求項 1 に記載の方法。

## 【請求項 3】

一塩基多型が S N P c . 9 4 2 G > T に等しい請求項 1 もしくは 2 に記載の方法。

10

## 【請求項 4】

個体の C A P 2 遺伝子における一塩基多型が、A R M S 法およびサザンブロッティング技術のような制限断片長多型、S S C P 解析、ミスマッチ化学切断法および変性高速液体クロマトグラフィーから選択される方法により決定される請求項 1 ~ 3 のいずれか一つに記載の方法。

## 【請求項 5】

個体の C A P 2 遺伝子における一塩基多型が、一塩基多型を含有する C A P 2 遺伝子の断片を増幅する 1 対の P C R プライマーを用いて決定される請求項 1 ~ 3 のいずれか一つに記載の方法。

## 【請求項 6】

一塩基多型が S N P c . 9 4 2 G > T からなる請求項 5 に記載の方法。

20

## 【請求項 7】

C A P 2 遺伝子の断片を増幅する P C R プライマー対が、配列番号 9 および配列番号 1 0 の配列を含んでなるフォワードおよびリバースプライマーからなる請求項 5 に記載の方法。

## 【請求項 8】

一塩基多型を含有する C A P 2 遺伝子の断片を増幅することができる 1 対の P C R プライマー。

## 【請求項 9】

配列番号 9 および配列番号 1 0 の配列を含んでなるフォワードおよびリバースプライマーからなる 1 対の P C R プライマー。

30

## 【請求項 10】

請求項 8 もしくは 9 に記載の P C R プライマー対を含んでなる診断キット。

## 【請求項 11】

個体の C A P 2 遺伝子における一塩基多型が、対立遺伝子特異的オリゴヌクレオチドプローブによって決定される請求項 1 ~ 3 のいずれか一つに記載の方法。

## 【請求項 12】

個体の C A P 2 遺伝子における一塩基多型 S N P c . 9 4 2 G > T を検出することができる対立遺伝子特異的オリゴヌクレオチドプローブ。

## 【請求項 13】

請求項 12 に記載の対立遺伝子特異的オリゴヌクレオチドプローブを含んでなる診断キット。

40

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、広範には精神医学的健康と関連する遺伝因子の決定に関する。さらに特に、本発明は、罹患個体およびそれらの家系における気分障害もしくは関連疾患に関係があるヒト遺伝子に関する。特に、本発明は、細胞質アンチプロテイナーゼ 2 ( C A P 2 ) をコードする遺伝子に関する。該遺伝子は第 1 8 染色体上に位置し、そして脳組織において発現され、そして双極性障害の診断マーカーとして用いることができる。

50

## 【背景技術】

## 【0002】

薬理遺伝学背景：

あらゆる個体は、それらの遺伝子および環境の相互作用の産物である。薬理遺伝学は、遺伝的な違いが薬剤に対する患者応答の変動にどのように影響を及ぼすかという研究である。薬理遺伝学の使用によって、我々はまもなく特定の医薬に対する応答を予測するために個体のDNA間の変異の概略を提示することができる。ヒトにおける臨床徴候の忍容性に優れ且つ有効な医薬を予測する標的検証は広く認められている問題であるが；真の難題は標的選択である。高処理量スクリーニングが現在可能である、受容体および酵素を包含する、限られた数の分子標的ファミリーが同定されている。優れた標的は、有効な分子（ヒット）を同定するためにそれに対して多数の化合物を迅速にスクリーニングすることができるものである。これらのヒットは、忍容性に優れ且つ有効な医薬の性質を有する最適化分子（リード（lead））に発展させることができる。疾病もしくは臨床徴候について検証することができる標的の選択は、製薬産業が直面する主要な問題である。最もよく検証された標的は、ヒトにおける忍容性に優れ且つ有効な医薬をすでに産み出しているもの（先例標的）である。多数の標的が科学的仮説に基づいて選択され、そして最初の仮説がたいていその後反証されるので有効な医薬につながらない。

10

## 【0003】

高処理量標的としての用途に遺伝子を同定しそしてそれらのタンパク質生成物を発現するために2つの幅広い戦略が用いられている。ゲノム学および遺伝学のこれらの方法は技術を共有するが、異なる科学的戦略および投資を意味する。発見ゲノム学は、疾病に遺伝学的に関連することが知られていない扱いやすいもしくはスクロール可能な標的の遺伝子および遺伝子のファミリーを同定するためにDNA配列情報の増加する数のデータベースを用いる。

20

## 【0004】

患者から得られる疾病感受性遺伝子に関する情報の利点は、定義により、これらの遺伝子が疾病への患者の遺伝的寄与に関係があることである。しかしながら、大部分の感受性遺伝子は、扱いやすい標的ではないかもしくは有効な化合物を同定するための高処理量スクリーニング方法が容易にできない。

## 【0005】

関連する遺伝子変異体に関する差次的代謝は、疾病の発症もしくは進行の機構を発見するために集中的な機能性ゲノムおよびプロテオーム技術において研究することができる。

30

## 【0006】

改変された代謝と関連する受容体の重要な酵素は、標的として用いることができる。適切な細胞代謝に対する特定の感受性遺伝子変異体の役割に焦点を合わせる遺伝子機能標的（gene-to-function-to-target）戦略は重要になる。

## 【0007】

強力な生物情報学手段でのヒトゲノムプロジェクトおよび同様のプログラムからの配列のデータ採掘は、同様の配列を有するドメインを位置づけることにより遺伝子ファミリーを同定することを可能にしている。これらのゲノム戦略により同定される遺伝子は、一般に、ある種の機能的検証もしくは疾病プロセスとの関連性を必要とする。遺伝子と疾病との間の関連性を示すために差次的遺伝子発現、トランスジェニック動物モデル、プロテオミクス、in situハイブリダイゼーションおよび免疫組織化学のような技術が用いられる。

40

## 【0008】

ゲノム方法と遺伝的方法との間の主要な違いは標的選択であり、それは疾病プロセスに関連することがすでに既知である遺伝子および変異体特異的標的を遺伝学的に特定した。疾病との関連性の探求における各遺伝子での非特異的な大規模な遺伝子同定のための発見ゲノム学の最近の流行は、医薬の開発の大きな機会を生み出す。

## 【0009】

50

薬剤開発の中核となる問題は乏しい標的選択であると認識することもまた重要である。疾病関連検証を示すための実証されていない技術のスクリーニング使用、およびヒトにおける概念の実証へ各選択遺伝子を進めるのに必要な莫大な投資は、「検証」という言葉の実証されていないそして無頓着な使用に基づく。各失敗は、損失時間および金銭において非常に高つく。例えば、差次的遺伝子発現 (DGE) およびプロテオミクスは、標的検証に広く用いられるスクリーニング技術である。これらは、組織における遺伝子およびタンパク質発現の異なるレベルおよび/もしくはパターンを検出し、その組織を冒す疾病との関連性を示すためにそれらを用いることができる。

気分障害背景：

気分障害もしくは関連疾患には、精神障害の診断および統計マニュアル (Diagnostic and statistical Manual of Mental Disorders)、バージョン4 (DSM-IV) 分類学DSM-IVコードを括弧に入れて) において特定されるような以下の疾患：気分障害 (296.XX, 300.4, 311, 301.13, 295.70)、統合失調症および関連疾患 (295.XX, 297.1, 298.8, 297.3, 298.9)、不安障害 (300.XX, 309.81, 308.3)、適応障害 (309.XX) および人格障害 (コード301.XX) が包含されるがこれらに限定されるものではない。

【0010】

本発明は、特に、双極性 (BP) スペクトル障害として知られている気分障害の家系と関連する遺伝因子に関する。双極性障害 (BP) は、高揚感の極度の状態 (躁病) から抑うつ (鬱病) までにおよぶ気分の障害を特徴とする重度の精神医学症状である。2つのタイプの双極性障害が記述されている：I型BP障害 (BPI) は、躁病の相と交互に起こる大鬱病エピソードを特徴とし、そしてII型BP障害 (BPII) は、軽躁の相と交互に起こる大鬱病エピソードを特徴とする。BP発端者の親族は、BP、単極性障害 (鬱病エピソードを経験するだけの患者；UP)、循環気質 (小鬱病および軽躁エピソード；cy)、並びに躁病 (SAM) および鬱病 (SAD) タイプの統合失調性感情障害の増大した危険性を有する。これらの結果に基づいて、BP、cY、UPおよびSAはBPスペクトル障害と分類される。

【0011】

BPスペクトル障害の病因への遺伝因子の関与は、家系、双子および養子縁組研究により示唆された (非特許文献1)。しかしながら、伝達の正確なパターンは不明である。いくつかの研究において、複雑な分離解析によりBPの単一の主要な遺伝子座の存在が裏付けられる (非特許文献2)。他の研究者等は、罹患性閾値モデルを提示し、ここで、疾病を発症する罹患性は、多数の遺伝および環境作用の加法的組み合わせに起因する (非特許文献3)。

【0012】

遺伝の複雑な形式のために、BP障害がメンデルの法則のように伝達されると思われる家系においてパラメーターおよび非パラメーター連鎖戦略が適用される。染色体11p15 (非特許文献4) およびXq27-q28 (非特許文献5；非特許文献6) 上の初期の連鎖結果は議論の余地があり、そして最初は再現することができなかった (非特許文献7；非特許文献8)。高度に多型のマーカーで飽和したヒト遺伝子地図の開発およびデータ解析技術の絶え間ない発展とともに、多数の新しい連鎖検索が行われた。いくつかの研究において、第4、12、18、21およびX染色体上の特定の領域への連鎖の証拠もしくは示唆的証拠が見出された (非特許文献9、非特許文献10、非特許文献11、非特許文献12および非特許文献13)。報告された連鎖結果の正当性を試験するために、これらの結果は他の独立した研究において再現されなければならない。

【0013】

最近、第18染色体上のセントロメア周囲 (pericentromeric) 領域への双極性障害の連鎖が報告された (非特許文献11)。18pter-p11および18q23-qterで切断点および欠失領域を有する環状第18染色体もまた、BP障害もし

くは関連症候群にかかっている3人の関係のない患者において報告された(非特許文献10)。染色体18p連鎖はStine et al. (非特許文献14)によって再現され、彼らはまた、同じ研究において18q21.2-q21.32上の遺伝子座の示唆的証拠も報告した。

【0014】

興味深いことに、Stine et al. は親由来効果(parent-of-origin effect)を認め、連鎖の証拠は父方の家系において最も強く、ここで、発端者の父もしくは発端者の父の兄弟姉妹の一人が罹患している。いくつかの研究は、BP障害を伝達する家系における表現促進を記述し(非特許文献15、非特許文献16)、CAG/CTG、CCG/CGGもしくはGAA/TTG反復の増加によって引き起こされる多数の疾病が表現促進を示すこと(非特許文献17により概説される)を考慮して、トリヌクレオチド反復増加(TRE)の関与を示唆する。潜在的に増加した反復を見出そうとする以前の取り組みは、CCG/CGG反復の検索が増えているが、主としてCAG/CTG反復に集中している(非特許文献18、非特許文献19、非特許文献20、非特許文献21)。以前に、我々は三塩基反復の領域特異的単離の新規な方法：三塩基反復YAC断片化に関して報告した(非特許文献22)。これはCAG/CTG反復の単離の有効な方法であると判明し、そしてこの方法を用いて、我々は18q21.33-q23の範囲内から双極性障害へのCAG/CTG反復の関与を除外した(非特許文献23)。本発明はCCG/CGG反復の領域特異的単離の方法を適応させ、そしてそれを染色体18q21.33-q23 BP候補領域に適用した。

10

20

【非特許文献1】

Tsuang and Faraone (1990), the Genetics of Mood Disorders, Baltimore, The Johns Hopkins University Press

【非特許文献2】

Spence et al. (1995), Am J. Med. Genet (Neuro psych. Genet.) QQ pp 370 - 376

【非特許文献3】

MuGuffin et al. (1994), Affective Disorders; Seminars in Psychiatric Genetics Gaske 30  
ll, London pp 110 - 127

30

【非特許文献4】

Egeland et al. (1987), Nature ~ pp 783 - 787

【非特許文献5】

Mendlewicz et al. (1987, the Lancet 1 pp 1230 - 1232)

【非特許文献6】

Baron et al. (1987) Nature 12 & pp 289 - 292

【非特許文献7】

Kelsoe et al. (1989) Nature ~ pp 238 - 243

40

【非特許文献8】

Baron et al. (1993) Nature Genet ~ pp 49 - 55

【非特許文献9】

Blackwood et al. (1996) Nature Genetics ~ pp 427 - 430

【非特許文献10】

Craddock et al. (1994) Brit J. psychiatry ~ pp 355 - 358

【非特許文献11】

Berrettini et al. (1994), Proc Natl Acad S 50

50

c i U S A ~ p p 5 9 1 8 - 5 9 2 1

【非特許文献12】

S t r a u b e t a l . ( 1 9 9 4 ) N a t u r e G e n e t i c s ~ p p 2 9 1 - 2 9 6

【非特許文献13】

P e k k a r i n e n e t a l . ( 1 9 9 5 ) G e n o m e R e s e a r c h 2 p p 1 0 5 - 1 1 5

【非特許文献14】

S t i n e e t a l . ( 1 9 9 5 ) A m J . H u m G e n e t 2 2 p p 1 3 8 4 - 1 3 9 4

【非特許文献15】

M c I n n i s e t a l 1 9 9 3

【非特許文献16】

N y l a n d e r e t a l 1 9 9 4

【非特許文献17】

M a r g o l i s e t a l . , 1 9 9 9

【非特許文献18】

K l e i d e r l e i n e t a l 1 9 9 8

【非特許文献19】

M a n g e l e t a l 1 9 9 8

【非特許文献20】

E i c h h a m m e r e t a l 1 9 9 8

【非特許文献21】

K a u s h i k e t a l 2 0 0 0

【非特許文献22】

D e l F a v e r o e t a l 1 9 9 9

【非特許文献23】

G o o s s e n s e t a l 2 0 0 0

【発明の開示】

【0015】

[発明の要約]

本発明は、新規な単離された核酸配列および単離された核酸配列によりコードされる細胞質アンチプロテイナーゼ2 (CAP2) タンパク質に関する。

【0016】

新規な単離された核酸配列は、18q21.33-q23でD18S68とD18S979との間にある8.9cM染色体領域に位置する。酵母人工染色体(YAC)(Verheyen et al 1999)を用いて物理的地図を構築した。

【0017】

以前に記述した方法をCCG/CGG反復の領域特異的単離に適応させ、そして染色体18q21.33-q23 BP候補領域に適用した。YACコンティグ地図により、セリンプロテイナーゼインヒビター(セルピン)をコードする6個の遺伝子のクラスターのBP候補領域内の局在が裏付けられた。セルピンは、 $\alpha_1$ -プロテイナーゼインヒビターに高い全体的な相同性を有するタンパク質分解タンパク質のスーパーファミリーである。6個のセルピンは全て、典型的なアミノ末端の切断可能なシグナルペプチドを欠きそして細胞内もしくは両方であることができるセルピンのオボアルブミンファミリーに属する。CAP2もしくは18q21.33に位置するP18は、その5'UTR領域に組み合わされたCAG-CGG三塩基反復配列を含有し、そして脳において発現される。この研究において、我々はCAP2のゲノム構成およびエキソン/イントロン境界を決定し、そして配列変異体の一本鎖高次構造多型(SSCP)解析および変性高速液体クロマトグラフィー(DHPLC)により該遺伝子を調べた。シーケンス、RFLP-PCRもしくはパイ

10

20

30

40

50

ロシーケンスによる6種の一塩基多型(SNP)の解析を75人の症例および75人の一致するコントロールのサンプルにおいて行った。

[発明の詳細な記述]

本発明は、第18染色体の18q染色体候補領域に位置する遺伝子を含んでなる新規な単離された核酸配列に関する。

【0018】

該遺伝子は、双極性スペクトル障害のような気分障害と関連する染色体領域に位置し、従って、双極性スペクトル障害の診断マーカーとして有用である。問題の領域はまた、ヒトゲノムの全体から取り除いた場合に、精神医学的健康および気分に影響を与える他の遺伝子を位置づけ、単離しそしてシーケンスするために用いることもできる。

10

【0019】

特に、BP候補領域は、多数の細胞内および細胞外機能に関与する脳で発現されるセルピン、細胞質アンチプロテイナーゼ2(CAP2)をコードする遺伝子を含む。本研究において、我々はCAP2のゲノム構成を決定し、そして全てのイントロン/エキソン境界を特定した。CAP2は、概算される17kbのゲノム領域内に7個のエキソンを含んでなる。

【0020】

CAP2の突然変異解析により、3種の非同義一塩基多型(SNP): c.203G>A (Arg69Gln)、c.910A>G (Thr304Ala)およびc.1076G>A (Arg359His); 2種と同義SNP c.477>Gおよびc.942>T 20  
ならびに1種のイントロンSNP IVS4+98A>Gが同定された。関係のないBP症例および一致するコントロールにおけるCAP2多型の解析により、対立遺伝子および遺伝子型頻度においてSNP c.942C>Tで統計的に有意な差が示された(p=0.03)。

20

新規な遺伝子の単離および同定:

関連する遺伝子を同定しそして特性化する過程において、同定されたYACクローンにそして適用できる場合には本明細書において特定するような気分障害に苦しむ個体からのDNAに、当業者に周知である標準的方法を適用した。例えば、本発明者等は、ヒト第18染色体上の対象の領域中の候補遺伝子を同定する目的で、選択したYACクローンにおけるTREについてスクリーニングするために、ヒトゲノム内のトリヌクレオチド反復増加 (TRE) と気分障害における表現促進の現象との間の以前に同定された明白な関連 (Lindblad et al. (1995), Neurobiology of Disease 2, pp 55-62 および O'Donovan et al. (1995), Nature Genetics 1Q pp 380-381) を利用することができる。様々な他の既知の方法もまた、以下に説明するように候補遺伝子を同定するために該YACクローンに適用することができる。

30

【0021】

従って、第一の態様として、本発明は、上記に特定するような気分障害もしくは関連疾患と関連する、その突然変異したおよび多型の変異体を包含する、少なくとも一つのヒト遺伝子を同定するための多型マーカーD18S68とD18S979との間に配置されるヒト染色体18qの8.9cM領域もしくはその断片の使用を含んでなる。以下に記述するように、本発明者等は、D18S51とD18S61との間に以前に位置づけられた12のSTR多型マーカーと家系MAD31における双極性障害との共分離の解析およびその後の対立遺伝子共有解析により、そのような遺伝子の染色体18qのこの候補領域を同定している。

40

【0022】

本発明に従って用いることができる候補領域を含む特定のYACは、961.h-9、942-c.3、766-f-12、731-c-7、907.e.1、752-g-8および717-d-3であり、好ましいものは961h-9、766.f.12および907-e.1であり、これらは、候補領域にわたる最小のタイリング経路 (tiling 50

