20040083９A

厚生労働科学研究費補助金

難治性疾患克服研究事業

罹患組織における遺伝子発現プロファイル解析からの病因解明に関する研究

平成 16 年度　総括・分担研究報告書

主任研究者　油　谷　浩　幸

平成 17（2005）年 3 月

目 次

Ⅰ 構成員名簿 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 1

Ⅱ 平成16年度総括研究報告書・・・・・・・・・・・・・・・・・・・・・・・・・ 5

Ⅲ 平成16年度分担研究報告書

発現プロファイル解析からの疾病遺伝子同定に関する研究・・・・・・・ 15
東京大学国際・産学共同研究センター教授 油谷 浩幸

知的統合による遺伝子、蛋白質相互作用探索の研究・・・・・・・・・・・ 23
東京大学先端科学技術研究センター教授 井原 茂男

特発性心筋症の解析に関する研究・・・・・・・・・・・・・・・・・・・・・ 29
千葉大学大学院医学研究院循環病態医科学教授 小室 一成

福山型筋ジストロフィーと長期記憶に関連したラット海馬遺伝子の解析・・ 31
大阪大学大学院医学系研究科教授 戸田 達史

難治性水疱性疾患およびその他の難治性皮膚疾患における罹患組織の
遺伝子発現プロファイルに関する研究・・・・・・・・・・・・・・・・・・ 34
九州大学大学院医学研究院皮膚科学教授 古江 増隆

血管内皮細胞における血管疾患（血栓、動脈硬化症）誘発刺激応答に
関する研究・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 37
東京大学先端科学技術研究センター助教授 南 敬

Ⅳ 研究成果の刊行に関する一覧表・・・・・・・・・・・・・・・・・・・ 43

構成員名簿

区　分	氏　名	所　　属	職　名
主任研究者	油谷　　浩幸	東京大学国際・産学共同研究センター	教　授
分担研究者	井原　　茂男	東京大学先端科学技術研究センター	教　授
分担研究者	小室　　一成	千葉大学大学院医学系研究院 循環病態医科学	教　授
分担研究者	戸田　　達史	大阪大学大学院医学系研究科	教　授
分担研究者	古江　　増隆	九州大学大学院医学研究院皮膚科学	教　授
分担研究者	南　　　敬	東京大学先端科学技術研究センター	助教授

（事務局）経理事務連絡担当責任者　安田　道義
東京大学生産技術研究所　経理課長
〒153-8904　東京都目黒区駒場 4-6-1
電話：03-5452-6041（直通）　ＦＡＸ：03-5452-6075

E-mail: haburata-tky@umin.ac.jp　（油谷）/ myasuda@iis.u-tokyo.ac.jp　（安田）

Ⅱ　平成 16 年度総括研究報告書

厚生労働科学研究費補助金（難治性疾患克服研究事業）
総括研究報告書

罹患組織における遺伝子発現プロファイル解析からの病因解明に関する研究

主任研究者　油谷浩幸　東京大学国際・産学共同研究センター　教授

研究要旨
ヒト疾患は個体あるいは臓器というシステムの破綻であり、それは生体固有のフィードバック機構により必ずしやトランスクリプトーム（遺伝子転写の総体）に反映されていると推定される。本研究班では、希少性疾患について罹患組織のトランスクリプトーム解析により、病態解明あるいは治療法開発のための標的遺伝子候補を探索すべく、本年度は1）発現プロファイルデータベースの拡充、2）ヒトゲノム中に存在する染色体コピー数変異の解析法の開発、3）慢性疲労症候群をはじめとする希少性疾患解析への応用に関する研究を行った。複数のアレイプラットフォームを用いてデータベースの拡充を進めた。ゲノムワイドな染色体コピー数の多型性の解析が可能となり、X 染色体数の異常や免疫グロブリン遺伝子座の再構成を検出することが出来たことから、今後疾患との関連の解明が待たれる。原因不明の病態である慢性疲労症候群については末梢血球由来 RNA の発現プロファイルから診断、病型分類を試みたところ、疾患群と罹患群は明瞭に識別可能であった。共通なプロファイルを示す遺伝子を制御する転写因子の同定を試みたところ、E2F が有意な因子として抽出された。発現変動の生じているパスウェイを同定するために文献からの相互作用検索は有効であった。罹患組織あるいは末梢血由来の RNA のプロファイル解析から難治性疾患の診断や病因に関連する候補遺伝子が同定され、免疫組織染色あるいはモデル動物を用いての検証が進められた。

分担研究者
油谷　浩幸　東京大学国際・産学共同研究センター・教授
井原　茂男　東京大学先端科学技術研究センター・教授
小室　一成　千葉大学大学院医学研究院循環病態医科学・教授
戸田　達史　大阪大学大学院医学系研究科・教授
古江　増隆　九州大学大学院医学研究院皮膚科学・教授
南　敬　東京大学先端科学技術研究センター・助教授

A.研究目的

希少性疾患は通常の遺伝解析のみによる病因解明が困難であるがゆえ、未だ原因不明であり、適切な治療法の開発も進められておらず、原因遺伝子が解明されても適当な治療標的分子の同定が必要となる。ヒトゲノム計画の進展は遺伝子塩基配列や発現情報などの網羅的データを基礎として新たな作業仮説を立て検証するという研究手法を可能とした。ヒト疾患は個体あるいは臓器というシステムの破綻であり、それは生体固有のフィードバック機構により必ずしやトランスクリプトーム（遺伝子転写の総体）に反映されていると推定される。罹患組織における多数の遺伝子の発現変動を同

時に正確に捉えることによる病態解明、いわば「臨床ゲノム学（clinical genomics）」の樹立が肝要である。本研究班では、希少性疾患について罹患組織のトランスクリプトーム解析により、病態解明あるいは治療法開発のための標的遺伝子候補を探索すべく下記の項目について研究を行った。

1. 発現プロファイルデータベースの拡充、データ交換

マイクロアレイによるヒト全遺伝子の発現プロファイルを測定し、データベース化した。さらに異なるアレイプラットフォームによるデータベースの拡充と、知的統合による遺伝子、蛋白質相互作用探索の研究を進めた。

2. 微量組織からのゲノム解析技術の確立

微量組織からの高感度トランスクリプトーム解析技術の確立を進めた。コピー数多型 CNP (copy number polymorphism)をゲノムタイピングマイクロアレイを用いて解析する方法の開発を目指した。

3. 希少性疾患解析への応用

慢性疲労症候群（chronic fatigue syndrome, CFS）は6ヶ月以上継続的・断続的に全身倦怠感を訴える原因不明の病態であり、患者末梢血球を採取後 RNA を抽出し、マイクロアレイ解析を行った。

福山型先天性筋ジストロフィ（FCMD)、II型滑脳症を伴う常染色体劣性遺伝性神経筋疾患である。以前 FCMD の原因遺伝子として同定した *fukutin* は筋細胞膜上の α-dystroglycan（αDG）の糖鎖修飾に関する糖転移酵素と推定されている。その一連の糖転移酵素に変異が同定されている muscle-eye-brain 病、Walker-Warburg 症候群、LARGE においても、同様の症状をきたし、α-dystroglycano pathy (α-DGpathy)と総称される。FCMD 患者の筋線維は他の筋ジストロフィーに比べ未熟である一方で、筋線維の再生は活発でないことが知られているが、その詳細に関し遺伝子レベルでの変化はわかっていない。また、長期記憶の成立に新規遺伝子の発現が関与することは示唆されているが、詳細な遺伝子発現の変化にはまだ未知の部分が大きいため、ラットをモデルとした解析に着手した。

天疱瘡や類天疱瘡などの難治性水疱性疾患は表皮細胞の接着装置に対する自己免疫性疾患である。ステロイド内服や血漿交換療法、免疫抑制剤内服などの治療が用いられ、予後はある程度改善されてきているものの、種々の治療に抵抗し死に至るケースも多い。天疱瘡ではデスモゾームを構成するデスモグレインに対する自己抗体が、類天疱瘡ではヘミデスモゾームを構成する BPAG などに対する自己抗体が産生されることは分かっているが、水疱形成を引き起こす病態については未だ不明の点が多い。本研究では、天疱瘡や類天疱瘡などの難治性水疱性疾患の罹患組織中の遺伝子発現を、種々の難治性皮膚疾患（乾癬、皮膚炎、腫瘍など）での遺伝子発現と比較検討することによって、これら難治性皮膚疾患の病態形成機序を明らかにするとともに、新たな治療法の開発に寄与したい。