50

path)を有するからである。使用のために適当なYACクローンは、D18S68とD18S979との間の絞り込んだ候補領域にわたる人工染色体を有するものである。

【0023】

気分障害もしくは関連疾患と関連する1個もしくは複数の候補遺伝子を同定するために、YACに存在するかどうかを問わず、上記に特定するような染色体18qの候補領域に適用することができる多数の方法がある。例えば、上記のように、ヒトゲノムにおけるトリヌクレオチド反復増加(TRE)の程度と気分障害の存在との間には明白な関連がある。

【0024】

従って、第三の態様として、本発明は、多型マーカーD18S68とD18S979との間に配置されるヒト染色体18qの領域におけるヌクレオチド三塩基反復を検出することを含んでなる本明細書において特定するような気分障害もしくは関連疾患と関連する、その突然変異したおよび多型の変異体を包含する、少なくとも一つのヒト遺伝子を同定する方法を含んでなる。

【0025】

該遺伝子もしくは複数の遺伝子を同定する代わりに方法は、多型マーカーD18S60とD18S61との間に配置されるヒト染色体18qの部分を含んでなるYACクローン、例えば7個の上記のYACクローンの一つもしくはそれ以上を断片化すること、および該断片における任意のヌクレオチド三塩基反復、特にCAGもしくはCTGの反復を検出することを含んでなる。少なくとも5個そして好ましくは少なくとも10個のCTGおよび/もしくはCAG三塩基反復を含んでなる核酸プローブは、適切に標識した場合に検出の適当な手段である。トリヌクレオチド反復はまた、既知のRED(反復増加検出)系(S Halling et al. (1993), Nature Genetics ~ p 135-139)を用いて決定することもできる。

【0026】

第四の態様として、本発明は、気分障害もしくは関連疾患と関連しそして多型マーカーD18S60とD18S61との間のヒト染色体18qの領域にわたるYACクローンに存在する、その突然変異したおよび多型の変異体を包含する、少なくとも一つの遺伝子を同定する方法を含んでなり、該方法は、少なくとも8個であるが好ましくは少なくとも12個の連続したグルタミン残基の列を含んでなるアミノ酸配列を有するタンパク質を認識することができる抗体の使用によりヌクレオチド三塩基反復を含む遺伝子の発現生成物を検出するという工程を含んでなる。そのような方法は、ヒトDNA発現ライブラリーに、例えば7個の上記のYACクローンから、YAC DNAをサブクロニングすることにより実施することができる。関連する発現生成物を検出する好ましい手段はモノクローナル抗体、特にmAB1C2の使用により、その製造および特性は国際特許出願公開第WO 97/17445号に記述されている。

【0027】

本発明のさらなる態様は、BAC(細菌人工染色体)もしくはPAC(P1もしくはファージ人工染色体)のようなベクターまたはエキソン-トラップ(exon-trap)コスミドベクターのようなコスミドベクターへの上記に特定するような(YAC)DNAの(サブ)クロニングを含む関連する遺伝子もしくは複数の遺伝子を同定する方法に関する。そのような方法の出発点は、多型マーカーD18S60とD18S61との間のヒト染色体18qの領域のコンティグ地図の構築である。この目的のために、本発明者等は、7個の上記のYACクローンの各々におけるヒトDNAの断片の末端領域をシーケンスし、そしてこれらの配列を本明細書に開示する。上記のような他のベクターへのYAC DNAのサブクロニングの後、該サブクロンとコンティグ地図との間の重複部分を検出することができるように、本明細書に示すYACクローンコンティグに記述するような、この領域における任意の既知の配列標識部位(sequence tagged site)と一緒に、これらの末端配列もしくはその一部、特に本明細書において図1~11に示す配列を含んでなるプローブを構築することができる。現在のYACコンティグにおける既知の配列もまた、コンティグ地図サブクロンの作製に用いることができる。

10

20

30

40

50



## 【0028】

気分障害もしくは関連疾患と関連する1個もしくは複数の遺伝子を同定することができる一つの手段は、エキソントラッピング(exon trapping)の既知の技術の使用による。これは、特殊化したエキソン-トラップコスミドベクターの現在のプロトコルにおいてほとんどの場合利用する人工RNAスプライシングアッセイである。該ベクターは、複製起点および強力なプロモーター配列を含有するSV40ゲノムの断片、多数のクローニング部位を含有するイントロンで隔てられた2個のスプライシング能力を有するエキソンならびにSV40ポリアダニル化部位からなる人工ミニ遺伝子を含有する。

## 【0029】

YAC DNAをエキソン-トラップベクターにサブクローニングし、そして組換えDNAを哺乳動物細胞の株にトランスフェクションする。SV40プロモーターからの転写はRNA転写産物をもたらす、それは、通常、ミニ遺伝子の2個のエキソンを含むようにスプライシングする。クローニングしたDNA自体が機能性のエキソンを含有する場合、それはベクターのミニ遺伝子に存在するエキソンにスプライシングされることができる。逆転写酵素を用いてcDNAコピーを作ることができ、そして特定のPCRプライマーを用いて、インサートDNAのエキソンに関するスプライシング事象を同定することができる。そのような方法は、関連する遺伝子を同定するために気分障害もしくは関連疾患に苦しむ個体からのDNAの同等領域と比較することができるYAC DNAにおけるコーディング領域を同定することができる。

## 【0030】

従って、第五の態様として、本発明は、

(1) 上記のように調製しそしてマッピングしたエキソントラップコスミドベクターで哺乳動物細胞をトランスフェクションすること；

(2) 該哺乳動物細胞を適切な培地において培養すること；

(3) SV40プロモーターから発現されるRNA転写産物を単離すること；

(4) 該RNA転写産物からcDNAを調製すること；

(5) 該サブクローニングしたDNAにおけるコーディング領域の部分を解明するために該エキソントラップコスミドベクターにサブクローニングしたDNAのエキソンに関するスプライシング事象を同定すること；

(6) 該コーディング領域と該気分障害もしくは関連疾患に苦しむ個体のDNAにおける同等領域との間の違いを検出すること；および

(7) 該気分障害もしくは関連疾患と関連する該遺伝子またはその突然変異したもしくは多型の変異体を同定すること、

の工程を含んでなる気分障害もしくは関連疾患と関連する、その突然変異した変異体および多型を包含する、少なくとも一つのヒト遺伝子を同定する方法を含んでなる。

## 【0031】

エキソントラッピングの代替の方法として、YAC DNAをBAC、PAC、コスミドもしくは他のベクターにサブクローニングし、そして上記のようにコンティグ地図を構築することができる。以下のように、サブクローニングしたDNA上の関連する遺伝子の位置を定めることができる利用可能な様々な既知の方法がある：

(a) cDNA選択もしくは捕獲(直接選択およびcDNA選択とも呼ばれる)：この方法は、特定の(例えば脳)cDNAライブラリーからの全てのcDNAクローンのインサートのようなcDNAの複雑な混合物に、クローニングしたDNA(例えばYAC DNAのインサート)をハイブリダイズさせることによるゲノムDNA/cDNAヘテロ二本鎖の形成を含む。関連する配列はハイブリダイズし、そしてピオチン-ストレプトアビジン捕獲およびPCR(もしくは関連技術)を用いて次の工程において濃縮することができる；

(b) mRNA/cDNAへのハイブリダイゼーション：ゲノムクローン(例えば、特定のコスミドのインサート)を培養細胞系のパネルからのmRNAのノーザンプロットにもしくは適切な(例えば脳)cDNAライブラリーに対してハイブリダイズさせることがで

きる。陽性シグナルは、クローニングした断片内の遺伝子の存在を示すことができる；

(c) CpGアイランド同定：CpGもしくはHTFアイランドは、遺伝子の5'末端に存在することが多い短い(約1kb)低メチル化GCリッチ(>60%)配列である。CpGアイランドは、いくつかのまれにしか切断しない制限酵素の制限部位を有することが多い。まれにしか切断しない制限部位のクラスター化は、CpGアイランド、従って可能性がある遺伝子を示す。CpGアイランドは、まれにしか切断しない酵素で消化したゲノムDNAのサザンプロットへのDNAクローンのハイブリダイゼーションにより、もしくはアイランド-レスキューPCR(アイランドと隣接するAlu反復との間の配列を増幅することによるYACからのCpGアイランドの単離)により検出することができる。

(d) ズー・プロットイング：DNAクローン(例えば、特定のコスミドのインサート)を様々な動物種からのゲノムDNAサンプルのサザンプロットに対して減少したストリンジェンシーでハイブリダイズさせること。ハイブリダイゼーションシグナルの検出は保存配列を示唆し、可能性がある遺伝子を示すことができる。従って、第六の態様として、本発明は、

(1) 上記のようなYAC DNAをコスミド、BAC、PACもしくは他のベクターにサブクローニングすること；

(2) サブクローン間の重複部分を検出しそしてその地図を構築するために、本明細書に記述するYACクローンコンテグにおけるようなこの領域におけるヌクレオチド配列もしくは任意の他の配列標識部位(STS)、または14個以上の連続した塩基からなるその一部もしくはその相補物を用いること；

(3) CpGアイランド同定、ズー・プロットイング、cDNAライブラリーもしくは培養細胞系のパネルからのmRNAのノーザンプロットへのサブクローニングしたDNAのハイブリダイゼーションの一つもしくはそれ以上によりサブクローニングしたDNA内の遺伝子の位置を同定すること；

(4) 該遺伝子と気分障害もしくは関連疾患に苦しむ個体のDNAの同等領域との間の違いを検出すること；および

(5) 該気分障害もしくは関連疾患と関連する該遺伝子を同定すること、の工程を含んでなる気分障害もしくは関連疾患と関連する、その突然変異したおよび多型の変異体を包含する、少なくとも一つのヒト遺伝子を同定する方法を含んでなる。

#### 【0032】

クローニングしたYAC DNAをシーケンスする場合、関連する遺伝子の存在を確立するためにコンピューター解析を用いることができる。相同性検索およびエキソン予測のような技術を適用することができる。

#### 【0033】

本発明の方法に従っていったん候補遺伝子が単離されると、正常な個体からの遺伝子と双極性スペクトル障害のような気分障害に苦しむ個体からのものとの間でさらに詳細な比較を行うことができる。例えば、DNA配列における突然変異もしくは多型を同定することができる、「突然変異試験」と記述されている2つの方法がある。第一のものとして、DNAサンプルを一つの特定の突然変異の有無に関して試験することができるが、これは、突然変異が何であるかもしれないかという知識を必要とする。第二のものとして、DNAのサンプルを標準(正常)DNAからの任意の偏差に関してスクリーニングする。この後者の方法は、突然変異が前もって同定されない場合に候補遺伝子を同定するのにいっそう有用である。さらに、正常なもしくは健康な個体からの遺伝子と気分障害もしくは関連疾患に苦しむ個体からのものとの間の違いを同定するために上記の方法により同定される遺伝子に以下の技術をさらに適用することができる：

(a) サザンプロットイング技術：患者および健康な個体の異なる制限酵素で消化したゲノムDNAを含有するナイロン膜にクローンをハイブリダイズさせる。患者と健康な個体との間の大きな違いは、放射性ラベリングプロトコルを用いて視覚化することができる；

(b) ポリアクリルアミドゲルにおけるヘテロ二本鎖移動度：この技術は、非変性ポリアクリルアミドゲルにおけるヘテロ二本鎖の移動度がホモ二本鎖の移動度より小さいという

10

20

30

40

50

ことに基づく。これは200bp未満の断片に最も有効である；

(c) 一本鎖高次構造多型解析(SSCPもしくはSSCA)：一本鎖DNAは、弱い分子内結合により安定化される複雑な構造を形成するように折り畳まる。

【0034】

非変性ポリアクリルアミドゲル上のこれらの構造の電気泳動移動度は、それらの鎖長およびそれらの高次構造により決まる；

(d) ミスマッチ化学切断法(CCM)：放射性標識したプローブを試験DNAにハイブリダイズさせ、そしてミスマッチの部位でDNAの一方の鎖を切断する一連の化学反応によりミスマッチを検出する。これは非常に高感度の方法であり、そしてキロ塩基長のサンプルに適用することができる；

(e) ミスマッチ酵素切断法：該アッセイはCCMと同様であるが、ある種のバクテリオファージもしくは真核生物の酵素により切断を行う；

(f) 変性勾配ゲル電気泳動：この技術では、変性剤(化学物質もしくは温度)の増加する量の勾配がある電気泳動ゲルを通してDNA二本鎖を移動させる。鎖が融解しそして分離するゲル上の位置にDNA二本鎖が達するまで移動は続き、その後、変性したDNAはさらに移動しない。正常および突然変異体DNA二本鎖の間の単一塩基対の違いは、それらをゲルの異なる位置に移動させるのに十分である。

(g) ダイレクトDNAシーケンス。

【0035】

これらの適当なアッセイ技術のさらに詳細な説明を以下に提供する。

遺伝子型判定 本明細書において用いる場合、「遺伝子型判定」という用語は、CAP2をコードするポリヌクレオチドが位置942でチミジン(T)を含むかどうかを決定することを意味する。「遺伝子型判定」という用語は、「遺伝子検査」、「遺伝子スクリーニング」、「対立遺伝子もしくは多型を決定することもしくは同定すること」、「分子診断」のような用語または任意の他の同様の語句と同義である。

【0036】

適切なサンプルDNA配列におけるヌクレオチドの違いを識別することができる任意の方法もまた用いることができる。実際、多数の既知の異なる方法が、本発明のCAP2をコードするポリヌクレオチドの遺伝子型判定(すなわち、遺伝子型を決定すること)における使用に適している。これらの方法には、ダイレクトシーケンス、PCR-RFLP、ARM

遺伝子型スクリーニング 一つの態様として、本発明は、個体からのCAP2をコードするポリヌクレオチドを含んでなる核酸の遺伝子型スクリーニングの方法を提供する。個体からのCAP2をコードするポリヌクレオチドを含んでなる核酸の遺伝子型スクリーニングの方法は、その個体からの標的サンプルからの核酸の増幅を必要とし得る。

標的サンプル 本発明の標的サンプルは、解析する個体からのCAP2をコードするポリヌクレオチドを含んでなる任意の標的核酸であることができる。そのような核酸のアッセイには、実質的に任意の生物学的サンプル(真正(pure)赤血球以外)が適当である。例えば、都合の良い標的サンプルには、全血、白血球、精液、唾液、涙液、尿、糞便物質、汗、口腔(buccal)、皮膚および毛髪が包含されるがこれらに限定されるものではない。cDNAもしくはmRNAのアッセイには、標的サンプルは、典型的に、標的核酸が発現される細胞もしくは器官から得られる。

遺伝子型判定SNP 多数の異なる方法がSNPの遺伝子型を決定することにおける使用に適している。これらの方法には、ダイレクトシーケンス、PCR-RFLP、ARM

適切なDNA配列における単一ヌクレオチドの違いを識別することができる任意の方法もまた用いる

10

20

30

40

50

ことができる。

増幅 本明細書において用いる場合、「増幅」という用語は、鋳型特異性を伴う核酸複製を意味する。鋳型特異性は、「標的サンプル」もしくは「標的配列」特異性に関する。標的配列は、それらが他の核酸から選別されることが求められるという意味で「標的」である。その結果として、増幅技術は、主としてこれを選別するために設計されている。増幅方法の例には、ポリメラーゼ連鎖反応（PCR）、特定の対立遺伝子のポリメラーゼ連鎖反応（PASA）、リガーゼ連鎖反応（LCR）、転写増幅、自律配列複製および核酸に基づく配列増幅（NASBA）が包含されるがこれらに限定されるものではない。

Taqman 遺伝子型を決定するのに適当な手段は、TaqmanTM技術に基づくことができる。TaqmanTM技術は、以下の米国特許4,683,202; 4,683,195および4,965,188に開示されている。TaqmanTM対立遺伝子識別アッセイに含まれるウラシルN-グリコシラーゼの使用は、米国特許5,035,996に開示されている。

PCR PCR技術は当該技術分野において周知である（例えば、EP-A-0200362およびEP-A-0201184ならびに米国特許第4,683,195号および第4,683,202号を参照）。標的配列を増幅する工程は、所望の標的配列を含有するDNA混合物に過剰の2種のオリゴヌクレオチドプライマーを導入することおよびその後のDNAポリメラーゼの存在下での正確な一連の温度サイクリングからなる。PCRでは、例えばゲノムDNAにおける特定の標的配列の単一コピーを、いくつかの異なる方法論（標識したプローブでのハイブリダイゼーション、ビオチン化したプライマーの導入およびその後のアビジン-酵素結合体検出、ならびに増幅配列へのdCTPもしくはdATPのような32pで標識したデオキシヌクレオチド三リン酸の導入のような）による検出可能なレベルに増幅することが可能である。あるいはまた、差次的に標識し、従って各々検出することができるプライマーで異なる多型部位（マーカー）を増幅することが可能である。多数のマーカーを解析する一つの手段は、各マーカーを異なる蛍光プローブで標識することを含む。次にPCR生成物を蛍光に基づく自動シーケンサーで解析する。ゲノムDNAに加えて、任意のオリゴヌクレオチド配列をプライマー分子の適切な組で増幅することができる。特に、PCR工程自体によって生じる増幅断片は、それら自体が次のPCR増幅の効率のよい鋳型である。例として、PCRはまた、ヒトにおけるもしくはヒトからのCAP2をコードするポリヌクレオチドの適当な部分を増幅するプライマーを同定するために用いることもできる。

プライマー 本発明はまた、一連の有用なプライマーも提供する。

#### 【0037】

本明細書において用いる場合、「プライマー」という用語は、適切なバッファーにおいてそして適当な温度で適切な条件下で（すなわち、4種の異なるヌクレオチド三リン酸およびDNAもしくはRNAポリメラーゼまたは逆転写酵素のような重合のための因子の存在下で）鋳型指示DNA合成の開始点として作用することができる一本鎖オリゴヌクレオチドをさす。プライマーの適切な長さは、プライマーの意図される用途により決まるが、典型的には15~30ヌクレオチドである。短いプライマー分子は、一般に、鋳型と十分に安定なハイブリッド複合体を形成するのにいっそう冷たい温度を必要とする。プライマーは鋳型の正確な配列を反映する必要はないが、鋳型とハイブリダイズするために十分に相補的でなければならない。

#### 【0038】

「プライマー部位」という用語は、プライマーがハイブリダイズする標的DNAの領域をさす。

#### 【0039】

「プライマー対」という用語は、増幅するDNA配列の5'末端とハイブリダイズする5'上流プライマーおよび増幅する配列の3'末端の相補物とハイブリダイズする3'下流プライマーを含む一組のプライマーを意味する。

#### 【0040】

10

20

30

40

50

本発明のプライマーはDNAもしくはRNA、および一本鎖もしくは二本鎖であることができる。あるいはまた、プライマーは、天然に存在するかもしくは合成であることができるが、典型的には合成手段により製造される。