動脈硬化、血栓症、病的血管新生等の血管疾患発症の機序を培養細胞系で構築し解明するには、炎症性因子の刺激に応答する血管内皮細胞での遺伝子発現制御機構を綿密に再現し組み立てることが重要である。そこでまず、内皮細胞における刺激応答をえ包括的に理解するためマイクロアレイを用いてトランスクリプトーム解析を進め、機序解明に呈することを目的とした。

B.研究方法

1. 遺伝子発現プロファイルデータベースの拡充、データ交換

1-1 データベースの拡充

GeneChip（Affymetrix 社製）U133 アレイのデータに加えて U133Plus2 アレイ、CodeLink 55K アレイなどのデータを取得し、多種のアレイなどにも対応可能なデータベース構築を行なった。データベースは、

PostgreSQL 7.2.3 RDBMS (Relational DataBase Management System)環境の上に構築した。一方、データを閲覧するためのブラウザーは Windows 環境では Internet Explore 6.0 あるいは MacOS9 または X については Internet Explorer 5.1 以上である。

1-2 知的統合による遺伝子、蛋白質相互作用探索の研究

ヒト正常臓器あるいは細胞株について GeneChip（Affymetrix 社製）U133 アレイを用いたトランスクリプトーム解析実験データに対し、情報解析するためのデータベースを中心とする情報処理基盤を構築し、情報処理解析ツールを作成しつつ、技術開発を行い、検証を行う。特に、文献情報から、単に名前が同時に現れる共起性ではなく、自然言語処理を用い蛋白質名と相互作用関係を表す動詞を含んだ二項関係を抽出し、探索すべき遺伝子あるいは蛋白質の相互作用パートナーを再帰的に探索するアプローチを採用した。

2． 微量組織からのゲノム解析技術の確立

2-1 染色体コピー数の多様性解析

高密度合成オリゴヌクレオチドマイクロアレイを用いて、CNP を精確に検出する技術開発を行った。これらの実験においては全ゲノムの一部を PCR で増幅したものを分析する工程（Representational analysis）が含まれ、こうした過程における実験条件の差によって無視できないノイズデータを含んでしまうという欠点があった。我々はこれらの実験条件に対する補正を行うことで、染色体のゲノム DNA 量を測定する正確に測定する方法を開発し、GIM（ゲノムインバランスマップ）と名づけた。

3． 希少性疾患解析への応用

3-1 慢性疲労症候群の病因解析

2002 年から 2003 年に採取された CFS I 群 7 症例（男性 3 例、女性 4 例）、CFSⅢ群 10 症例（男性 7 例、女性 3 例）、健常者 9 例（男性 3 例、女性 6 例）の末梢血単核球 (PBMCs)より RNeasy Mini Kit（QIAGEN）を用いて total RNA を抽出した。RNA の質は 2100 Bioanalyzer (Agilent Technologies)を用いて確認した。検体は大阪大学内科の倉恒らにより収集された。mRNA 発現解析には、CodeLink Human UniSet 1 20K Bioarray（Amersham Biosciences Corp.）に収載されている約 20000 種類のオリゴ DNA を Duplicate でスポットしたカスタムアレイを用いた。Gene Ontology 解析には Onto-express（http://vortex.cs.wayne.edu/index.htm）、階層的クラスター解析には GeneSpring 7.2（Silicon Genetics）を用いた。さらに転写因子解析には TRANSFAC®（BIOBASE GmbH）を用いた。

3-2 特発性心筋症の解析に関する研究

心臓縮小手術（バチスタ手術）にて得られた不全心筋から RNA サンプルを調整し、DNA chip 解析を行った。その結果によって得られたいくつかの遺伝子についてはマウスモデルを作成し、心不全病態生理への関与を検討した。

3-3 福山型筋ジストロフィーと長期記憶に関連したラット海馬遺伝子の解析

今回我々は FCMD 患者の骨格筋を独自に開発した cDNA マイクロアレイにより発現解析を行い、また α-DGpathy の一疾患である LARGE 変異モデルマウス（*myd* マウス）を用い、α-DGpathy における発現解析及び病態解析を行った。

長期記憶に関する解析では、長期記憶との関連が示唆されている LTP をラット海馬 CA 1 に誘発し経時的遺伝子発現の変化（10 分、30 分、60 分後）を GeneChip（Affymetrix 社）にて網羅的に解析した。統計学的手法を用いて発現変化の著しい遺伝子群の空間的、機能的解析及びパスウェイ分析を行った。

3-4 難治性水疱性疾患およびその他の難治性皮膚疾患における罹患組織の遺伝子発現プロファイルに関する研究

前年度までに天疱瘡や類天疱瘡をはじめ、

コントロールとして種々の難治性皮膚疾患（乾癬、皮膚炎、腫瘍など）の病変部および健常部の遺伝子発現プロファイルを比較した。遺伝子発現プロファイルの違いが病態形成にどのように関与するのかを、表皮細胞、免疫担当細胞、線維芽細胞の培養系を用いて検証を行い、新規の機序や分子を明らかにする。

3-5　血管内皮細胞における血管疾患（血栓、動脈硬化症）誘発刺激応答に関する研究

生理的な刺激反応を維持し、かつ安定的な供給が見込めるものとしてヒト臍帯静脈血管内皮細胞 (HUVEC)を用いた。Sub-confluent HUVEC にトロンビン (1.5U/ml) あるいは　tumor necrosis factor (TNF)-a (10ng/ml) vascular endothelial cell growth factor (VEGF) (50ng/ml) を添加し、経時的に total RNA を回収した。網羅的なマイクロアレイ遺伝子解析は Affymetrix の奨める方法に従い行なった。Down syndrome critical region (DSCR)-1 遺伝子の強制発現系は、自己増幅能欠損型のアデノウイルスを用いて行った。またマウスへの DSCR-1 発現も東京大学動物実験指針にのっとり、行った。

（倫理面への配慮）

厚生労働省の「遺伝子解析研究に付随する倫理問題等に対応するための指針」、３省庁合同の「ヒトゲノム・遺伝子解析研究に関する倫理指針(平成１３年３月２９日文科省、厚労省・経済省告示第一号）に従い指針に適合する倫理委員会の承認を得、各臨床機関において患者からの血液あるいは組織試料等の匿名化をはかった試料に対して得られた遺伝子発現解析実験データに対して情報処理解析を行なう。

C.研究結果

１．遺伝子発現プロファイルデータベースの拡充、データ交換

1-1　データベースの拡充

本年度は、GeneChip (Affymetrix社製) U133アレイのデータに加えてU133Plus2アレイ、CodeLink 55Kアレイなどさらに多種のアレイなどにも対応可能なデータベース構築を行なった。最大55,000遺伝子についての発現プロファイルデータの収集に着手し、データベース化をはかっている。１遺伝子に複数のプローブが存在する場合にバリアントにより発現プロファイルは異なる可能性があることから、今後は個別のプロファイリングが必要と考えられる。U133アレイとU133Plus2アレイはプローブ配列は同一であるにも関わらず、シグナル値が相関しない例も認められた。

1-2　知的統合による遺伝子、蛋白質相互作用探索の研究

文献データベースで最も大きく広く使われているMedlineには膨大なデータが蓄積されているのでそこに含まれている抄録600万件を利用できるようなシステムをまず構築した。一方で、入手可能なデータベースから疾患解析に必要な用語を抽出した。次に、文献から必要な情報を信頼度高く抽出し、分かりやすい形で表示する自然言語処理技術を応用した文献検索システムの構築をPCクラスター上で行った。脊椎動物に対して蛋白質・遺伝子名辞書の用語の数を二十万件にまで拡大した。相互作用を表す動詞を整備することにより、２対の蛋白質・遺伝子名の相互作用を文献から抽出することにより、疾患遺伝子探索を試みられるようにした。

２．微量組織からのゲノム解析技術の確立

2-1　染色体コピー数多様性解析

マイクロアレイにおけるサンプルとコントロールのシグナル比を測定することにより、染色体上の最大約 10 万箇所のポイントにおいて 1 コピーの変化が見られる解像度で見ることができるようになった。染色体上の約 1 万箇所を見ることの出来るオリゴヌクレオチドマイクロアレイを用いて男