**プライマーハイブリダイゼーション条件** 本明細書において用いる場合、「ハイブリダイゼーション」という用語は、相補的な核酸の対合をさす。ハイブリダイゼーションおよびハイブリダイゼーションの強さ（すなわち、核酸間の会合の強さ）は、核酸間の相補性の程度、関与する条件のストリンジェンシー、形成されるハイブリッドの融解温度（ $T_m$ ）および核酸内のG:C比のような因子により影響を受ける。

【0041】

本明細書において用いる場合、「ストリンジェンシー」という用語は、核酸ハイブリダイゼーションを行う、温度、イオン強度および有機溶媒のような他の化合物の存在の条件に関して用いる。 10

【0042】

ハイブリダイゼーションは、典型的に、ストリンジェントな条件下で、例えば、1M以下の塩濃度および少なくとも25の温度で行う。例えば、5X SSPE（750mM NaCl、50mM リン酸Na、5mM EDTA、pH7.4）および25-30の温度の条件は、対立遺伝子特異的プライマーハイブリダイゼーションに適している。

**対立遺伝子特異的プライマー** 対立遺伝子特異的プライマーは、多型に重なる標的DNA上の部位にハイブリダイズし、そしてプライマーが完全な相補性を示す対立遺伝子型の増幅のみをプライミングする（Gibbs, Nucleic Acid Res. 17, 2427-2448 (1989)を参照）。このプライマーは、遠位部位でハイブリダイズする第二のプライマーとともに用いることができる。増幅はこれら2個のプライマーから進行し、特定の対立遺伝子型が存在することを示す検出可能な生成物をもたらす。コントロールは、一方が多型部位で単一塩基ミスマッチを示しそしてもう一方が遠位部位に完全な相補性を示すもう一つのプライマーで行うことができる。単一塩基ミスマッチは増幅を妨げ、そして検出可能な生成物は形成されない。多型と一列に並ぶオリゴヌクレオチドの最も3'の位置は、プライマーからの伸長に最も不安定であるので、この位置にミスマッチが含まれる場合に該方法は最もうまくいく（例えばWO 93/22456を参照）。 20

【0043】

単一塩基ミスマッチを検出するために特異的ハイブリダイゼーションが可能なハイブリダイゼーションプローブは、当該技術分野において既知でありそしてManiatis et al. Molecular Cloning: A Laboratory Manual, 2nd Ed (1989) Cold Spring Harbourに記述されている方法に従って設計することができる。 30

【0044】

(i) PCRプライマー 好ましくは、スクリーニングは、ヌクレオチド942を含むヒトのCAP2をコードするポリヌクレオチド（遺伝子）の部分を増幅するように設計されたPCRプライマーを用いて実施する。

【0045】

そのようなPCRプライマーの例を配列番号9および10として示す。 40

**増幅した標的配列における多型の検出** 増幅した核酸配列は、対立遺伝子特異的プローブ、タイピングアレイ、ダイレクトシーケンス、変性勾配ゲル電気泳動および一本鎖高次構造多型（SCCP）解析が包含されるがこれらに限定されるものではない方法を用いて検出することができる。

**対立遺伝子特異的プローブ** 二人の個体からのそれぞれの断片における異なる多型の型の存在のために一方の個体からの標的DNAの断片にハイブリダイズするがもう一方の個体からの対応する断片にハイブリダイズしない対立遺伝子特異的プローブを設計することができる。

【0046】

本明細書において用いる場合、「プローブ」という用語は、精製された制限消化産物にお 50

けるように天然に存在しようともしくは合成で製造しようとして、対象の別のオリゴヌクレオチド配列にハイブリダイズすることができるオリゴヌクレオチド(すなわち、ヌクレオチドの配列)をさす。プローブは、特定の遺伝子配列の検出、同定および単離において有用である。本発明のハイブリダイゼーションプローブは、典型的に、核酸の相補鎖に塩基特異的に結合することができるオリゴヌクレオチドである。

#### 【0047】

本発明のプローブは、酵素(例えば、ELISA、ならびに酵素に基づく組織化学アッセイ)、蛍光、放射性および発光系が包含されるがこれらに限定されるものではない任意の検出系においてそれが検出可能であるように任意の「レポーター分子」で標識することができる。対象の標的配列(すなわち、検出される配列)もまた、レポーター分子で標識することができる。本発明は、任意の特定の検出系もしくは標識に限定されない。

10

#### 【0048】

本発明のプローブに選択するハイブリダイゼーション条件は、対立遺伝子間でハイブリダイゼーション強度に有意な差、好ましくは本質的に2値の応答(binary response)があり、それによりプローブが対立遺伝子の一方のみにハイブリダイズするように十分にストリンジェントである。典型的なハイブリダイゼーション条件は、1塩基対ミスマッチを決定することができるように本発明の対立遺伝子特異的プライマーに上記に示したようなストリンジェントな条件である。

タイリングアレイ 本発明の多型はまた、核酸アレイへのハイブリダイゼーションにより同定することもでき、そのいくつかの例はWO 95/11995に記述されている。「タイリング」という用語は、一般に、解析する配列(「標的配列」)に相補的な配列ならびにその配列のあらかじめ選択したバリエーションで構成されている特定の1組のオリゴヌクレオチドプローブの合成を意味する。これらのバリエーションには、通常、1個もしくはそれ以上のヌクレオチドでの1個もしくはそれ以上の塩基位置での置換が含まれる。

ダイレクトシーケンス 本発明の多型の配列の直接解析は、ジデオキシ連鎖停止法もしくはマクサム・ギルバート法(Sambrook et al., Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989を参照)のいずれかを用いて、または例えば、ABI Prism 377 DNAシーケンサーで解析するBig Dyeターミネーターサイクルシーケンス化学を用いる標準的なABIシーケンス技術を用いて実施することができる。

好ましくは、本発明のアッセイに用いる多型は、同定された配列のPstI制限解析で生じる断片の有無により同定する。

20

30

1.5変性勾配ゲル電気泳動 PCRを用いて生成される本発明の増幅生成物はまた、変性勾配ゲル電気泳動の使用により解析することもできる。異なる対立遺伝子は、溶液中のDNAの異なる配列依存的融解特性および電気泳動移動度に基づいて同定することができる。Erllich, ed., PCR Technology, Principles and Applications for DNA Amplification, (W. H. Freeman and Co, New York, 1992), 第7章。

一本鎖高次構造多型(SSCP)解析 本発明の標的配列の対立遺伝子はまた、Orta et al., Proc. Nat. Acad. Sci. 86, 2766-2770 (1989)に記述されているように、一本鎖PCR生成物の電気泳動移動度の変化により塩基の違いを同定する一本鎖高次構造多型(SSCP)解析を用いて識別することもできる。増幅PCR生成物を上記のように生成せしめ、そして加熱するかもしくはそうでなければ変性させて、一本鎖増幅生成物を生成せしめることができる。一本鎖核酸は、塩基配列に部分的に依存する二次構造を再生するかもしくは形成することができる。一本鎖増幅生成物の異なる電気泳動移動度は、標的配列の対立遺伝子間の塩基配列の違いに関係することができる。

40

試験配列とコントロール配列との間の違いを同定すること 増幅された核酸配列のこれらの検出方法は、基準核酸配列と試験核酸配列との間の変異の1つもしくはそれ以上の点の違いを同定するためにまたは2人もしくはそれ以上の個体からのCAP2遺伝子の異なる

50

多型の型を比較するために用いることができる。

基準核酸配列 本明細書において用いる場合、「基準核酸配列」は、そのアッセイにおいて試験する対立遺伝子の各々のホモ接合体である一人もしくはそれ以上の個体に相当するコントロールDNA配列のようなコントロール核酸配列を意味する。例として、コントロールDNA配列には：(i)ホモ接合体の個体からのゲノムDNA；(ii)ホモ接合体の個体から増幅される関連したSNPを含有するPCR生成物；または(iii)プラスミドもしくは他の適当なベクターにクローニングされている関連したSNPを含有するDNA配列を包含することができるがこれらに限定されるものではない。コントロールサンプルはまた、対立遺伝子の既知の組からの複数の対立遺伝子を含んでなる対立遺伝子ラダーであることもできる。各々異なる対立遺伝子もしくは対立遺伝子の組を含有する複数のコントロールサンプルがあることができる。他の基準/コントロールサンプルには、典型的に、図表示、文書表示、鋳型またはPCR生成物もしくは核酸の他の断片における多型の存在を同定するのに適当な任意の他の手段が包含される。「基準核酸配列」、基準サンプルおよびコントロールサンプルという用語は、本文の全体にわたって互換的に用いる。

H. 治療用途 本発明の一つの態様は、双極性気分障害にかかりやすい素因に関して個体をスクリーニングすること、およびポリヌクレオチド素因が同定される場合に、双極性気分障害を遅らせるかもしくは軽減するかもしくは防ぐためにその個体を処置することを提供する。

10

#### 【0049】

本発明のこの態様の実施形態として、双極性気分障害にかかりやすい個体の素因は、上記に説明する検出の方法を用いて個体がヌクレオチド942がチミジン(T)であるCAP2をコードするポリヌクレオチドのホモ接合体であるか、位置942のグアノシン(G)がチミジン(T)で置換されるCAP2をコードするポリヌクレオチドのヘテロ接合体であるか、もしくはヌクレオチド942がグアノシン(G)であるCAP2をコードするポリヌクレオチドのホモ接合体であるかを決定することにより評価する。

20

#### 【0050】

従って、該多型に関して、位置942でT/Tホモ接合体である個体は、危険性が最も高いと分類される。G/Tヘテロ接合体である個体は、中程度の危険性を有すると分類される。G/Gホモ接合体である個体は、最も低い危険性のカテゴリーにあると分類される。

#### 【0051】

場合により、個体の危険因子の評価は、CAP2をコードするポリヌクレオチド多型の存在並びに既知のポリヌクレオチド指標もしくは生理学的指標もしくは他の指標の両方を参照することにより計算される。このように本発明は、個体の危険性の測定が基づくことができるさらなる情報を提供する。

30

一般的な方法論参考文献 一般に、本明細書に記述する技術は当該技術分野において周知であるが、特にSambrook et al., Molecular Cloning, A Laboratory Manual (1989)およびAusubel et al., Short Protocols in Molecular Biology (1999) 4th Ed, John Wiley & Sons, Inc.を参照することができる。

40

#### 【0052】

本明細書に記述する方法に関して、YACからのコーディング領域と気分障害もしくは関連疾患に苦しむ個体のDNAとの間の違いを検出する工程において、該個体はその疾患にかかっている誰でもあることができ、そして必ずしも家系MAD31の一員であるとは限らないと理解される。

#### 【0053】

さらなる態様によれば、本発明は、気分障害もしくは関連疾患と関連しそして上記の方法のいずれかにより得ることができる単離されたヒト遺伝子およびその変異体、該遺伝子によりコードされる単離されたヒトタンパク質ならびに該タンパク質をコードするcDNAを提供する。

50

## 【0054】

いったん遺伝子が同定されると、コードされるタンパク質の機能を決定するために多数の方法が利用可能である。これらの方法は、Eisenberg et al (Nature vol. 15, June 2000)により記述されており、そして引用することにより本明細書に組み込まれる。一つの方法は、ゲノム配列から機能連鎖を示す算定法を含み、そして遺伝子近接法 (gene neighbor method) と呼ばれる。いくつかのゲノムにおいて、2個のタンパク質をコードする遺伝子が染色体上で近接する場合、これらのタンパク質は機能的に関連している傾向がある。この方法は、オペロンが一般的である原核生物における機能連鎖を明らかにすることにおいて有効であることができるが、真核生物における相互作用するタンパク質を解析することにも効果を発揮する。

10

CAP-2 遺伝子

17.1 kbのゲノムにわたる5' UTRおよび6個のコーディングエクソンを含有する細胞質アンチプロテイナーゼ2遺伝子 (CAP2) の完全なイントロン-エクソン構造を本明細書に開示する。イントロンのサイズを測るために、CAP2 cDNA配列にわたるプライマーの異なる組み合わせを用いる。このようにして、5個のイントロンのサイズおよびエクソン-イントロン境界配列を得た。スプライス連結部の5' 供与および3' 受容部位は、共通配列と相関関係があった (表1)。第一の5' UTRエクソンは非常に小さく (73 bp)、そして多型であるがMAD31ベルギー人家系においても罹患およびコントロール集団においても増加していないと判明した (CAG)<sub>2</sub> (CGG)<sub>6</sub> (CAG)<sub>6</sub> 配列を含有する。

20

## 【0055】

CAP2から得られるアミノ酸配列は、胎盤トロニンインヒビターもしくはPI6 (68%の同一性) およびプロテイナーゼインヒビター9 PI9 (63%) を包含する細胞質セルピンのオボアルブミンファミリーの他のヒトメンバーに高度の同一性を示す。5種のエクソン多型が同定され、これらから3種はアミノ酸変化をもたらす。PI6およびPI9のアミノ酸配列とCAP2の推定一次構造との整列により、アミノ酸位置68で、CAP2はArgもしくはGlnのいずれかであることができ、そしてPI6ではGlnであることが示された。同様に、アミノ酸位置359で、CAP2はArgもしくはHisのいずれかを示し、そしてPI6はHisを示す<sup>15</sup>。さらに、アミノ酸位置304で、CAP2はThrもしくはAlaのいずれかを示し、そしてPI9はAlaを示す。逆に、アミノ酸位置314で、CAP2はAlaを示し、一方、PI6およびPI9はValを示す。

30

## 【0056】

DHPLCにより検出されるこれらの変異体のうち2種 (c. 910A > G および c. 942C > T) は、SSCP解析により以前に同定されなかった。

## 【0057】

ベルギー系の75/75症例-コントロールサンプルにおける関連解析を適用した。症例およびコントロールは、民族性、性別、年齢に関して厳密に一致させた。6種のSNPの対立遺伝子および遺伝子型頻度の比較により、6種のSNPのうち5種で患者とコントロールとの間に有意な関連は示されなかった。エクソン7におけるSNP c. 942C > T置換の頻度は、BP患者とコントロールとの間で有意に異なり: BP患者は、コントロールと比較した場合にT対立遺伝子の頻度が高かった (p = 0.03)。6種のSNPは同じ遺伝子内に位置するが、これらのうち4種は連鎖不平衡になかった。SNP c. 203G > AとSNP IVS4 + 98A > Gとの間に非常に強いLDがあったが、コントロールにおいてのみであった。BP症例では、これら2種のSNP間のLDは弱かった。さらに、CAP2-CAG-CGG反復とCAP2 SNPとの間でLDは見出されなかった。BP障害およびCAP2 SNPで関連の強い証拠が見出されなかったことは、我々の集団における有意なLD結果の欠如と一緒に、CAP2がBP障害の病因において主要な役割を果たしていないかもしれないことを示唆する。

40

## 【0058】

50



家系 M A D 3 1 の 1 人 の 罹 患 個 体 にお いて、D 1 8 S 6 8 と D 1 8 S 9 6 9 と の 間 で 近 位 組 換 え が 起 こ っ た<sup>1 1</sup>。C A P 2 遺 伝 子 は こ れ ら の 2 種 の マ ー カ ー の 間 に 位 置 す る。C A P 2 - S N P は、こ の 近 位 組 換 え が 遺 伝 子 の 下 流 で 起 こ っ た こ と を 示 し た。

【実施例 1】

【0059】

A. 家系、患者およびコントロール被験体

B P I I 発 端 者 を 有 す る ベ ル ギ ー 人 家 系、M A D 3 1 に お け る 家 系 お よ び 臨 床 診 断 は 他 の 所 で 詳 細 に 記 述 さ れ た<sup>1 0</sup>。簡 潔 に 言 え ば、家 系 M A D 3 1 に お け る 異 な る 臨 床 診 断 は 以 下 の と お り で あ る：1 B P I、2 B P I I、2 U P、4 大 鬱 病 性 障 害 ( M D D )、1 統 合 失 調 性 躁 病 ( S A m ) お よ び 1 統 合 失 調 性 鬱 病 ( S A d )。

10

【0060】

症 例 - コ ン ト ロ ー ル サ ン プ ル は、ブ リ ュ ッ セ ル の E r a s m e 病 院 で 確 か め ら れ た ベ ル ギ ー 系 の 7 5 人 の 関 係 の な い 患 者、な ら び に 病 院 に お け る 告 示 に よ っ て 募 集 さ れ た 7 5 人 の 年 齢、性 別 お よ び 民 族 性 の 一 致 す る コ ン ト ロ ー ル 被 験 体 か ら な っ た。全 て の コ ン ト ロ ー ル 個 体 は、精 神 医 学 症 状 を 除 く た め に 面 接 し た。患 者 は、B P 障 害 の 研 究 診 断 基 準 ( R e s e a r c h D i a g n o s t i c C r i t e r i a )<sup>1 6</sup> を 満 た し た。

B. P C R 増 幅

ゲ ノ ム D N A お よ び c D N A は、C A P 2 c D N A 配 列 ( G e n B a n k 受 託 番 号 L 4 0 3 7 7 )<sup>1 5</sup> に わ た る 6 つ の 重 複 す る プ ラ イ マ ー 組 を 用 い て 増 幅 し た。約 5 0 n g の ゲ ノ ム D N A も し く は 1 n g の c D N A お よ び 1 0 p g の 各 プ ラ イ マ ー を 標 準 的 な P C R 反 応 に お いて 使 用 し た。増 幅 条 件 は 以 下 の と お り で あ っ た：9 4 で 4 分 間 の 最 初 の 変 性 工 程、続 い て 1 分 間 9 4、1 分 間 5 5、2 分 間 7 2 で 3 5 サ イ ク ル、お よ び 7 2 で 1 0 分 間 の 最 後 の 伸 長 期 間。

20

C. サザンプロット解析

家 系 M A D 3 1 の 3 人 の 罹 患 メ ン バ ー お よ び 2 人 の 非 罹 患 メ ン バ ー か ら の ゲ ノ ム D N A を H i n d I I I お よ び B a m H I で 別 個 に 消 化 し、そ し て 1 % ア ガ ロ ー ス ゲ ル 上 で 泳 動 し た。サ ザ ン プ ロ ッ テ ィ ン グ は、標 準 プ ロ ト コ ル<sup>1 7</sup> に 従 っ て 実 施 し た。

【0061】

5 0 n g の C A P 2 c D N A を ラ ン ダ ム プ ラ イ ム ラ ベ リ ン グ ( G i b c o - B R L ) に よ り ( - <sup>3 2</sup> P ) d C T P で 標 識 し た。ハ イ プ リ ダ イ ゼ ー シ ョ ン は、6 5 で C h u r c h バ ッ フ ェ ー に お いて 一 晩 実 施 し た。次 に、膜 を 6 5 で 1 X S S C、0 . 1 % S D S に お いて 1 回、0 . 5 X S S C、0 . 1 % S D S に お いて 1 回、そ し て 0 . 1 X S S C、0 . 1 % S D S に お いて 2 回 洗 浄 し、続 い て K o d a k X 線 フ ィ ル ム に - 7 0 で 7 2 時 間 露 出 し た。