女のゲノムより得られるシグナルを比較したところX染色体の本数に応じた変化が検出され1コピーの差でも検出可能であった。また従来の CGH アレイとの比較では、X 染色体とそれ以外の常染色体でのシグナル強度比(median±SD)を求めると、GIM 法では 0.99±0.09/0.63±0.08 であるのに対して、CGH アレイでの測定では 1.00±0.09/0.71±0.14 になる。したがって従来の CGH アレイよりも明確に遺伝子コピー数を識別することができた。

さらにより高密度な染色体上の約 10 万箇所を見ることの出来るオリゴヌクレオチドマイクロアレイの解析でも同様に X 染色体の本数の差が明確に検出できると同時に、一方の検体を株化されたリンパ球にすることで免疫グロブリン遺伝子の再構成による局所的なコピー数の変化も検出することができた。またそれ以外の染色体でも 100Kb 以上の領域において複数のプローブにおけるシグナル比の有意な変化が捉えられるところが数箇所みられコピー数の多型と考えられた。そのうちの一つは血球における発生、分化に関わるとされる prominin1(CD133) という遺伝子を含む 100Kb 超の領域であった。これらの手法によって CNP という小さなコピー数変化でも、高密度オリゴヌクレオチドマイクロアレイを用いたシグナル評価法によって十分解析可能であることが示された。

3．希少性疾患解析への応用

3-1　慢性疲労症候群の病因解析

慢性疲労症候群と健常者の遺伝子発現プロファイルは有意に異なっており、CFS Ⅰ群とⅢ群および健常群を識別する遺伝子として細胞接着、免疫応答、炎症反応関連の遺伝子が有意に多かった。共通なプロファイルを示す遺伝子を制御する転写因子の同定を試みたところ、CFS 群と健常群を識別する遺伝子にはその上流配列に転写因子として E2F1 の結合配列を有するものが多数認められた。

3-2　特発性心筋症の解析に関する研究

本年度は、不全心で発現の亢進していた遺伝子である酸化酵素 12-lipoxygenase に注目しマウスモデルなどを用いて解析を進めた。まず培養系の実験では、12-lipoxygenase の活性が低酸素によって誘導されること、その活性上昇が心筋細胞のアポトーシスを誘導することを確認することができた。12-lipoxygenase ノックアウトでは、心筋梗塞後のアポトーシス抑制、線維化、リモデリングの縮小がみられた。

3-2　神経筋疾患の解析

FCMD 及び *myd* の筋線維では筋分化過程後期に発現する転写因子である *myf 6* や *myosin heavy chain2,7* の発現低下をきたし、また電顕画像において、筋細胞の核及び筋線維の成熟障害を呈した。また神経筋接合部では免疫染色および電顕的観察において、著しい異常像を認め、正常ではおこる αDG の凝集も見られなかった。

LTP 誘発後 10 分後より 30 分後にかけて著しい遺伝子発現の変化が見られ、それらには糖鎖修飾酵素、細胞骨格、Rab family、ECM 成分、に関連する遺伝子が特異的に関与していることが見出された。またパスウェイ分析では c－*myc* が遺伝子発現の変化のあった転写カスケードに関与している可能性が示唆された。

3-3　難治性水疱性疾患およびその他の難治性皮膚疾患における罹患組織の遺伝子発現プロファイルに関する研究

今年度は、1）乾癬における Stat3 の発現上昇、2）ヒト乳頭腫ウイルス発癌における P16 ならびに hTERT 蛋白の発現上昇、の２点を明らかにした。乾癬表皮の核内で Stat3 が高発現していることは表皮の過増殖を引き起こす病態形成に密接に関与していると考えられた。またヒト乳頭腫ウイルス(HPV)の中の high risk 型は上皮細胞に感染すると高い発癌性を有することが知られている。我々は high risk HPV の感染によって癌化した表皮細胞は、p16 および Human

telomerase reverse transcriptase (hTERT)を高発現していることを見出した。

3-4 血管内皮細胞における血管疾患（血栓、動脈硬化症）誘発刺激応答に関する研究

血管新生関連刺激に伴う遺伝子変動の網羅的探索、時間的制御の比較解析から、これら活性化シグナルは転写カスケードを介して経時的に接着因子、増殖因子の発現誘導へと推移していくことが共通に認められた。そして特に刺激初期においては thrombin と VEGF のシグナルの共通項が大きいこと、VEGF 及び thrombin の共通の反応として転写因子 NF-AT の活性化が挙げられ、その活性化のネガティブフィードバック因子である DSCR-1 が強力に誘導されてくること、また DSCR-1 が安定に存在した状態では血管新生、炎症を含む内皮の活性化状態を顕著に抑制することが認められた。