30

D. S S C P お よ び D H P L C 解 析

P C R 増 幅 し た D N A は、あ ら か じ め 成 形 さ れ た す ぐ に 使 用 で き る ゲ ル お よ び H y d r o l i n k 5 % グ リ セ ロ ー ル ゲ ル ( P h a r m a c i a B i o t e c h ) で D N A 解 析 シ ス テ ム を 用 い て S S C P に よ り 解 析 し た。

【0062】

変 性 高 速 液 体 ク ロ マ ト グ ラ フ ィ ー ( D H P L C ) は、T r a n s g e n o m i c ( S a n t a C l a r a , C A , U S A ) か ら 購 入 し た 自 動 機 器 で 実 施 し た。粗 P C R 生 成 物 を D N A S e p カ ラ ム 上 に 載 せ、そ し て 0 . 9 m l / 分 の 一 定 の 流 速 で、0 . 1 M ト リ エ チ ル ア ミ ン ア セ テ ー ト バ ッ フ ェ ー ( T E A A )、p H 7 に お け る ア セ ト ニ ト リ ル 勾 配 を 用 い て カ ラ ム か ら 溶 出 し た。勾 配 は、溶 離 剤 A お よ び B を 混 合 す る こ と に よ り 作 製 し た。溶 離 剤 A は 0 . 1 M T E A A、0 . 1 M N a <sub>4</sub> E D T A で あ っ た。溶 離 剤 B は 0 . 1 M T E A A 中 2 5 % の ア セ ト ニ ト リ ル で あ っ た。ヘ テ ロ ニ 本 鎖 分 子 の 成 功 し た 分 離 に 必 要 な 勾 配 お よ び 温 度 は、W a v e m a k e r バ ー ジ ョ ン 3 . 4 . 4 に よ り 予 測 し た。

40

E. D N A シ ー ケ ン ス

シ ー ケ ン ス は、製 造 業 者 の プ ロ ト コ ル に 従 っ て、P e r k i n - E l m e r A B I 3 7 7 自 動 シ ー ケ ン サ ー お よ び B i g D y e タ ー ミ ネ ー タ ー サ イ ク ル シ ー ケ ン ス キ ャ ッ ト (

50

Applied Biosystems, PE)を用いてプラスミドDNAもしくはゲル精製したPCR鑄型に行った。PCR断片は最初にアガロースゲル上で視覚化し、そして次にUltrafree-DNAフィルター装置(Millipore)を用いてゲル精製した。

#### F. パイロシーケンス

ビオチン化したPCR生成物をストレプトアビジンで被覆した常磁性ビーズ(Dynal AS, Oslo, Norway)上に固定した。ssDNAは、固定したPCR生成物を50 $\mu$ lの0.5M NaOHにおいて5分間インキュベーションし、続いて100 $\mu$ lの10mM Tris-アセテート pH7.6において2回連続して洗浄することにより得た。プライマー エキソン7-1025(5'-GTG CCT CTG TCC AAG GTT GC-3')は、エキソン7におけるSNP c.942の検出のためのパイロシーケンスプライマーとして用いた。プライマーアニーリングは、72 $^{\circ}$ Cで2分間そして次に室温で5分間のインキュベーションにより行った。パイロシーケンスは、PSQ96パイロシーケンサー(Pyrosequencing AB, Uppsala, Sweden)で行った。

10

#### G. 統計解析

BP症例およびコントロールの全対立遺伝子および遺伝子型分布を比較し、そしてGenePop<sup>18</sup>を用いてハーディー・ワインバーグ平衡を試験した。対立遺伝子および遺伝子型特異的比較は、カイ二乗解析もしくは必要に応じてフィッシャーの直接確率検定を用いて行った。遺伝子型判定した全てのSNPのデータを合わせる複数点関連解析にはDisMUltプログラム<sup>19</sup>を用いた。連鎖不平衡(LD)は、GenePopのLinkD<sub>os</sub>を用いて計算した。患者およびコントロールにおけるハプロタイプ頻度を概算するために、最尤法に基づくArlequinアルゴリズムを適用した<sup>20</sup>。

20

#### H. CAP2遺伝子のゲノム構造

我々は、CAP2遺伝子のゲノム構造を決定した。第一に、家系MAD31の5人の選択したメンバー(3人の罹患および2人の非罹患)からのHindIIIでそしてBamHIで消化したゲノムDNAを含有する2枚の異なるサザンプロットに対してPCRで得られたCAP2 cDNA断片をハイブリダイズさせることによりCAP2遺伝子をゲノム再編成に関して解析した。観察されるハイブリダイズしたバンドに基づき、25kbの最小ゲノムサイズが概算された。罹患および非罹患個体のハイブリダイゼーションパターン

30

#### 【0063】

cDNAプライマーを用いて、ゲノムDNAに対するPCRからイントロンの位置およびサイズが得られた。シーケンスの後、cDNAおよびゲノム配列の比較により正確なエキソン-イントロン境界を決定した(表1)。これらのゲノム配列からイントロンプライマーを設計した(表4)。

#### 【0064】

#### 【表1】

表4 イントロンCAP2プライマー

## エキソン3のPCRプライマー

フォワード 5' ACTTTCAAT TTCTTTGTCATC 3' (配列番号3)  
 リバース 5' TACAAAGCAGGAGATATTCACC 3' (配列番号4)

## イントロン4のPCRプライマー

フォワード 5' GAAGCATATAAATGACTGGGTG 3' (配列番号5)  
 リバース 5' GATAAGAAATGACAGAGTTGC 3' (配列番号6)

10

## エキソン5のPCRプライマー

フォワード 5' CCAAGAGAATATTTCTCTG 3' (配列番号7)  
 リバース 5' AGTCGATCCCCTGACAAAGC 3' (配列番号8)

## エキソン7のPCRプライマー

フォワード 5' AGCTGGAGGAGAGTTATGACTT 3' (配列番号9)  
 リバース 5' GCAAGATAGGTAGAAGGAAAGG 3' (配列番号10)

20

## 【0065】

この解析により、CAP2は73bp(エキソン1)~405bp(エキソン7)におよぶサイズを有する1個の非コーディングエキソンおよび6個のコーディングエキソンを含有することが示された(図1および表1)。イントロン2~6のサイズはPCRにより決定し、そして1.3Kb(イントロン2)~1.8Kb(イントロン3)であった(図1)。これらの実験が進行中であった間に、CAP2遺伝子を含有するBAC 793J2の完全な配列が利用できるようになった(Genbank受託番号AC009802)。エキソン-イントロン境界配列、イントロンおよびエキソンサイズを確かめ、そしてイントロン1のサイズを8.1kbと決定した。全体で、CAP2遺伝子は17.1kbのゲノム領域にわたる。

30

## 【0066】

CAP2遺伝子の方向を定めるために、CAP2-CAG断片化YAC<sup>21</sup>をL40377(CAP2エキソン7)を含む遺伝子に対してセントロメアおよびテロメアのSTSマーカーの存在に関して解析した。PCR解析により、CAP2に対してセントロメアのマーカーでの陽性ヒットおよび遺伝子に対してテロメアのマーカーでの増幅の欠如が示され、CAP2の転写方向がセントロメアからテロメアへであることが示された。

## 【0067】

## 【表2】

表1. CAP2におけるイントロン-エキソン境界

エキソン番号	サイズ(bp)	スプライス番号	スプライス供与
1	73		GCAGCAGGAG/gtgggggcct
2	178	tttgatgcag/ACCTTCTCTG	GATGTCCCAG/gtatgtgtgc
3	138	tttgatgcag/ACCTTCTCTC	TTCTTCCAG/taagtagtat
4	118	gtgttcag/GACTTTAAAGA	AAGACTGAAG/gtgagacagt
5	143	tctttatag/GTAAGATTTC	AACCAACGAG/gtaggaaag
6	153	ttccgtag/GAAAAAAGA	CCTCGCCGTG/gtaagctcca
7	405	cttatcctag/GTGGAAAAG	TTCTCCGTAA

10

## 【0068】

## 突然変異検出および解析

24人のBP患者のDNAから全てのエキソンをPCR増幅するためにイントロンプライマーを設計した。SSCP解析を用いてPCR生成物を突然変異に関してスクリーニングした。2種の非同義SNPがc.203G>A(Arg68Gln)およびc.1076G>A(Arg359His)で同定された。Leuをコードする1種と同義SNPがc.477A>G(コドン159)で同定された。さらに、1種のSNPがイントロン4、IVS4+98A>Gにおいて検出された。これらの結果はDHPLC解析により確かめられ、そしてc.910A>G(Thr304Ala)およびAlaをコードするc.942C>T(Ala314Ala)におけるものの、エキソン7における2種の追加のSNPの同定をもたらした(表2)。

20

## 【0069】

## 【表3】

30

表2. CAP2多型

多型の位置	ヌクレオチド位置 cDNA	制限部位 変化	タンパク質位置 コドン	アミノ酸 変化
エキソン3	c.203G>A		68	Arg → Gln
エキソン4	IVS4+98A>G	<i>Rsa</i> Iの獲得		
エキソン5	c.477A>G	<i>Mae</i> Iの喪失	159	Leu → Leu
エキソン7	c.910A>G	<i>Pvu</i> IIの獲得	304	Thr → Ala
エキソン7	c.942C>T		314	Ala → Ala
エキソン7	c.1076G>A	<i>Hha</i> Iの喪失	359	Arg → His

10

20

## 【0070】

75人の関係のない双極性患者および75人の一致するコントロールにおけるPCR-RFLP解析をこれらの変異体のうち3種について行った：SNP IVS4+98A>Gには*Rsa* I-RFLPアッセイを適用した。SNP c.910A>Gには*Pvu* II-RFLP解析を適用した。SNP c.1076G>Aには*Hha* I-RFLP解析を適用した。SNP c.477A>Gおよびc.203G>Aは、ゲノムDNAから作製したPCR断片のダイレクトシーケンスにより解析した。SNP c.942C>Tを解析するためにはパイロシーケンスを用いた。

## 【0071】

これらのSNPのうち5種において対立遺伝子頻度もしくは遺伝子型分布におけるBP患者とコントロールとの間の有意な差はなかった。しかしながら、SNP c.203G>A ( $p = 0.05$ ) およびc.477A>G ( $p = 0.03$ ) のハーディー・ワインバーグ平衡のわずかな偏差が、両方ともBPにおいてあり、これは過剰のヘテロ接合体によってもたらされる ( $ex3$ では $p = 0.03$ ;  $ex5$ では $p = 0.02$ )。さらに、コントロールにおいてSNP c.203G>Aのわずかに過剰のヘテロ接合体がある ( $p = 0.04$ )

30

SNP c.942C>TのT対立遺伝子は、コントロール(1%)におけるよりBP症例(6%)において有意に高い頻度を有した ( $\chi^2 = 4.83$ ;  $p = 0.03$ )。遺伝子型を比較する場合、2/75(3%)のコントロールと比較して9/73(12%)が1個のT対立遺伝子を有した ( $\chi^2 = 5.02$ ;  $p = 0.03$ )。興味深いことに、性別の層化後にデータを解析した場合、有意な差が認められた。男性では、T対立遺伝子はBP患者において8%の頻度を有し、一方、コントロールにおいてそれは認められなかった (フィッシャーの直接確率検定、 $p = 0.03$ )。女性では、対立遺伝子もしくは遺伝子型分布における違いは症例とコントロールとの間で認められなかった。

40

## 【0072】

これらの結果を確かめるために、関係のない患者および一致するコントロール群を拡張し、そして113人のBP患者および163人の年齢、性別および民族性の一致するコントロールにおいて最終的な遺伝子型判定関連解析を行った。表3は、患者およびコントロール集団におけるこれらの多型の対立遺伝子および遺伝子型頻度を示す。

50

【 0 0 7 3 】

【 表 4 】

表3. CAP2遺伝子におけるSNPの遺伝子型および対立遺伝子頻度

SNP	遺伝子型							対立遺伝子		
	症例			コントロール				BP	コントロール	
	AA	AG	GG	AA	AG	GG	p値	A	A	p値
	%	%	%	%	%	%		%	%	
c.203G>A	2	48	50	2	49	49	0.90	26	27	0.81
IVS4+98A>G	71	29	0	68	32	0	0.63	86	84	0.66
c.477A>G	5	47	48	8	43	49	0.76	28	30	0.77
c.910A>G	80	19	1	79	20	1	1.00	89	89	1
c.942C>T	0	10	90	0	3	97	<b>0.02</b>	5	1	<b>0.02</b>
c.1076G>A	21	43	36	17	50	33	0.93	43	42	0.93

10

20

【 0 0 7 4 】

【 表 5 】

## 参考文献リスト

以下の公開は、引用することにより本明細書に組み込まれる。

1. Akiskal, H.S. *et al.* Psychopathology, temperament, and past course in primary major depressions. 1. Review of evidence for a bipolar spectrum. *Psychopathology* **22**, 268-277 (1989).
2. Tsuang MT & Faraone SV. The Genetics of Mood Disorders. The Johns Hopkins University Press, Baltimore (1990). 10
3. Spence, M.A. *et al.* Bipolar disorder: evidence for a major locus [see comments]. *Am. J. Med. Genet.* **60**, 370-376 (1995).
4. Conneally PM. Genetics and Psychiatric Disorders. (1998).
5. McInnis, M.G. *et al.* Anticipation in bipolar affective disorder. *Am. J. Hum. Genet.* **53**, 385-390 (1993).
6. Nylander, P.O., Engstrom, C., Chotai, J., Wahlstrom, J. & Adolfsson, R. Anticipation in Swedish families with bipolar affective disorder. *J. Med. Genet.* **31**, 686-689 (1994). 20
7. Li, R. & el-Mallakh, R.S. Triplet repeat gene sequences in neuropsychiatric diseases. *Harv. Rev. Psychiatry* **5**, 66-74 (1997).
8. Van Broeckhoven, C. & Verheyen, G. Chromosome 18 workshop. *Psychiatr. Genet.* **8**, 97-108 (1998).
9. Van Broeckhoven, C. & Verheyen, G. Report of the chromosome 18 workshop. *Am. J. Med. Genet.* **88**, 263-270 (1999). 30
10. De bruyn, A., Souery, D., Mendelbaum, K., Mendlewicz, J. & Van Broeckhoven, C. Linkage analysis of families with bipolar illness and chromosome 18 markers. *Biol. Psychiatry* **39**, 679-688 (1996).
11. Verheyen GR *et al.* Genetic refinement and physical mapping of a chromosome 18q candidate region for bipolar disorder. *Eur. J. Hum. Genet.* **7**, 427-434 (1999).
12. Huber, R. & Carrell, R.W. Implications of the three-dimensional structure of alpha 1-antitrypsin for structure and function of serpins. *Biochemistry* **28**, 8951-8966 (1989). 40
13. Remold-O'Donnell, E. The ovalbumin family of serpin proteins. *FEBS Lett.* **315**, 105-108 (1993).
14. Bartuski, A.J., Kamachi, Y., Schick, C., Overhauser, J. & Silverman, G.A. Cytoplasmic antiproteinase 2 (PI8) and bomapin (PI10) map to the serpin cluster at 18q21.3. *Genomics* **43**, 321-328 (1997).

15. Sprecher, C.A. *et al.* Molecular cloning, expression, and partial characterization of two novel members of the ovalbumin family of serine proteinase inhibitors. *J. Biol. Chem.* **270**, 29854-29861 (1995).
16. Endicott J & Spitzer RL. Schedule for Schizophrenia and Affective Disorders-Lifetime Version. New York: N.Y. state Psychiatric Institute., N.Y. (1978).
17. Sambrook J, Fritsch E & Maniatis T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, N.Y. (1989).
18. Raymond, M. & Rousset, F. GENEPOP (version 1.2): a population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249 (1995). 10
19. Terwilliger, J.D. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56**, 777-787 (1995).
20. Schneider S, Roessli D & Excoffier L. Arlequin ver. 2.000: A software for population genetics data analysis. (2000).
21. Goossens, D. *et al.* No evidence for the involvement of CAG/CTG repeats from within 18q21.33- q23 in bipolar disorder. *Eur. J. Hum. Genet.* **8**, 385-388 (2000). 20

【図面の簡単な説明】

【0075】

【図1】18q21.33-q23 BP候補領域 (Verheyen *et al.* 1999) の最小のYACタイリング経路。YACを実線で、CCG/CGG断片化生成物を点線で表す。括弧の間のYACサイズは、PFGE解析により概算される。詰まった丸は、陽性のSTS/STRヒットを示す。斜線付きボックスは、YAC断片化により単離されたCCG/CGG反復および3つのCpGアイランドを強調する。 30

【図2】細胞質アンチプロテイナーゼ2 (CAP2) 遺伝子のゲノム構造。黒色のボックスはエキソンを表し、そしてbp単位のそれらのサイズをボックスの上を示す。イントロンサイズはkb単位である。組み合わせられたCAG-CGG反復を示す。転写開始および終始コドンを示す。



## 【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
27 March 2003 (27.03.2003)

PCT

(10) International Publication Number  
WO 03/025222 A1

- (51) International Patent Classification: C12Q 1/68, C12N 15/15
- (21) International Application Number: PCT/EP02/10667
- (22) International Filing Date: 17 September 2002 (17.09.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 01203558.0 17 September 2001 (17.09.2001) EP
- (71) Applicant (for all designated States except US): JANSSEN PHARMACEUTICA N.V. [BE/BE]; Turnhoutseweg 30, B-2340 Beerse (BE).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): DEL-FAVERO, Jorgen, Peter, Lode [BE/BE]; Hlanders Interuniversity Institute for Biotechnology (VIB), Rijvissechstraat 120, B-9052 Zwijnaarde (BE); VAN BROECKHOVEN, Christine [BE/BE]; Hlanders Interuniversity Institute for Biotechnology (VIB), Rijvissechstraat 120, B-9052 Zwijnaarde (BE).
- (74) Common Representative: JANSSEN PHARMACEUTICA N.V.; Turnhoutseweg 30, B-2340 Beerse (BE).
- (81) Designated States (national): AE, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HN, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LI, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Declarations under Rule 4.17:**  
— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GF, GH, GM, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MY, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LI, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).  
— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations of inventorship (Rule 4.17(iv)) for US only
- Published:**  
with international search report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/025222 A1

(54) Title: BRAIN EXPRESSED CAP-2 GENE AND PROTEIN ASSOCIATED WITH BIPOLAR DISORDER

(57) Abstract: We previously identified 18q21.33-q23 as a candidate region for bipolar (BP) disorder and constructed a yeast artificial chromosome (YAC) contig map. In a next step we isolated and analysed all CAGACTG repeats from this region and excluded them from involvement in BP disorder. Here, in the process of identifying all CCGCGG repeats from the region, we found a cluster of 6 genes encoding for serpins. For one of them, CAP2 located at 18q21.3, a mutation analysis was performed. Analysis of the thus identified single nucleotide polymorphisms (six in total) revealed a statistically significant association of SNP c.942G>T with BP disorder. It is thus an object of the present invention to provide a method of diagnosing BP or susceptibility to BP in an individual which method comprises determining, in a sample from the individual, the single nucleotide polymorphism in the CAP2 gene of the individual, and determining the status of the individual by reference to polymorphism in the CAP2 gene.

WO 03/025222

PCT/EP02/10667

-1-

**Novel Brain Expressed CAP-2 Gene and Protein associated with Bipolar Disorder**

---

5 **FIELD OF THE INVENTION:**

The invention is broadly concerned with the determination of genetic factors associated with psychiatric health. More particularly, the present invention is directed to a human gene which is linked to a mood disorder or related disorder in affected individuals and their families. Specifically, the present invention is directed to a gene encoding cytoplasmic antipeptidase 2 (CAP2). The gene is located on the eighteenth chromosome and is expressed in brain tissue and can be used as a diagnostic marker for bipolar disorder.

15 **BACKGROUND OF THE INVENTION:**

**Pharmacogenetics background:**

Every individual is a product of the interaction of their genes and the environment.

20 Pharmacogenetics is the study of how genetic differences influence the variability in patients responses to drugs. Through the use of pharmacogenetics, we will soon be able to profile variations between individuals' DNA to predict responses to a particular medicine. Target validation that will predict a well-tolerated and effective medicine for a clinical indication in humans is a widely perceived problem; but the real challenge is target selection. A limited number of molecular target families have been identified, including receptors and enzymes, for which high throughput screening is currently possible. A good target is one against which many compounds can be screened rapidly to identify active molecules (hits). These hits can be developed into optimized molecules (leads), which have the properties of well-tolerated and effective medicines.

30 Selection of targets that can be validated for a disease or clinical symptom is a major problem faced by the pharmaceutical industry. The best-validated targets are those that have already produced well-tolerated and effective medicines in humans (precedent targets). Many targets are chosen on the basis of scientific hypotheses and do not lead to effective medicines because the initial hypotheses are often subsequently disproved.

WO 03/025222

PCT/EP02/10667

-2-

Two broad strategies are being used to identify genes and express their protein products for use as high-throughput targets. These approaches of genomics and genetics share technologies but represent distinct scientific tactics and investments. Discovery  
5 genomics uses the increasing number of databases of DNA sequence information to identify genes and families of genes for tractable or scrollable targets that are not known to be genetically related to disease.

The advantage of information on disease-susceptibility genes derived from patients is  
10 that, by definition, these genes are relevant to the patients' genetic contributions to the disease. However, most susceptibility genes will not be tractable targets or amenable to high-throughput screening methods to identify active compounds.

The differential metabolism related to the relevant gene variants can be studied in  
15 focused functional genomic and proteomic technologies to discover mechanisms of disease development or progression.

Critical enzymes of receptors associated with the altered metabolism can be used as  
20 targets. Gene-to-function-to-target strategies that focus on the role of the specific susceptibility gene variants on appropriate cellular metabolism become important.

Data mining of sequences from the Human Genome Project and similar programmes  
25 with powerful bioinformatic tools has made it possible to identify gene families by locating domains that possess similar sequences. Genes identified by these genomic strategies generally require some sort of functional validation or relationship to a disease process. Technologies such as differential gene expression, transgenic animal models, proteomics, in situ hybridization and immunohistochemistry are used to imply  
relationships between a gene and a disease.

The major distinction between the genomic and genetic approaches is target selection,  
which genetically defined genes and variant-specific targets already known to be  
30 involved in the disease process. The current vogue of discovery genomics for nonspecific, wholesale gene identification, with each gene in search of a relationship to a disease, creates great opportunities for development of medicines.

It is also critical to realize that the core problem for drug development is poor target  
selection. The screening use of unproven technologies to imply disease-related

WO 03/025222

-3-

PCT/EP02/10667

validation, and the huge investment necessary to progress each selected gene to proof of a concept in humans, is based on an unproven and cavalier use of the word 'validation'. Each failure is very expensive in lost time and money. For example, differential gene expression (DGE) and proteomics are screening technologies that are widely used for target validation. They detect different levels and/or patterns of gene and protein expression in tissues, which may be used to imply a relationship to a disease affecting that tissue.

**Mood Disorder Background:**

Mood disorders or related disorders include but are not limited to the following disorders as defined in the Diagnostic and statistical Manual of Mental Disorders, version 4 (DSM-IV) taxonomy DSM-IV codes in parenthesis): mood disorders (296.XX,300.4,311,301.13,295.70) , schizophrenia and related disorders (295.XX,297.1,298.8,297.3,298.9), anxiety disorders (300.XX,309.81,308.3), adjustment disorders (309.XX) and personality disorders (codes 301.XX) .

The present invention is particularly directed to genetic factors associated with a family of mood disorders known as bipolar (BP) spectrum disorders. Bipolar disorder (BP) is a severe psychiatric condition that is characterized by disturbances in mood, ranging from an extreme state of elation (mania) to a severe state of dysphoria (depression).

Two types of bipolar illness have been described: type I BP illness (BPI) is characterized by major depressive episodes alternated with phases of mania, and type II BP illness (BPII) , characterized by major depressive episodes alternating with phases of hypomania. Relatives of BP probands have an increased risk for BP, unipolar disorder (patients only experiencing depressive episodes; UP), cyclothymia (minor depression and hypomania episodes; cy) as well as for schizoaffective disorders of the manic (SAm) and depressive (SAd) type. Based on these observations BP, cY, UP and SA are classified as BP spectrum disorders.

The involvement of genetic factors in the etiology of BP spectrum disorders was suggested by family, twin and adoption studies (Tsuang and Faraone (1990), the Genetics of Mood Disorders, Baltimore, The John Hopkins University Press) However, the exact pattern of transmission is unknown. In some studies, complex segregation analysis supports the existence of a single major locus for BP (Spence et al. (1995), Am J.Med. Genet (Neuropsych. Genet.) QQ pp 370-376). Other researchers propose a liability-threshold-model, in which the liability to develop the disorder results from the

WO 03/025222

PCT/EP02/10667

-4-

additive combination of multiple genetic and environmental effects (McGuffin et al. (1994), *Affective Disorders; Seminars in Psychiatric Genetics* Gaskell, London pp 110-127).

- Due to the complex mode of inheritance, parametric and non-parametric linkage strategies are applied in families in which BP disorder appears to be transmitted in a Mendelian fashion. Early linkage findings on chromosomes 11p15 (Egeland et al. (1987), *Nature* ~ pp 783-787) and Xq27-q28 (Mendlewicz et al. (1987), *the Lancet* 1 pp 1230-1232; Baron et al. (1987) *Nature* 12& pp 289-292) have been controversial and could initially not be replicated (Kelsoe et al. (1989) *Nature* ~ pp 238-243; Baron et al. (1993) *Nature Genet* ~ pp 49-55). With the development of a human genetic map saturated with highly polymorphic markers and the continuous development of data analysis techniques, numerous new linkage searches were performed. In several studies, evidence or suggestive evidence for linkage to particular regions on chromosomes 4, 12, 18, 21 and X was found (Black wood et al. (1996) *Nature Genetics* ~ pp 427-430, Craddock et al. (1994) *Brit J. psychiatry* ~ pp355-358, Berrettini et al. (1994), *Proc Natl Acad Sci USA* ~ pp 5918-5921, Straub et al. (1994) *Nature Genetics* ~ pp 291-296 and Pekkarinen et al. (1995) *Genome Research* 2 pp 105-115). In order to test the validity of the reported linkage results, these findings have to be replicated in other, independent studies.
- Recently, linkage of bipolar disorder to the pericentromeric region on chromosome 18 was reported (Berrettini et al. 1994). Also a ring chromosome 18 with break-points and deleted regions at 18pter-p11 and 18q23-qter was reported in three unrelated patients with BP illness or related syndromes (Craddock et al. 1994). The chromosome 18p linkage was replicated by Stine et al. (1995) *Am J. Hum Genet* 22 pp 1384-1394, who also reported suggestive evidence for a locus on 18q21.2-q21.32 in the same study.

- Interestingly, Stine et al. observed a parent-of-origin effect: the evidence of linkage was the strongest in the paternal pedigrees, in which the proband's father or one of the proband's father's sibs is affected. Several studies described anticipation in families transmitting BP disorder (McInnis et al 1993, Nylander et al 1994) suggesting the involvement of trinucleotide repeat expansions (TREs), considering a number of diseases caused by an expansion of a CAG/CTG, a CCG/CGG or a GAA/TTT repeat show anticipation (reviewed by Margolis et al., 1999). Previous efforts to find potentially expanded repeats have primarily focused on CAG/CTG repeats although the

WO 03/025222

-5-

PCT/EP02/10667

search for CCG/CGG repeats is increasing (Kleiderlein et al 1998, Mangel et al 1998, Eichhammer et al 1998, Kaushik et al 2000). Previously, we reported on a new method for the region specific isolation of triplet repeats: triplet repeat YAC fragmentation (Del Favero et al 1999). This proved to be a valid method for the isolation of CAG/CTG repeats and using this method, we excluded the involvement of CAG/CTG repeats from within 18q21.33-q23 in bipolar disorder (Goossens et al 2000). The present invention adapted the method for the region specific isolation of CCG/CGG repeats and applied it to the chromosome 18q21.33-q23 BP candidate region.

#### 10 **SUMMARY OF THE INVENTION:**

The present invention is directed to novel isolated nucleic acid sequence and the cytoplasmic antipeptidase 2 (CAP 2) protein encoded by isolated nucleic acid sequences.

The novel isolated nucleic acid sequence is located at an 8.9 cM chromosome region located between D18S68 and D18S979 at 18q21.33-q23. A physical map was constructed using yeast artificial chromosomes (YACs) (Verheyen et al 1999).

The previously described method was adapted for the region specific isolation of CCG/CGG repeats and applied to the chromosome 18q21.33-q23 BP candidate region. The YAC contig map confirmed the localization within the BP candidate region of a cluster of 6 genes coding for serine proteinase inhibitors (serpins). Serpins are a superfamily of proteolytic proteins with high overall homology to  $\alpha_1$ -proteinase inhibitor. All 6 serpins belong to the ovalbumin family of serpins that lack a typical amino-terminal cleavable signal peptide and can be intracellularly or both. CAP2 or P18 located at 18q21.33, contains a combined CAG-CGG triplet repeat sequence in its 5'UTR region and is expressed in brain. In this study, we determined the genomic organization and exon/intron boundaries of CAP2 and examined the gene by single-strand conformation polymorphism (SSCP) analysis and denaturing high-performance liquid chromatography (DHPLC) for sequence variants. Analysis of six single nucleotide polymorphisms (SNPs) by sequencing, RFLP-PCR or pyrosequencing was performed in a sample of 75 cases and 75 matched controls.

#### **BRIEF DESCRIPTION OF THE DRAWING**

WO 03/025222

-6-

PCT/EP02/10667

**Figure 1:** Minimal YAC tiling path of the 18q21.33-q23 BP candidate region (Verheyen et al 1999). The YACs are represented by solid lines, the CCG/CGG fragmentation products by dotted lines. YAC sizes, between brackets, are estimated by PFGE analysis. Solid circles indicate positive STS/STR hits. Shaded boxes highlight

5 the CCG/CGG repeat and the three CpG islands isolated by YAC fragmentation.

**Figure 2:** Genomic structure of Cytoplasmic antiproteinase 2 (CAP2) gene. Black boxes represent exons and their sizes in bp are indicated above the box. Introns sizes are in kb. The combined CAG-CGG repeat is indicated. Transcription initiation and stop codons are indicated.

10

#### DETAILED DESCRIPTION OF THE INVENTION:

The present invention is directed to a novel isolated nucleic acid sequence comprising gene located at the 18q chromosomal candidate region of chromosome 18.

15

The gene is located at a chromosomal region associated with mood disorders such as bipolar spectrum disorders and therefore is useful as a diagnostic marker for bipolar spectrum disorders. The region in question when removed from the totality of the human genome may also be used to locate, isolate and sequence other genes which influences psychiatric health and mood.

20

Specifically the BP candidate region contains the gene coding for cytoplasmic antiproteinase 2 (CAP2), a brain expressed serpin implicated in a number of intra-and extracellular functions. In this study we determined the genomic organization of CAP2 and defined all intron/exon boundaries. CAP2 comprises 7 exons within an estimated

25 17-kb genomic region.

Mutation analysis of CAP2 identified 3 non-synonymous single nucleotide polymorphism (SNPs): c.203G>A (Arg69Gln), c.910A>G (Thr304Ala) and c.1076G>A (Arg359His); 2 synonymous SNPs c. 477>G and c.942>T and 1 intronic SNP IVS4+98A>G. Analysis of CAP2 polymorphisms in unrelated BP cases and

30 matched controls showed a statistical significant difference with SNP c.942C>T in allele and genotype frequencies ( $p=0.03$ ).

**Isolation and identification of novel gene:**

Standard procedures well-known to one skilled in the art were applied to the identified YAC clones and, where applicable, to the DNA from an individual afflicted with a mood disorder as defined herein, in the process of identifying and characterizing the relevant gene. For example, the inventors are able to make use of the previously identified apparent association between trinucleotide repeat expansions (TRE) within the human genome and the phenomenon of anticipation in mood disorders (Lindblad et al. (1995), *Neurobiology of Disease* 2, pp 55-62 and O'Donovan et al. (1995), *Nature Genetics* 1Q pp 380-381) to screen for TRE's in the selected YAC clones in order to identify candidate genes in the region of interest on human chromosome 18. A variety of other known procedures can also be applied to the said YAC clones to identify the candidate gene as discussed below.

Accordingly, in a first aspect the present invention comprises the use of an 8.9 cM region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with mood disorders or related disorders as defined above. As will be described below, the present inventors have identified this candidate region of chromosome 18q for such a gene, by analysis of co-segregation of bipolar disease in family MAD31 with 12 STR polymorphic markers previously located between D18S51 and D18S61 and subsequent allele sharing analysis.

Particular YACs covering the candidate region which may be used in accordance with the present invention are 961.h-9, 942.c.3, 766.f.12, 731.c- 7, 907.e.1, 752.g-8 and 717-d-3, preferred ones being 961h-9, 766.f.12 and 907-e.1 since these have the minimum tiling path across the candidate region. Suitable YAC clones for use are those having an artificial chromosome spanning the refined candidate region between D18S68 and D18S979.

There are a number of methods which can be applied to the candidate regions of chromosome 18q as defined above, whether or not present in a YAC, to identify a candidate gene or genes associated with mood disorders or related disorders. For example, as aforesaid, there is an apparent association between the extent of trinucleotide repeat expansions (TRE) in the human genome and the presence of mood disorders.



WO 03/025222

PCT/EP02/10667

-8-

Accordingly, in a third aspect the present invention comprises a method of identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with a mood disorder or related disorder as defined herein which comprises detecting nucleotide triplet repeats in the region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979.

5 An alternative method of identifying said gene or genes comprises fragmenting a YAC clone comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61, for example one or more of the seven aforementioned YAC clones, and detecting any nucleotide triplet repeats in said fragments, in particular  
10 repeats of CAG or CTG. Nucleic acid probes comprising at least 5 and preferably at least 10 CTG and/or CAG triplet repeats are a suitable means of detection when appropriately labelled. Trinucleotide repeats may also be determined using the known RED (repeat expansion detection) system (Shalling et al. (1993), Nature Genetics ~ pp 135-139).

15 In a fourth embodiment the invention comprises a method of identifying at least one gene, including mutated and polymorphic variants thereof, which is associated with a mood disorder or related disorder and which is present in a YAC clone spanning the region of human chromosome 18q between polymorphic markers D18S60 and D18S61, the method comprising the step of detecting the expression  
20 product of a gene incorporating nucleotide triplet repeats by use of an antibody capable of recognizing a protein with an amino acid sequence comprising a string of at least 8, but preferably at least 12, continuous glutamine residues. Such a method may be implemented by sub-cloning YAC DNA, for example from the seven aforementioned YAC clones, into a human DNA expression library. A preferred means of detecting the  
25 relevant expression product is by use of a monoclonal antibody, in particular mAB1C2, the preparation and properties of which are described in International Patent Application Publication No WO 97/17445.

Further embodiments of the present invention relate to methods of identifying the relevant gene or genes which involve the (sub-)cloning of (YAC) DNA as defined  
30 above into vectors such as BAC (bacterial artificial chromosome) or PAC (P1 or phage artificial chromosome) or cosmid vectors such as exon-trap cosmid vectors. The starting point for such methods is the construction of a contig map of the region of human chromosome 18q between polymorphic markers D18S60 and D18S61. To this end the present inventors have sequenced the end regions of the fragment of human

WO 03/025222

-9-

PCT/EP02/10667

- DNA in each of the seven aforementioned YAC clones and these sequences are disclosed herein. Following sub-cloning of YAC DNA into other vectors as described above, probes comprising these end sequences or portions thereof, in particular those sequences shown in Figures 1 to 11 herein, together with any known sequenced tagged site (STS) in this region, as described in the YAC clone contig shown herein, as can be used to detect overlaps between said sub-clones and a contig map can be constructed. Also the known sequences in the current YAC contig can be used for the generation of contig map sub-clones.
- One route by which a gene or genes which is associated with a mood disorder or associated disorder can be identified is by use of the known technique of exon trapping. This is an artificial RNA splicing assay, most often making use in current protocols of a specialized exon-trap cosmid vector. The vector contains an artificial mini-gene consisting of a segment of the SV40 genome containing an origin of replication and a powerful promoter sequence, two splicing-competent exons separated by an intron which contains a multiple cloning site and an SV40 polyadenylation site.
- The YAC DNA is sub-cloned in the exon-trap vector and the recombinant DNA is transfected into a strain of mammalian cells. Transcription from the SV40 promoter results in an RNA transcript which normally splices to include the two exons of the minigene. If the cloned DNA itself contains a functional exon, it can be spliced to the exons present in the vector's minigene. Using reverse transcriptase a cDNA copy can be made and using specific PCR primers, splicing events involving exons of the insert DNA can be identified. Such a procedure can identify coding regions in the YAC DNA which can be compared to the equivalent regions of DNA from a person afflicted with a mood disorder or related disorder to identify the relevant gene.
- Accordingly, in a fifth aspect the invention comprises a method of identifying at least one human gene, including mutated variants and polymorphisms thereof, which is associated with a mood disorder or related disorder which comprises the steps of:
- (1) transfecting mammalian cells with exon trap cosmid vectors prepared and mapped as described above;
  - (2) culturing said mammalian cells in an appropriate medium;
  - (3) isolating RNA transcripts expressed from the SV40 promoter;
  - (4) preparing cDNA from said RNA transcripts;
  - (5) identifying splicing events involving exons of the DNA sub-cloned into said exon trap cosmid vectors to elucidate positions of coding regions in said sub-cloned DNA;

WO 03/025222

-10-

PCT/EP02/10667

(6) detecting differences between said coding regions and equivalent regions in the DNA of an individual afflicted with said mood disorder or related disorder; and

(7) identifying said gene or mutated or polymorphic variant thereof which is associated with said mood disorder or related disorders.