D.考察

1．発現プロファイルデータベースの拡充、データ交換

1-1 データベースの拡充

本年度は、GeneChip（Affymetrix 社製）U133 アレイのデータに加えて U133Plus2 アレイ、CodeLink アレイなどさらに多種のアレイなどにも対応可能なデータベース構築を行なった。データの互換性はプローブ配列が同一の U133 アレイと U133Plus2 アレイであっても完全には保証されなかった。個々のプローブの feature サイズが小さくなり、アレイ上での配置の変更なども影響したものと考えられた。

1-2 知的統合による遺伝子、蛋白質相互作用探索の研究

ユーザビリティの向上を考慮して再帰計算のできる文献情報処理技術を開発・改良した。相互作用の抽出のための技術として相互作用を記述する文章の鋳型を用いた抽出方法と、統計的分析に基づく抽出法を組み合わせて用い、さらに文章の鋳型を用いた抽出方法において、DIPRE アルゴリズムを応用することにより自動化を図った。さらに相互作用を表す単語のオントロジーの導入により相互作用の分類を可能とした。なお、本技術開発が網羅性に主体をおいたことにもよるが、もともとの人間が記述する文章におけるシノニムが混乱しているため正答率は、８５％以下にとどまった。また本研究では様々な相互作用を扱ったが、蛋白質同士の結合で重要な bind および interact に相互作用を限定すれば、相互作用パートナーとして蛋白質同士が物理的に相互作用する状態を表すことができる。

今後は様々な解析を積極的に行い、プログラム、データにおけるシノニムエラー、機械学習に起因するエラーを取り除き、早期公開をめざすとともに、様々な技術を融合化することによって知識ベースのより進んだシステムの構築をめざす必要がある。

2．発現プロファイル解析技術の確立

発現プロファイル解析から染色体コピー数変異が遺伝子発現量に相関することを明らかにし、報告してきた（緑川、2004）。コピー数の多様性がゲノム全体にどの程度存在しており、疾患と関連するかについては不明であるが、網羅的な解析手法がまさに求められており、本研究においてゲノタイピング用に開発された短鎖オリゴヌクレオチドアレイを用いて１アレルの違いを定量的に検出できるアルゴリズム GIM を開発できた。CNP の頻度について多くの人種についてそれぞれ解析を進めて、HapMap プロジェクトの SNP タイピングデータとの関連についても検証していく必要があると考えられた。また CNP の同定には複数のプローブが連続的に変動することが求められるので、さらに高密度なアレイによる解析が求められる。

3．希少性疾患解析への応用

DNA チップを用いた網羅的遺伝子発現の解析とそれらに基づくマウスモデルの作成･解析によって、いくつかの心不全関連遺伝子が同定された。心不全増悪因子として同定された酸化酵素 12-lipoxygenase はこ

れまで血球系での役割は検討されていたが、心臓における役割は知られていなかった。本研究により心筋細胞における 12-lipoxygenase が心不全の病態生理に深く関わっていることが明らかとなった。

神経筋接合部由来の筋分化シグナルが基底膜障害により不完全となり、筋線維の成熟障害を起こすことが、α-DGpathy の重要な病態生理である可能性が FCMD の発現解析を通じて示唆された。また、ラット海馬 CA 1 領域では、長期記憶成立には 10 分後、30 分後での遺伝子発現の変化が著しいこと、またその変化にはある機能的遺伝子群の変化が関与していると考えられた。

天疱瘡および類天疱瘡における新治験を今年度は得ることはできなかったが、発現プロファイル解析から見出された変動遺伝子の解析を引き続き進める予定である。HPV 感染のマイクロアレイ解析データは p16 および hTERT の高発現は HPV 癌化のよい指標となることを示している。