5 As an alternative to exon trapping the YAC DNA may be sub-cloned into BAC, PAC, cosmid or other vectors and a contig map constructed as described above. There are a variety of known methods available by which the position of relevant genes on the sub-cloned DNA can be established as follows:

- 10 (a) cDNA selection or capture (also called direct selection and cDNA selection) : this method involves the forming of genomic DNA/cDNA heteroduplexes by hybridizing a cloned DNA (e.g. an insert of a YAC DNA), to a complex mixture of cDNAs, such as the inserts of all cDNA clones from a specific (e.g. brain) cDNA library. Related sequences will hybridize and can be enriched in subsequent steps using biotin-streptavidine capturing and PCR (or related techniques);
- 15 (b) hybridization to mRNA/cDNA: a genomic clone (e.g. the insert of a specific cosmid) can be hybridized to a Northern blot of mRNA from a panel of culture cell lines or against appropriate (e.g. brain) cDNA libraries. A positive signal can indicate the presence of a gene within the cloned fragment;
- 20 (c) CpG island identification: CpG or HTF islands are short (about 1 kb) hypomethylated GC-rich (> 60%) sequences which are often found at the 5' ends of genes. CpG islands often have restriction sites for several rare-cutter restriction enzymes. Clustering of rare-cutter restriction sites is indicative of a CpG island and therefore of a possible gene. CpG islands can be detected by hybridization of a DNA clone to Southern blots of genomic DNA digested with rare-cutting enzymes, or by
- 25 island-rescue PCR (isolation of CpG islands from YACs by amplifying sequences between islands and neighbouring Alu-repeats) ;
- (d) zoo-blotting: hybridizing a DNA clone (e.g. the insert of a specific cosmid) at reduced stringency against a Southern blot of genomic DNA samples from a variety of animal species. Detection of hybridization signals can suggest conserved sequences,
- 30 indicating a possible gene. Accordingly, in a sixth aspect the invention comprises a method of identifying at least one human gene including mutated and polymorphic variants thereof which is associated with a mood disorder or related disorder which comprises the steps of:

WO 03/025222

-11-

PCT/EP02/10667

- (1) sub-cloning the YAC DNA as described above into a cosmid, BAC, PAC or other vector;
- (2) using the nucleotide sequences or any other sequenced tagged site (STS) in this region as in the YAC clone contig described herein, or part thereof consisting of not  
5 less than 14 contiguous bases or the complement thereof, to detect overlaps amongst the sub-clones and construct a map thereof;
- (3) identifying the position of genes within the sub-cloned DNA by one or more of CpG island identification, zoo-blotting, hybridization of the sub-cloned DNA to a cDNA library or a Northern blot of mRNA from a panel of culture cell lines;
- 10 (4) detecting differences between said genes and equivalent region of the DNA of an individual afflicted with a mood disorder or related disorder; and
- (5) identifying said gene which is associated with said mood disorders or related disorders.

If the cloned YAC DNA is sequenced, computer analysis can be used to establish the presence of relevant genes. Techniques such as homology searching and exon prediction may be applied.

Once a candidate gene has been isolated in accordance with the methods of the invention more detailed comparisons may be made between the gene from a normal individual and one afflicted with a mood disorder such as a bipolar spectrum disorder.  
20 For example, there are two methods, described as "mutation testing", by which a mutation or polymorphism in a DNA sequence can be identified. In the first the DNA sample may be tested for the presence or absence of one specific mutation but this requires knowledge of what the mutation might be. In the second a sample of DNA is screened for any deviation from a standard (normal) DNA. This latter method is more  
25 useful for identifying candidate genes where a mutation is not identified in advance. In addition the following techniques may be further applied to a gene identified by the above-described methods to identify differences between genes from normal or healthy individuals and those afflicted with a mood disorder or related disorder:

- (a) Southern blotting techniques: a clone is hybridized to nylon membranes containing  
30 genomic DNA digested with different restriction enzymes of patients and healthy individuals. Large differences between patients and healthy individuals can be visualized using a radioactive labelling protocol;
- (b) heteroduplex mobility in polyacrylamide gels: this technique is based on the fact that the mobility of heteroduplexes in non-denaturing polyacrylamide gels is less than

WO 03/025222

PCT/EP02/10667

-12-

- the mobility of homoduplexes. It is most effective for fragments under 200 bp;
- (c) single-strand conformational polymorphism analysis (SSCP or SSCA) : single stranded DNA folds up to form complex structures that are stabilized by weak intramolecular bonds.
- 5 The electrophoretic mobilities of these structures on non-denaturing polyacrylamide gels depends on their chain lengths and on their conformation;
- (d) chemical cleavage of mismatches (CCM) : a radiolabelled probe is hybridized to the test DNA, and mismatches detected by a series of chemical reactions that cleave one strand of the DNA at the site of the mismatch. This is a very sensitive method and can
- 10 be applied to kilobase-length samples;
- (e) enzymatic cleavage of mismatches: the assay is similar to CCM, but the cleavage is performed by certain bacteriophage or eukaryotic enzymes.
- (f) denaturing gradient gel electrophoresis: in this technique, DNA duplexes are forced to migrate through an electrophoretic gel in which there is a gradient of increasing
- 15 amounts of a denaturant (chemical or temperature). Migration continues until the DNA duplexes reach a position on the gel wherein the strands melt and separate, after which the denatured DNA does not migrate much further. A single base pair difference between a normal and a mutant DNA duplex is sufficient to cause them to migrate to different positions in the gel;
- 20 (g) direct DNA sequencing.

A more detailed discussion of these suitable assay techniques is provided below.

- GENOTYPING As used herein, the term "genotyping" means determining whether a CAP2 encoding polynucleotide includes a thymidine (T) at position 942. The term
- 25 "genotyping" is synonymous with terms such as "genetic testing", "genetic screening", "determining or identifying an allele or polymorphism", "molecular diagnostics" or any other similar phrase.

- Any method capable of distinguishing nucleotide differences in the appropriate sample DNA sequences may also be used. In fact, a number of known different methods are
- 30 suitable for use in genotyping (that is, determining the genotype) for a CAP2 encoding polynucleotide of the present invention. These methods include but are not limited to direct sequencing, PCR-RFLP, ARMS-PCR, Taqman™, Molecular beacons,

WO 03/025222

-13-

PCT/EP02/10667

hybridization to oligonucleotides on DNA chips and arrays, single nucleotide primer extension and oligo ligation assays.

5 GENOTYPE SCREENING In one embodiment, the present invention provides a method for genotype screening of a nucleic acid comprising a CAP2 encoding polynucleotide from an individual. The methods for genotype screening of a nucleic acid comprising a CAP2 encoding polynucleotide from an individual may require amplification of a nucleic acids from a target sample from that individual.

10 TARGET SAMPLE The target samples of the present invention may be any target nucleic acid comprising a CAP2 encoding polynucleotide from an individual being analyzed. For assay of such nucleic acids, virtually any biological sample (other than pure red blood cells) is suitable. For example, convenient target samples include but are not limited to whole blood, leukocytes, semen, saliva, tears, urine, faecal material, sweat, buccal, skin and hair. For assay of cDNA or mRNA, the target sample is typically obtained from a cell or organ in which the target nucleic acid is expressed.

15 GENOTYPING SNPS A number of different methods are suitable for use in determining the genotype for an SNP. These methods include but are not limited to direct sequencing, PCR- RFLP, ARMS-PCR, Taqman™, Molecular beacons, hybridization to oligonucleotides on DNA chips and arrays, single nucleotide primer extension and oligo ligation assays. Any method capable of distinguishing single nucleotide differences in the appropriate DNA sequences may also be used.

20 AMPLIFICATION As used herein, the term "amplification means nucleic acid replication involving template specificity. The template specificity relates to a "target sample" or "target sequence" specificity. The target sequences are "targets" in the sense that they are sought to be sorted out from other nucleic acids. Consequently, amplification techniques have been designed primarily for sorting this out. Examples of amplification methods include but are not limited to polymerase chain reaction (PCR), polymerase chain reaction of specific alleles (PASA), ligase chain reaction (LCR), transcription amplification, self-sustained sequence replication and nucleic acid based sequence amplification (NASBA).

30 TAQMAN Suitable means for determining genotype may be based on the Taqman™ technique. The Taqman™ technique is disclosed in the following US patents

WO 03/025222

-14-

PCT/EP02/10667

4,683,202; 4, 683,195 and 4,965,188. The use of uracil N-glycosylase which is included in Taqman™ allelic discrimination assays is disclosed in US patent 5,035,996.

5 PCR PCR techniques are well known in the art (see for example, EP- A-0200362 and EP-A- 0201184 and US patent Nos 4 683 195 and 4 683 202). The process for amplifying the target sequence consists of introducing an excess of two oligonucleotide primers to the DNA mixture containing the desired target sequence, followed by a precise sequence of thermal cycling in the presence of a DNA polymerase. With PCR, it is possible to amplify a single copy of a specific target sequence in, for example, 10 genomic DNA to a level detectable by several different methodologies (such as hybridisation with a labelled probe, incorporation of biotinylated primers followed by avidin-enzyme conjugate detection and incorporation of <sup>32</sup>p labelled deoxynucleotide triphosphates, such as dCTP or dATP, into the amplified sequence). Alternatively, it is possible to amplify different polymorphic sites (markers) with primers that are 15 differentially labelled and thus can each be detected. One means of analysing multiple markers involves labelling each marker with a different fluorescent probe. The PCR products are then analysed on a fluorescence based automated sequencer. In addition to genomic DNA, any oligonucleotide sequence may be amplified with the appropriate set of primer molecules. In particular, the amplified segments created by the PCR process 20 itself are, themselves, efficient templates for subsequent PCR amplifications. By way of example, PCR can also be used to identify primers for amplifying suitable sections of a CAP2 encoding polynucleotide in or from a human.

PRIMERS The present invention also provides a series of useful primers.

As used herein, the term "primer" refers to a single-stranded oligonucleotide capable of 25 acting as a point of initiation of template-directed DNA synthesis under appropriate conditions (i.e., in the presence of four different nucleoside triphosphates and an agent for polymerization, such as, DNA or RNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. The appropriate length of a primer depends on the intended use of the primer but typically ranges from 15 to 30 30 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the

WO 03/025222

-15-

PCT/EP02/10667

exact sequence of the template but must be sufficiently complementary to hybridize with a template.

The term "primer site" refers to the area of the target DNA to which a primer hybridizes.

- 5 The term "primer pair" means a set of primers including a 5' upstream primer that hybridizes with the 5' end of the DNA sequence to be amplified and a 3' downstream primer that hybridizes with the complement of the 3' end of the sequence to be amplified.

- 10 The primers of the present invention may be DNA or RNA, and single- or double-stranded. Alternatively, the primers may be naturally occurring or synthetic, but are typically prepared by synthetic means.

- PRIMER HYBRIDISATION CONDITIONS As used herein, the term "hybridisation" refers to the pairing of complementary nucleic acids. Hybridisation and the strength of hybridisation (i.e. the strength of association between the nucleic acids) is impacted by such factors as the degree of complementarity between nucleic acids, stringency of conditions involved, the melting temperature ( $T_m$ ) of the formed hybrid and the G:C ratio within the nucleic acids.

- As used herein, the term "stringency" is used in reference to the conditions of temperature, ionic strength and the presence of other compounds such as organic solvents under which the nucleic acid hybridizations are conducted.

Hybridizations are typically performed under stringent conditions, for example, at a salt concentration of no more than 1M and a temperature of at least 25°C. For example, conditions of 5X SSPE (750 mM NaCl, 50 mM NaPhosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30°C. are suitable for allele-specific primer hybridizations.

- 25 ALLELE SPECIFIC PRIMERS An allele-specific primer hybridises to a site on target DNA overlapping a polymorphism and only primes amplification of an allelic form to which the primer exhibits perfect complementarity (See Gibbs, Nucleic Acid Res. 17, 2427- 2448 (1989)). This primer may be used in conjunction with a second primer which hybridises at a distal site. Amplification proceeds from the two primers leading to a detectable product signifying the particular allelic form is present. A control may



WO 03/025222

-16-

PCT/EP02/10667

be performed with a second pair of primers, one of which shows a single base mismatch at the polymorphic site and the other of which exhibits perfect complementarity to a distal site. The single-base mismatch prevents amplification and no detectable product is formed. The method works best when the mismatch is  
5 included in the 3'-most position of the oligonucleotide aligned with the polymorphism because this position is most destabilizing to elongation from the primer (see, for example WO 93/22456).

Hybridisation probes capable of specific hybridisation to detect a single base mismatch may be designed according to methods known in the art and described in Maniatis et al  
10 Molecular Cloning: A Laboratory Manual, 2nd Ed (1989) Cold Spring Harbour.

(i) PCR PRIMERS Preferably the screening is carried out using PCR primers designed to amplify portions of the human a CAP2 encoding polynucleotide (gene) that include nucleotide 942.

Examples of such PCR primers are shown as SEQ ID's Nos. 9 and 10.

15 **DETECTION OF POLYMORPHISMS IN AMPLIFIED TARGET SEQUENCES** The amplified nucleic acid sequences may be detected using procedures including but not limited to allele-specific probes, tiling arrays, direct sequencing, denaturing gradient gel electrophoresis and single-strand conformation polymorphism (SCCP) analysis.

20 **ALLELE-SPECIFIC PROBES** Allele-specific probes can be designed that hybridize to a segment of target DNA from one individual but do not hybridize to the corresponding segment from another individual due to the presence of different polymorphic forms in the respective segments from the two individuals.

As used herein, the term "probe" refers to an oligonucleotide (i.e. a sequence of nucleotides), whether occurring naturally as in a purified restriction digest or produced  
25 synthetically, which is capable of hybridizing to another oligonucleotide sequence of interest. Probes are useful in the detection, identification and isolation of particular gene sequences. The hybridisation probes of the present invention are typically oligonucleotides capable of binding in a base-specific manner to a complementary strand of nucleic acid.

WO 03/025222

-17-

PCT/EP02/10667

The probes of the present invention may be labeled with any "reporter molecule" so that it is detectable in any detection system, including but not limited to enzyme (for example, ELISA, as well as enzyme based histochemical assays), fluorescent, radioactive and luminescent systems. The target sequence of interest (that is, the sequence to be detected) may also be labeled with a reporter molecule. The present invention is not limited to any particular detection system or label.

The hybridization conditions chosen for the probes of the present invention are sufficiently stringent that there is a significant difference in hybridization intensity between alleles, and preferably an essentially binary response, whereby a probe hybridizes to only one of the alleles. The typical hybridization conditions are stringent conditions as set out above for the allele specific primers of the present invention so that a one base pair mismatch may be determined.

**TILING ARRAYS** The polymorphisms of the present invention may also be identified by hybridisation to nucleic acid arrays, some example of which are described in WO 95/11995. The term "tiling" generally means the synthesis of a defined set of oligonucleotide probes that is made up of a sequence complementary to the sequence to be analysed (the "target sequence"), as well as preselected variations of that sequence. The variations usually include substitution at one or more base positions with one or more nucleotides.

**DIRECT SEQUENCING** The direct analysis of the sequence of polymorphisms of the present invention may be accomplished using either the dideoxy chain termination method or the Maxam Gilbert method (see Sambrook et al., Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989) or using, for example, Standard ABI sequencing technology using Big Dye Terminator cycle sequencing chemistry analyzed on an ABI Prism 377 DNA sequencer. Preferably, the polymorphism used in the assays of the present invention are identified by the presence or absence of the fragments generated by PstI restriction analysis of the identified sequences.

**1.5 DENATURING GRADIENT GEL ELECTROPHORESIS** Amplification products of the present invention, which are generated using PCR, may also be analyzed by the use of denaturing gradient gel electrophoresis. Different alleles may be identified based on the different sequence-dependent melting properties and electrophoretic migration

WO 03/025222

PCT/EP02/10667

-18-

of DNA in solution. Erlich, ed., PCR Technology, Principles and Applications for DNA Amplification, (W.H. Freeman and Co, New York, 1992), Chapter 7.

#### SINGLE-STRAND CONFORMATION POLYMORPHISM (SCCP) ANALYSIS

Alleles of target sequences of the present invention may also be differentiated using  
5 single-strand conformation polymorphism (SCCP) analysis, which identifies base differences by alteration in electrophoretic migration of single stranded PCR products, as described in Orita et al., Proc. Nat. Acad. Sci. 86, 2766-2770(1989). Amplified PCR products can be generated as described above, and heated or otherwise denatured, to form single stranded amplification products. Single-stranded nucleic acids may refold  
10 or form secondary structures which are partially dependent on the base sequence. The different electrophoretic mobilities of single-stranded amplification products may be related to base-sequence difference between alleles of target sequences.

#### IDENTIFYING DIFFERENCES BETWEEN TEST AND CONTROL SEQUENCES

These detection procedures for amplified nucleic acid sequences may be used to  
15 identify difference of one or more points of variation between a reference and test nucleic acid sequence or to compare different polymorphic forms of the CAP2 gene from two or more individuals.

#### REFERENCE NUCLEIC ACID SEQUENCES

As used herein the term "reference nucleic acid sequence" means a control nucleic acid sequence such as a control DNA  
20 sequence representing one or more individuals homozygous for each of the alleles being tested in that assay. By way of example, control DNA sequences may include but are not limited to: (i) a genomic DNA from homozygous individuals; (ii) a PCR product containing a relevant SNP amplified from homozygous individuals; or (iii) a DNA sequence containing a relevant SNP that has been cloned into a plasmid or other  
25 suitable vector. The control sample may also be an allelic ladder comprising a plurality of alleles from known set of alleles. There may be a plurality of control samples, each containing different alleles or sets of alleles. Other reference/control samples typically include diagrammatic representations, written representations, templates or any other means suitable for identifying the presence of a polymorphism  
30 in a PCR product or other fragment of nucleic acid. The terms "reference nucleic acid sequence", reference samples and control samples are used interchangeable throughout the text.

WO 03/025222

PCT/EP02/10667

-19-

H. THERAPEUTIC USES An aspect of the invention provides screening an individual for a predisposition to bipolar mood disorder and, if a polynucleotidetic predisposition is identified, treating that individual to delay or reduce or prevent the bipolar mood disorder.

- 5 In an embodiment of this aspect of the invention, the predisposition of an individual to bipolar mood disorder is assessed by determining whether that individual is homozygous for a CAP2 encoding polynucleotide in which nucleotide 942 is thymidine (T), is heterozygous for a CAP2 encoding polynucleotide in which guanosine (G) at  
10 position 942 is replaced by thymidine (T), or is homozygous for a CAP2 encoding polynucleotide in which nucleotide 942 is guanosine (G) using methods of detection discussed above.

15 Thus, an individual who is T/T homozygous at position 942, for the polymorphism is classified as being at highest risk. An individual being G/T heterozygous is classified as having moderate risk. An individual being G/G homozygous is classified as being in the lowest risk category.

20 Optionally, the assessment of an individual's risk factor is calculated by reference both to the presence of a CAP2 encoding polynucleotide polymorphism and also to other known polynucleotidetic or physiological or other indications. The invention in this way provides further information on which measurement of an individual's risk can be based.

General methodology reference Although in general the techniques mentioned herein are well known in the art, reference may be made in particular to Sambrook et al., Molecular Cloning, A Laboratory Manual (1989) and Ausubel et al., Short Protocols in Molecular Biology (1999) 4th Ed, John Wiley & Sons, Inc.

25 It will be appreciated that with respect to the methods described herein, in the step of detecting differences between coding regions from the YAC and the DNA of an individual afflicted with a mood disorder or related disorder, the said individual may be anybody with the disorder and not necessary a member of family MAD31.

30 In accordance with further aspects the present invention provides an isolated human gene and variants thereof associated with a mood disorder or related disorder and

WO 03/025222

-20-

PCT/EP02/10667

which is obtainable by any of the above described methods, an isolated human protein encoded by said gene and a cDNA encoding said protein.

Once a gene has been identified a number of methods are available to determine the function of the encoded protein. These methods are described by Eisenberg et al (Nature vol. 15, June 2000) and is herein incorporated by reference. One method involves a computational method that reveals functional linkages from genome sequences and is called the gene neighbor method. If in several genomes the genes that encode two proteins are neighbors on the chromosome, the proteins tend to be functionally linked. This method can be powerful in uncovering functional linkages in prokaryotes, where operons are common, but also shows promise for analysing interacting proteins in eukaryotes.

#### CAP-2 Gene

The complete intron-exon structure of the Cytoplasmic antiproteinase 2 gene (CAP2), which contains a 5'UTR and 6 coding exons spanning a genomic of 17.1 kb is herein disclosed. To size the introns, different combinations of primers spanning the CAP2 cDNA sequence are used. In this way, the size and exon-intron boundary sequences of 5 introns were derived. The 5' donor and 3' acceptor sites at the splice junctions correlated with consensus sequences (Table 1). The first 5'UTR exon is very small (73 bp) and contains a (CAG)<sub>2</sub>(CGG)<sub>6</sub>(CAG)<sub>6</sub> sequence which proved to be polymorphic but not expanded in the MAD 31 Belgian family nor in the affected and the control population.

The CAP2 derived amino acid sequence exhibits a high degree of identity to other human members of the ovalbumin family of cytoplasmic serpins including Placental thrombin inhibitor or PI6 (68% identity) and proteinase inhibitor 9 PI9 (63%). 5 exonic polymorphisms were identified from which 3 result in amino acid change. Alignment of the deduced primary structure of CAP2 with the amino acid sequences of PI6 and PI9, showed that, at amino acid position 68, CAP2 can either be Arg or Gln and in PI6 is Gln. Similarly, at amino acid position 359, CAP2 exhibits either Arg or His and PI6 exhibits His.<sup>15</sup> In addition, at amino acid position 304, CAP2 exhibits either Thr or Ala and PI9 exhibits Ala. By contrary, at amino acid position 314, CAP2 exhibits Ala while PI6 and PI9 exhibit Val.

WO 03/025222

PCT/EP02/10667

-21-

Two of these variants (c.910A>G and c.942C>T) detected by DHPLC were not previously identified by SSCP analysis.

An association analysis in a 75/75 case-control sample of Belgian origin was applied. Cases and controls were strictly matched for ethnicity, gender and age. Comparison of the allele and genotype frequencies of the 6 SNPs indicated no significant association between patients and controls for 5 of the 6 SNPs. The frequency of the SNP c.942C>T substitution in exon 7 was significantly different between BP patients and controls: BP patients had a higher frequency of the T allele when compared to controls ( $p=0.03$ ). Although the 6 SNPs are located within the same gene, 4 of them were not in linkage disequilibrium. There was a very strong LD between the SNP c.203G>A and the SNP IVS4+98A>G, but only in controls. In BP cases, the LD between these 2 SNPs was weak. In addition, no LD was found between the CAP2-CAG-CGG repeat and the CAP2 SNPs. The fact that no strong evidence for association was found with BP disorder and the CAP2 SNPs, together with the lack of significant LD results in our population, suggests that the CAP2 might not play a major role in the etiology of BP disorders.

In one affected individual of family MAD31, a proximal recombination occurred between D18S68 and D18S969.<sup>11</sup> CAP2 gene is located between these two markers. The CAP2-SNPs indicated that this proximal recombination occurred downstream of the gene.

### Example 1

#### **A. Family, patients and control subjects**

The pedigree and the clinical diagnoses in MAD 31, a Belgian family with a BPII proband, were described in detail elsewhere.<sup>10</sup> Briefly, the different clinical diagnoses in family MAD31 are as follows: 1 BPI, 2 BPII, 2 UP, 4 major depressive disorder (MDD), 1 schizoaffective maniac (SAM) and 1 schizoaffective depressive (SAd). The case-control sample consisted of 75 unrelated patients of Belgian origin ascertained at the Erasme Hospital in Brussels, and 75 age, gender and ethnicity matched control subjects recruited through announcements in the hospital. All control individuals were interviewed to exclude psychiatric conditions. Patients fulfilled the Research Diagnostic Criteria<sup>16</sup> for BP disorder.

**B. PCR amplification**

Genomic DNA and cDNA were amplified using six overlapping primer sets spanning the CAP2 cDNA sequence (GenBank acc. no L40377).<sup>15</sup> Approximately 50 ng of genomic DNA or 1 ng of cDNA and 10 pg of each primer were used in a standard PCR reaction. Amplification conditions were as follows: initial denaturation step at 94°C for 4 min, followed by 35 cycles at 94°C for 1 min, 55°C for 1 min, 72°C for 2 min, and a final extension time at 72°C for 10 min.

**10 C. Southern blot analysis**

Genomic DNA from 3 affected and 2 non-affected members of family MAD31 was digested with *Hind* III and *Ban* HI separately and run on a 1% agarose gel. Southern blotting was performed according to the standard protocol<sup>17</sup>.

15 50 ng of CAP2 cDNA was labeled with ( $\alpha$ -<sup>32</sup>P) dCTP by random-primed labeling (Gibco-BRL). Hybridization was carried out overnight in Church buffer at 65°C. Subsequently, membranes were washed one time in 1XSSC, 0.1% SDS, one time in 0.5XSSC, 0.1% SDS and two times in 0.1XSSC, 0.1% SDS at 65°C followed by exposure to Kodak X-ray film at -70°C for 72 h.

20

**D. SSCP and DHPLC analysis**

PCR amplified DNA was analyzed by SSCP using the DNA Analysis System with precast ready-to-use gels and Hydrolink 5% glycerol gels (Pharmacia Biotech).

25 Denaturing high-performance liquid chromatography (DHPLC) was performed on automated instrumentation purchased from Transgenomic (Santa Clara, CA, USA). Crude PCR products, were loaded on a DNASep column and eluted from the column using an acetonitrile gradient in a 0.1 M triethylamine acetate buffer (TEAA), pH 7, at a constant flow rate of 0.9 ml/min. The gradient was created by mixing eluents A and  
30 B. Eluent A was 0.1 M TEAA, 0.1 M Na<sub>4</sub>EDTA. Eluent B was 25% acetonitrile in 0.1 M TEAA. The gradient and temperature required for successful resolution of heteroduplex molecules were predicted by Wavemaker version 3.4.4.

**E. DNA Sequencing**

Sequencing was performed on plasmid DNA or gel purified PCR templates using a Perkin-Elmer ABI 377 automated sequencer and the Big Dye terminator cycle  
5 sequencing kit (Applied Biosystems, PE), according to the manufacturer's protocol. PCR fragments were first visualized on an agarose gel and then gel purified, using Ultrafree-DA filter devices (Millipore).

**F. Pyrosequencing**

10 Biotinylated PCR products were immobilized onto streptavidin-coated paramagnetic beads (DynaL AS, Oslo, Norway). ssDNA was obtained by incubating the immobilized PCR product in 50µl 0.5 M NaOH for 5 min. followed by 2 sequential washes in 100 µl 10 mM Tris-Acetate pH 7.6. Primer Exon7-1025 (5'-GTG CCT CTG TCC AAG  
15 GTT GC-3') was used as pyrosequencing primer for the detection of the SNP c.942 in exon7. Primer annealing was performed by incubation at 72°C for 2 min and then at room temperature for 5 min. Pyrosequencing was performed on the PSQ96 pyrosequencer (Pyrosequencing AB, Uppsala, Sweden).

**G. Statistical Analysis**

Total allele and genotype distributions of BP cases and controls were compared and Hardy-Weinberg equilibrium was tested using Genepop.<sup>18</sup> Allele and genotype specific comparisons were done using a chi-square analysis or where appropriate, a Fisher exact  
25 test. The Dismult program<sup>19</sup> was used for multipoint association analysis combining the data of all SNPs genotyped. Linkage disequilibrium (LD) was calculated using Linkdos of Genepop. To estimate haplotype frequencies in patients and controls, the Arlequin algorithm based on the maximum likelihood method was applied.<sup>20</sup>

**H. Genomic Structure of CAP2 gene**

We determined the genomic structure of the CAP2 gene. First, the CAP2 gene was analyzed for genomic rearrangements by hybridizing a PCR-derived CAP2 cDNA fragment against two different Southern blots containing *Hind* III- and *Bam* HI-



WO 03/025222

-24-

PCT/EP02/10667

digested genomic DNA from the 5 selected members of family MAD31 (3 affected and 2 unaffected). Based on the observed hybridized bands, a minimal genomic size of 25 kb was estimated. No difference between the hybridization patterns of affected and non-affected individuals were observed.

- 5 Using cDNA primers, the position and size of introns were obtained from PCR on genomic DNA. After sequencing, the exact exon-intron boundaries were determined by comparison of cDNA and genomic sequences (Table 1). Intronic primers were designed from these genomic sequences (Table 4).

10 **Table 4** Intronic CAP2 primers

PCR primers for Exon 3	
15 Forward	5' ACTTTCAAT TTCITTTGTCATC 3' (SEQ ID NO 3)
Reverse	5' TACAAAGCAGGAGATATTCACC 3' (SEQ ID NO 4)
PCR primers for Intron 4	
20 Forward	5' GAAGCATATAAATGACTGGGTG 3' (SEQ ID NO 5)
Reverse	5' GATAAGAAATGACAGAGTTGC 3' (SEQ ID NO 6)
PCR primers for Exon 5	
25 Forward	5' CCAAGAGAATATTCCTG 3' (SEQ ID NO 7)
Reverse	5' AGTCGATCCCCTGACAAAAGC 3' (SEQ ID NO 8)
PCR primers for Exon 7	
30 Forward	5' AGCTGGAGGAGACTTATGACTT 3' (SEQ ID NO 9)
Reverse	5' GCAAGATAGGTAGAAGGAAAGG 3' (SEQ ID NO 10)

- 35 This analysis showed that *CAP2* contains 1 non-coding and 6 coding exons with sizes ranging from 73 bp (exon 1) to 405 bp (exon 7) (Fig 1 and table 1). The sizes of introns 2 to 6 were determined by PCR and ranged from 1.3 Kb (intron 2) to 1.8 Kb (intron 3) (Fig1). While these experiments were in progress, the complete sequence of the BAC 793J2 containing the *CAP2* gene became available (Genbank Acc. n° AC009802).
- 40 Exon-intron boundaries sequences, intron and exon sizes were confirmed and the size of intron 1 was determined at 8.1 kb. In total, the *CAP2* gene spans a genomic region of 17.1 kb.

To establish the orientation of the *CAP2* gene, a *CAP2*-CAG fragmented YAC<sup>21</sup> was analyzed for the presence of STS markers centromeric and telomeric to the gene

WO 03/025222

-25-

PCT/EP02/10667

including L40377 (CAP2 exon 7). PCR analysis showed positive hits with markers centromeric to CAP2 and the absence of amplification with markers telomeric to the gene indicating that the transcription orientation of CAP2 is from centromere to telomere.

5

**Table 1.** Intron-Exon Boundaries in CAP2

Exon N°	Size (bp)	Splice acceptor	Splice donor
1	73		GCAGCAGGAG/gtggggcct
2	178	ttgatgcag/ACCTTCTCTG	GATGTCCCAG/gtatgtgtgc
3	138	ttgatgcag/ACCTTCTCTC	TTCCCTCCAG/taagtagtat
4	118	gtgtttgcag/GACTTTAAAGA	AAGACTGAAG/gtgagacagt
5	143	ttctttag/GTAAGATTTC	AACCAACGAG/gtaggaaag
6	153	tttcgtag/GAAAAAAGA	CCTCGCGTG/gtaagtcca
7	405	cttatcctag/GTGGAAAAG	TTCTCCGTAA

10

**20 Mutation detection and analysis**

Intronic primers were designed in order to PCR amplify all exons from DNA of 24 BP patients. PCR products were screened for mutations using SSCP analysis. Two non-synonymous SNPs were identified at c.203G>A (Arg68Gln) and c.1076G>A (Arg359His). One synonymous SNP coding for Leu was identified c.477A>G (codon 159). In addition, 1 SNP was detected in intron 4, IVS4+98A>G. These results were confirmed by DHPLC analysis and resulted in the identification of 2 additional SNPs in exon 7, at c.910A>G (Thr304Ala) and one in c.942C>T (Ala314Ala) coding for Ala (Table 2).

25

30

Table 2. CAP2 polymorphisms

	Location of Polymorphism	Nucleotide position cDNA	Restriction site change	Protein position codon	Amino acid change
5	Exon 3	c.203G>A		68	Arg to Gln
10	Intron 4	IVS4+98A>G	Gain of <i>Rsa</i> I		
	Exon 5	c.477A>G	Loss of <i>Mae</i> I	159	Leu to Leu
	Exon 7	c.910A>G	Gain of <i>Pvu</i> II	304	Thr to Ala
15	Exon 7	c.942C>T		314	Ala to Ala
	Exon 7	c.1076G>A	Loss of <i>Hha</i> I	359	Arg to His

20

PCR-RFLP analysis in 75 unrelated bipolar patients and 75 matched controls was performed for 3 of these variants: A *Rsa* I-RFLP assay was applied for the SNP IVS4+98A>G. A *Pvu* II-RFLP analysis was applied for SNP c.910A>G. A *Hha* I-RFLP analysis was applied for the SNP c.1076G>A. SNPs c.477A>G and c.203G>A were analyzed by direct sequencing of PCR fragments generated from genomic DNA. Pyrosequencing was used to analyze the SNP c.942C>T.

25

There was no significant difference between BP patients and controls in allele frequencies or genotype distribution in 5 of these SNPs. However, there is a slight departure of Hardy-Weinberg equilibrium for SNPs c.203G>A ( $p=0.05$ ) and c.477A>G ( $p=0.03$ ), both in BP, which comes from an excess of heterozygotes ( $p=0.03$  for ex3;  $p=0.02$  for ex5). In addition there is a slight excess of heterozygotes for SNP c.203G>A in the controls ( $p=0.04$ ).

30

The T allele of SNP c.942C>T had a significantly higher frequency in BP cases (6%) than in controls (1%) ( $\chi^2=4.83$ ;  $p=0.03$ ). When comparing genotypes 9/73 (12%) had

35

WO 03/025222

-27-

PCT/EP02/10667

one T allele compared with 2/75 (3%) controls ( $\chi^2=5.02$ ;  $p=0.03$ ). Interestingly, when data was analyzed after stratification for gender, a significant difference was observed. In males the T allele had a frequency of 8% in BP patients while it was not observed in controls (Fisher exact test,  $p=0.03$ ). In females no difference in allele or genotype distribution was observed between cases and controls.

To confirm these results the unrelated patients and matched control groups were extended and final genotyping association analysis performed in 113 BP patients and 163 age, sex and ethnicity matched controls. Table 3 shows the allele and genotype frequencies for these polymorphisms in patient and control populations.

**Table 3.** Genotype and allele frequencies for SNPs in the CAP2 gene

SNP	Genotypes							Alleles		
	Cases			Controls				BP	Controls	
	AA	AG	GG	AA	AG	GG	p value	A	A	p value
%	%	%	%	%	%		%	%		
c.203G>A	2	48	50	2	49	49	0.90	26	27	0.81
IVS4+98A>G	71	29	0	68	32	0	0.63	86	84	0.66
c.477A>G	5	47	48	8	43	49	0.76	28	30	0.77
c.910A>G	80	19	1	79	20	1	1.00	89	89	1
c.942C>T	0	10	90	0	3	97	0.02	5	1	0.02
c.1076G>A	21	43	36	17	50	33	0.93	43	42	0.93

### Reference List

The following publications are herein incorporated by reference.

- 5
1. Akiskal, H.S. *et al.* Psychopathology, temperament, and past course in primary major depressions. 1. Review of evidence for a bipolar spectrum. *Psychopathology* **22**, 268-277 (1989).
  - 10 2. Tsuang MT & Faraone SV. *The Genetics of Mood Disorders*. The Johns Hopkins University Press, Baltimore (1990).
  3. Spence, M.A. *et al.* Bipolar disorder: evidence for a major locus [see comments]. *Am. J. Med. Genet.* **60**, 370-376 (1995).
  4. Conneally PM. *Genetics and Psychiatric Disorders*. (1998).
  - 15 5. McInnis, M.G. *et al.* Anticipation in bipolar affective disorder. *Am. J. Hum. Genet.* **53**, 385-390 (1993).
  6. Nylander, P.O., Engstrom, C., Chotai, J., Wahlstrom, J. & Adolfsson, R. Anticipation in Swedish families with bipolar affective disorder. *J. Med. Genet.* **31**, 686-689 (1994).
  - 20 7. Li, R. & el-Mallakh, R.S. Triplet repeat gene sequences in neuropsychiatric diseases. *Harv. Rev. Psychiatry* **5**, 66-74 (1997).
  8. Van Broeckhoven, C. & Verheyen, G. Chromosome 18 workshop. *Psychiatr. Genet.* **8**, 97-108 (1998).
  - 25 9. Van Broeckhoven, C. & Verheyen, G. Report of the chromosome 18 workshop. *Am. J. Med. Genet.* **88**, 263-270 (1999).
  10. De bruyn, A., Souery, D., Mendelbaum, K., Mendlewicz, J. & Van Broeckhoven, C. Linkage analysis of families with bipolar illness and chromosome 18 markers. *Biol. Psychiatry* **39**, 679-688 (1996).
  - 30 11. Verheyen GR *et al.* Genetic refinement and physical mapping of a chromosome 18q candidate region for bipolar disorder. *Eur. J. Hum. Genet.* **7**, 427-434 (1999).
  12. Huber, R. & Carrell, R.W. Implications of the three-dimensional structure of alpha 1-antitrypsin for structure and function of serpins. *Biochemistry* **28**, 8951-8966 (1989).
  - 35 13. Remold-O'Donnell, E. The ovalbumin family of serpin proteins. *FEBS Lett.* **315**, 105-108 (1993).
  14. Bartuski, A.J., Kamachi, Y., Schick, C., Overhauser, J. & Silverman, G.A. Cytoplasmic antiproteinase 2 (PI8) and bomapin (PI10) map to the serpin cluster at 18q21.3. *Genomics* **43**, 321-328 (1997).