原発性高脂血症では動脈硬化症の進展が臨床的な問題となる。内皮細胞の血管新生関連刺激に伴う遺伝子変動の網羅的探索、時間的制御の比較解析から、これら活性化シグナルは転写カスケードを介して経時的に接着因子、増殖因子の発現誘導へと推移していくことは共通であるが、特に早期においては thrombin と VEGF のシグナルの共通項が大きいこと。そして共通の反応として NF-AT の活性化が挙げられ、その活性化のネガティブフィードバック因子である DSCR-1 が強力に誘導されてくること。そして DSCR-1 が安定に存在した状態では血管新生、炎症を含む内皮の活性化状態を顕著に抑制する、しかも細胞内因子であることから、これが将来毒性の少ない新規な抗血管病治療法として用いられることが期待される。

E..結論

本年度は、GeneChip（Affymetrix 社製）U133 アレイのデータに加えて U133Plus2 アレイ、CodeLink アレイなどさらに多種のアレイなどにも対応可能なデータベース構築を行なった。今後は様々な臨床データへの応用を展開していくことによって、臨床研究者がより使いやすいシステムの要件を明確化し、この分野でより先端的な研究を行なうことができる。

井原らは GeneChip (Affymetrix 社製) U133 アレイのデータを解析対象とした再帰計算ができる文献情報システムを作成した。従来の文献検索ではできないようなクエリー間での相互作用パートナーの探索などに有力な手法と思われる。

高密度タイピングアレイを用いて染色体コピー数をアレル別に定量するアルゴリズム GIM を開発し、X 染色体数の異常や免疫グロブリン遺伝子座の再構成を検出することが出来た。今後疾患との関連の解明が待たれる。

油谷らは慢性疲労症候群の病因解析において、末梢血球から抽出した RNA の発現プロファイルにより患者群と健常群を識別することに成功した。

小室らは不全心筋を用いた網羅的な遺伝子発現解析は、特発性心筋症の原因遺伝子の同定や心不全病態生理の解明に有用である。心不全増悪因子として同定された酸化酵素 12-lipoxygenase は今後心不全治療の新たなターゲットとして同定した。

戸田らは FCMD では原因遺伝子フクチンのコードするフクチン蛋白の機能喪失により、筋細胞膜のαジストログリカンが糖鎖修飾をうけられず、筋基底膜成分のラミニンとの結合が弱くなり、筋基底膜の破綻が生じる結果、αジストログリカンが機能的に作用する筋細胞膜及び神経筋接合部の分化遅延をもたらし FCMD 及びα-DGpathy の病態を作り出していることを示した。また、ラット海馬 CA 1 領域では、経時的かつ網羅的遺伝子発現解析をすることで、多次元解析が可能となり、長期記憶に関連する機能

的遺伝子群やそのパスウェイが抽出された。

古江らはプロファイル解析を進め遺伝子発現に関する難治性皮膚疾患プロファイル一覧表を作成するとともに、様々な皮膚疾患における病態解析をさらに進めた。

各種血管疾患誘発因子を用いたトランスクリプトーム解析から、遺伝子応答は各種因子で個別に生じる早期の遺伝子応答や多くの共通誘導遺伝子が存在していること、このうち thrombin, VEGF 刺激で共通に最も強く早期に誘導される DSCR-1 は血管内皮細胞活性化のネガティブフィードバック因子として作用し、これを安定的に発現させることによって、血管疾患（病的血管新生、動脈硬化）に関わる多くの遺伝子の発現抑制をもたらし、この因子が病的状態の緩和に重要な役割をもつことが明らかとなった。

F.健康危険情報

本年度の研究成果からは健康に害となる情報は得られていないが、発現プロファイル解析から病態のマーカーとなる遺伝子群が見出されつつある。

G.研究発表

1. 論文発表

（分担研究者の項参照）

2. 学会発表

（分担研究者の項参照）

H.知的財産権の出願・登録状況（予定も含む）

1. 特許取得

（分担研究者の項参照）

2. 実用新案登録

なし

3. その他

なし

Ⅲ　平成16年度分担研究報告書

厚生労働科学研究費補助金（難治性疾患研究事業）
分担研究報告書

発現プロファイル解析からの疾病遺伝子同定に関する研究

分担研究者　油谷浩幸　東京大学国際・産学共同研究センター教授

研究要旨
発現プロファイル解析は病因解明の有効な手法として期待されている。本年度の研究においては１）染色体コピー数の多様性解析、２）転写ネットワークからの疾病遺伝子同定について慢性疲労症候群を例にとり検討した。転写制御に共通に関わる転写因子を同定することにより、疾病罹患群における遺伝子発現プロファイルを特徴づける識別遺伝子群から疾患において変動しているパスウェイを探索する手法を開発した。