WO 03/025222

-29-

PCT/EP02/10667

15. Sprecher, C.A. *et al.* Molecular cloning, expression, and partial characterization of two novel members of the ovalbumin family of serine proteinase inhibitors. *J. Biol. Chem.* **270**, 29854-29861 (1995).
- 5 16. Endicott J & Spitzer RL. Schedule for Schizophrenia and Affective Disorders- Lifetime Version. New York: N.Y. state Psychiatric Institute., N.Y. (1978).
17. Sambrook J, Fritsch E & Maniatis T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, N.Y. (1989).
18. Raymond, M. & Rousset, F. GENEPOP (version 1.2): a population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249  
10 (1995).
19. Terwilliger, J.D. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56**, 777-787 (1995).
- 15 20. Schneider S, Roessli D & Excoffier L. Arlequin ver. 2.000: A software for population genetics data analysis. (2000).
21. Goossens, D. *et al.* No evidence for the involvement of CAG/CTG repeats from within 18q21.33- q23 in bipolar disorder. *Eur. J. Hum. Genet.* **8**, 385-388 (2000).

WO 03/025222

PCT/EP02/10667

-37-

**CLAIMS**

What is claimed is:

- 5 1. A method of diagnosing BP or susceptibility to BP in an individual which method comprises determining, in a sample from the individual, the single nucleotide polymorphism in the CAP2 gene of the individual, and determining the status of the individual by reference to polymorphism in the CAP2 gene.
- 10 2. A method according to claim 1 wherein the single nucleotide polymorphism of the individual is in linkage disequilibrium with the polymorphism in the CAP2 gene.
3. A method according to claims 1 or 2 wherein the single nucleotide polymorphism equals SNP c.942G>T.
- 15 4. A method according to any one of claims 1 to 3 wherein the single nucleotide polymorphism in the CAP2 gene of the individual is determined by a method selected from amplification refractory mutation system and restriction fragment length polymorphism such as Southern blotting techniques, single-strand conformational polymorphism analysis, chemical cleavage of mismatches and denaturing high-performance liquid chromatography.
- 20 5. A method according to any one of claims 1 to 3 wherein the single nucleotide polymorphism in the CAP2 gene of the individual is determined using a pair of PCR primers that amplify a fragment of the CAP2 gene containing the single nucleotide polymorphism.
- 25 6. A method according to claim 5 wherein the single nucleotide polymorphism consists of SNP c.942G>T.
- 30 7. A method according to claim 5 wherein the pair of PCR primers that amplify a fragment of the CAP2 gene consist of a forward and reverse primer comprising the sequences of SEQ ID No 9 and SEQ ID No 10.

WO 03/025222

PCT/EP02/10667

-38-

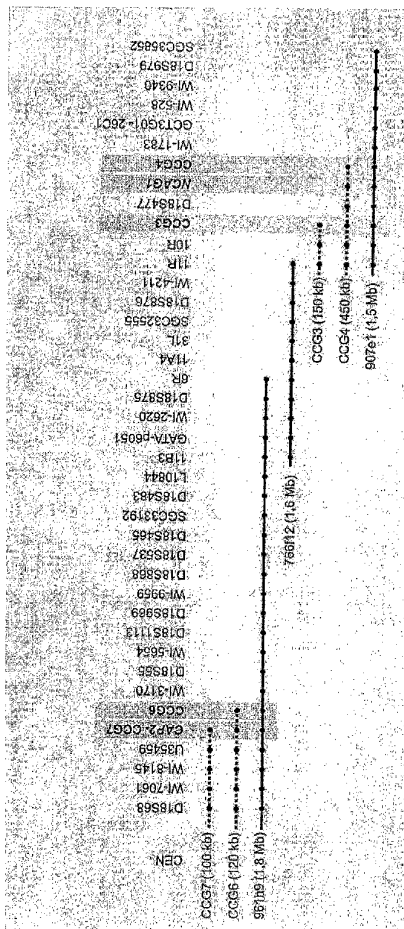
8. A pair of PCR primers capable of amplifying a fragment of the CAP2 gene containing the single nucleotide polymorphism.
- 5 9. A pair of PCR primers consisting of a forward and reverse primer comprising the sequences of SEQ ID No 9 and SEQ ID No 10.
10. A diagnostic kit comprising the pair of PCR primers according to claims 8 or 9.
- 10 11. A method according to any one of claims 1 to 3 wherein the single nucleotide polymorphism in the CAP2 gene of the individual is determined by means of an allele-specific oligonucleotide probe.
- 15 12. An allele-specific oligonucleotide probe capable of detecting the single nucleotide polymorphism SNP c.942G>T in the CAP2 gene of an individual.
13. A diagnostic kit comprising the allele-specific oligonucleotide probe according to claim 12.



WO 03/025222

PCT/EP02/10667

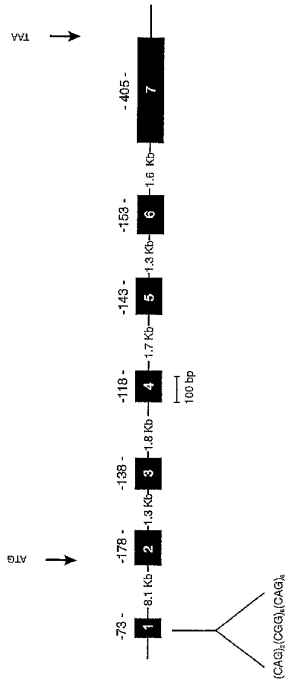
Fig. 12



WO 03/025222

PCT/EP02/10667

Fig. 22



WO 03/025222

-30-

PCT/EP02/10667

## SEQUENCE LISTING

<110> Janssen Pharmaceutica N.V.

5 <120> Novel Brain Expressed CAP-2 Gene and Protein associated with Bipolar Disorder

10 <130> JAB 1746

<150> EP 01203558.0  
 15 <151> 2001-09-17

<160> 10

20 <170> Patent in version 3.1

<210> 1

25 <211> 1325

<212> DNA

<213> Homo sapiens

30 <220>

<221> CDS  
 35 <222> (84)..(1205)

<223>

40 <400> 1  
 agcatctaca aaggaggaat agtcaaagca gcagcggcgg cggcggcggc ggcagcagca 60  
 gcagcagcag gagacottct ctg atg gat gac ctg tgt gaa gca aat ggc act 113  
 Met Asp Asp Leu Cys Glu Ala Asn Gly Thr  
 1 5 10

45 ttt gcc atc agc tta ttt aaa ata ttg ggg gaa gag gac aac tca aga 161  
 Phe Ala Ile Ser Leu Phe Lys Ile Leu Gly Glu Glu Asp Asn Ser Arg  
 15 20 25

50 aac gta ttc ttc tct ccc atg agc atc tcc tct gcc ctg gcc atg gtc 209  
 Asn Val Phe Phe Ser Pro Met Ser Ile Ser Ser Ala Leu Ala Met Val  
 30 35 40

55 ttc atg ggg gca aag gga agc act gca gcc cag atg tcc cag gca ctt 257  
 Phe Met Gly Ala Lys Gly Ser Thr Ala Ala Gln Met Ser Gln Ala Leu  
 45 50 55

60 tgt tta tac aaa gac gga gat att cac cga ggt ttc cag tca ctt ctg 305  
 Cys Leu Tyr Lys Asp Gly Asp Ile His Arg Gly Phe Gln Ser Leu Leu  
 60 65 70

65 agt gaa gtt aac aga act ggc act cag tac ttg ctt aga act gcc aac 353  
 Ser Glu Val Asn Arg Thr Gly Thr Gln Tyr Leu Leu Arg Thr Ala Asn  
 75 80 85 90



WO 03/025222

PCT/EP02/10667

-32-

```

ttc tgt gca gac cac cct ttt ctt ttc ttc atc agg cgc cac aaa acc 1169
Phe Cys Ala Asp His Pro Phe Leu Phe Phe Ile Arg Arg His Lys Thr
350 355 360
5 aac tgc atc ttg ttc tgt ggc agg ttc tct tct cgg taaagaggag 1215
Asn Cys Ile Leu Phe Cys Gly Arg Phe Ser Ser Pro
365 370
10 caattgctgt acataccctc ctttctctct acctatcttg cottaattaa cattccctgt 1275
gacctagtgt gtgcagtggc ttgaatgcca aaataaagcg tgtgcactgg 1325

<210> 2
15 <211> 374
<212> PRT
20 <213> Homo sapiens

<400> 2
25 Met Asp Asp Leu Cys Glu Ala Asn Gly Thr Phe Ala Ile Ser Leu Phe
1 5 10
30 Lys Ile Leu Gly Glu Glu Asp Asn Ser Arg Asn Val Phe Phe Ser Pro
20 25 30
35 Met Ser Ile Ser Ser Ala Leu Ala Met Val Phe Met Gly Ala Lys Gly
35 40 45
Ser Thr Ala Ala Gln Met Ser Gln Ala Leu Cys Leu Tyr Lys Asp Gly
50 55 60
40 Asp Ile His Arg Gly Phe Gln Ser Leu Leu Ser Glu Val Asn Arg Thr
65 70 75 80
45 Gly Thr Gln Tyr Leu Leu Arg Thr Ala Asn Arg Leu Phe Gly Glu Lys
85 90 95
50 Thr Cys Asp Phe Leu Pro Asp Phe Lys Glu Tyr Cys Gln Lys Phe Tyr
100 105 110
55 Gln Ala Glu Leu Glu Glu Leu Ser Phe Ala Glu Asp Thr Glu Glu Cys
115 120 125
Arg Lys His Ile Asn Asp Trp Val Ala Glu Lys Thr Glu Gly Lys Ile
130 135 140
60 Ser Glu Val Leu Asp Ala Gly Thr Val Asp Pro Leu Thr Lys Leu Val
145 150 155 160
65

```

WO 03/025222

-33-

PCT/EP02/10667

Leu Val Asn Ala Ile Tyr Phe Lys Gly Lys Trp Asn Glu Gln Phe Asp  
 165 170 175

5 Arg Lys Tyr Thr Arg Gly Met Leu Phe Lys Thr Asn Glu Gln Lys Lys  
 180 185 190

10 Thr Val Gln Met Met Phe Lys Glu Ala Lys Phe Lys Met Gly Tyr Ala  
 195 200 205

15 Asp Glu Val His Thr Gln Val Leu Glu Leu Pro Tyr Val Glu Glu Glu  
 210 215 220

Leu Ser Met Val Ile Leu Leu Pro Asp Asp Asn Thr Asp Leu Ala Val  
 225 230 235 240

20 Val Glu Lys Ala Leu Thr Tyr Glu Lys Phe Lys Ala Trp Thr Asn Ser  
 245 250 255

25 Glu Lys Leu Thr Lys Ser Lys Val Gln Val Phe Leu Pro Arg Leu Lys  
 260 265 270

30 Leu Glu Glu Ser Tyr Asp Leu Glu Pro Phe Leu Arg Arg Leu Gly Met  
 275 280 285

35 Ile Asp Ala Phe Asp Glu Ala Lys Ala Asp Phe Ser Gly Met Ser Thr  
 290 295 300

Glu Lys Asn Val Pro Leu Ser Lys Val Ala His Lys Cys Phe Val Glu  
 305 310 315 320

40 Val Asn Glu Glu Gly Thr Glu Ala Ala Ala Ala Thr Ala Val Val Arg  
 325 330 335

45 Asn Ser Arg Cys Ser Arg Met Glu Pro Arg Phe Cys Ala Asp His Pro  
 340 345 350

50 Phe Leu Phe Phe Ile Arg Arg His Lys Thr Asn Cys Ile Leu Phe Cys  
 355 360 365

Gly Arg Phe Ser Ser Pro  
 370

55 <210> 3

60 <211> 21

<212> DNA

<213> Artificial Sequence

WO 03/025222

PCT/EP02/10667

-34-

<220>  
5 <223> CAP2 - Exon3 Forward PCR primer  
<400> 3  
actttcaatt tctttgcat c 21  
10  
<210> 4  
<211> 22  
15 <212> DNA  
<213> Artificial Sequence  
20  
<220>  
<223> CAP2 - Exon3 Reverse PCR Primer  
25 <400> 4  
tacaagcag gagatattca cc 22  
30  
<210> 5  
<211> 22  
<212> DNA  
35 <213> Artificial Sequence  
40  
<220>  
<223> CAP2 - Intron 4 Forward PCR Primer  
<400> 5  
45 gaagcatata aatgactggg tg 22  
<210> 6  
<211> 21  
50 <212> DNA  
<213> Artificial Sequence  
55  
<220>  
<223> CAP2 - Intron 4 Reverse PCR Primer  
60 <400> 6  
gataagaat gacagagttg c 21

WO 03/025222 PCT/EP02/10667  
-35-

<210> 7  
<211> 18  
5 <212> DNA  
<213> Artificial Sequence

10  
<220>  
<223> CAP2 - Exon 5 Forward PCR Primer  
15 <400> 7  
ccaagagaat atttctctg 18

20  
<210> 8  
<211> 20  
<212> DNA  
25 <213> Artificial Sequence

30  
<220>  
<223> CAP2 - Exon 5 Reverse PCR Primer  
<400> 8  
35 agtcgatccc ctgacaaagc 20

40  
<210> 9  
<211> 22  
<212> DNA  
<213> Artificial Sequence

45  
<220>  
<223> CAP2 - Exon 7 Forward PCR Primer  
50 <400> 9  
agctggagga gagttatgac tt 22

55 <210> 10  
<211> 22  
<212> DNA  
60 <213> Artificial Sequence



WO 03/025222

PCT/EP02/10667

-36-

<220>

<223> CAP2 - Exon 7 Reverse PCR Primer

5 <400> 10  
gcaagatagg tagaaggaaa gg

22

## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/EP 02/10667
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7 C12Q1/68 C12N15/15		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
BIOSIS, MEDLINE, EMBASE, EPO-Internal, WPI Data		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	VERHEYEN G R ET AL: "Genetic refinement and physical mapping of a chromosome 18q candidate region for bipolar disorder" EUROPEAN JOURNAL OF HUMAN GENETICS, KARGER, BASEL, CH, vol. 7, no. 4, May 1999 (1999-05), pages 427-434, XP002112413 ISSN: 1018-4813 the whole document especially p.433, col. 2, lines 25-30. --- -/--	1-13
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
11 December 2002		07/01/2003
Name and mailing address of the ISA European Patent Office, P.O. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel (+31-70) 340-2040, Tx. 01 051 epo nl, Fax: (+31-70) 340-3016		Authorized officer Dumont, E

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		International Application No. PCT/EP 02/10667
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>STEVENSON G ET AL: "Discovery and genotyping of single nucleotide polymorphisms in candidate genes associated with bipolar disorder." AMERICAN JOURNAL OF HUMAN GENETICS, vol. 67, no. 4 Supplement 2, October 2000 (2000-10), page 385 XPO01121623</p> <p>50th Annual Meeting of the American Society of Human Genetics; Philadelphia, Pennsylvania, USA; October 03-07, 2000 ISSN: 0002-9297 cited in the application the whole document</p>	1-13
X	<p>WO 96 24650 A (ZYMOGENETICS INC) 15 August 1996 (1996-08-15) page 2, line 30 -page 4, line 7 page 10, line 27 -page 11, line 8</p>	8-10, 12, 13
A	<p>MCCARTHY J J ET AL: "THE USE OF SINGLE-NUCLEOTIDE POLYMORPHISM MAPS IN PHARMACOGENOMICS" NATURE BIOTECHNOLOGY, NATURE PUB. CO, NEW YORK, NY, US, vol. 18, May 2000 (2000-05), pages 505-508, XPO00941628 ISSN: 1087-0156 the whole document</p>	1-13
A	<p>JORDE L B: "Linkage disequilibrium and the search for complex disease genes." GENOME RESEARCH, vol. 10, no. 10, October 2000 (2000-10), pages 1435-1444, XPO02224534 ISSN: 1088-9051 the whole document</p>	1-13

International Application No. PCT/JP 02 10667

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 1-7 and claim 11 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compounds.

-----  
Continuation of Box I.1

Rule 39.1(iv) PCT - Diagnostic method practised on the human or animal body

INTERNATIONAL SEARCH REPORT		International application No. PCT/EP 02/10667
<b>Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)</b>		
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1.	<input checked="" type="checkbox"/>	Claims Nos.: — because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210
2.	<input type="checkbox"/>	Claims Nos.: — because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3.	<input type="checkbox"/>	Claims Nos.: — because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
<b>Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)</b>		
This International Searching Authority found multiple inventions in this International application, as follows:		
1.	<input type="checkbox"/>	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.	<input type="checkbox"/>	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.	<input type="checkbox"/>	As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.	<input type="checkbox"/>	No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
<b>Remark on Protest</b>		<input type="checkbox"/> The additional search fees were accompanied by the applicant's protest. <input type="checkbox"/> No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT  
 Information on patent family members

International Application No.  
 PCT/EP 02/10667

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9624650	A	15-08-1996	US 5712117 A 27-01-1998
			AU 4911096 A 27-08-1996
			CA 2212386 A1 15-08-1996
			EP 0808351 A2 26-11-1997
			JP 10513062 T 15-12-1998
			WO 9624650 A2 15-08-1996
			US 5710026 A 20-01-1998
			US 5578705 A 26-11-1996
			US 5747645 A 05-05-1998

## フロントページの続き

(81) 指定国 AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, N O, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(72) 発明者 デル・ファベロ, ユルゲン・ペーター・ロデ  
ベルギー・ビー - 9 0 5 2 ツイーナルデ・リービスシエストラート 1 2 0 ・フランダースインタ  
ーユニバーシテイインスチチュートフオーバイオテクノロジー (ブイアイビー)

(72) 発明者 バン・ブロークホーベン, クリステイン  
ベルギー・ビー - 9 0 5 2 ツイーナルデ・リービスシエストラート 1 2 0 ・フランダースインタ  
ーユニバーシテイインスチチュートフオーバイオテクノロジー (ブイアイビー)

F ターム(参考) 4B024 AA01 AA11 BA14 CA04 CA05 CA09 CA11 HA14  
4B063 QA13 QA19 QQ44 QR32 QR55 QR62 QS25 QS34