A. 研究目的

ヒトの疾患罹患性、薬剤応答性を規定する遺伝子多型としては SNP が最もよく知られているが、最近になってヒトには数百キロ bp の領域に及ぶコピー数の多型 CNP (copy number polymorphism)が従来の予想よりも高頻度に存在することが明らかになってきた。こうした長い領域にわたるコピー数の変異（多型）がヒト表現形質におよぼす意義や、疾患との関連性について少数の疾患で報告されている。例えば、神経難病の一つである Charcot-Marie-Tooth 病は PMP22 遺伝子重複(および欠失)に起因するほか、コピー数の変異（多型）がパーキンソン病等の疾患の原因となりうることも知られており、疾患の診断および治療という面においてもその解明が有益であることが期待され、ゲノムワイドに染色体コピー数を評価する手法の開発が待たれていた。現在までに異なるアレイ解析手法を用いた 2 つの報告があるが、両者の結果が一致するところは限られておりより正確，かつ網羅的に行うことが課題となっている。

癌やダウン症などの先天性疾患においては、そのコピー数の異常が重要であるとされ、その網羅的な解析には BAC(bacterial artificial chromosome、大腸菌人工染色体)を貼り付けた CGH(comparative genomic hybridization) アレイが従来最もよく使われてきた手法であった。しかしながら BAC は 100kb 以上におよぶ長いプローブであり、そのクローニングはオリゴヌクレオチドのように工業的に合成ができず、アレイ上にスポット可能なプローブ数（ゲノム上での密度）にも限界があった。また BAC のアレイはプローブよりも小さな～10kb 程度の領域の変化は捕らえることができない。近年短いオリゴヌクレオチドを用いたゲノムタイピングマイクロアレイを利用した方法も含めて、合成オリゴヌクレオチドマイクロアレイを使ったコピー数解析がいくつか報告されているものの、これらの方法ではデータが多くのノイズを含んでしまうために CNP のような 1 コピーの差を見るには不適な状態であった。ゲノムタイピングマイクロアレイを利用して CNP を解析する方法の開発を目指した。

マイクロアレイ解析は３万を越えるヒト遺伝子転写産物についての網羅的発現解析を可能とした。発現プロファイル解析により新たな病型分類あるいは病態メカニズムの理解が進むことが求められている。デー

タマイニングにはクラスタ解析や主成分分析などの数理解析のみならず、遺伝子の機能情報やパスウェイ解析により病態への理解が深まると期待されている。慢性疲労症候群（chronic fatigue syndrome, CFS）は6ヶ月以上継続的・断続的に全身倦怠感を訴える原因不明の病態であり、患者末梢血球を採取後 RNA を抽出し、マイクロアレイ解析を行った。

B. 研究方法

<u>染色体コピー数の多様性解析</u>

高密度合成オリゴヌクレオチドマイクロアレイを用いて、CNP を精確に検出する技術開発を行った。これらの実験においては全ゲノムの一部を PCR で増幅したものを分析する工程（Representational analysis）が含まれ、こうした過程における実験条件の差によって無視できないノイズデータを含んでしまうという欠点があった。我々はこれらの実験条件に対する補正を行うことで、染色体のゲノム DNA 量を測定する正確に測定する方法を開発した（GIM：ゲノムインバランスマップと名づけた）。

<u>慢性疲労症候群の発現プロファイル解析</u>

検体・RNA 抽出　2002 年から 2003 年に採取された CFS I 群 7 症例（男性 3 例、女性 4 例）、CFSⅢ群 10 症例（男性 7 例、女性 3 例）、健常者 9 例（男性 3 例、女性 6 例）の末梢血単核球(PBMCs)より RNeasy Mini Kit（QIAGEN）を用いて total RNA を抽出した。RNA の質は 2100 Bioanalyzer (Agilent Technologies)を用いて確認した。検体は大阪大学内科の倉恒らにより収集された。

マイクロアレイ解析　mRNA 発現解析には、CodeLink Human UniSet 1 20K Bioarray（Amersham Biosciences Corp.）に収載されている約 20000 種類のオリゴ DNA を Duplicate でスポットしたカスタムアレイを用いた。マイクロアレイのアッセイは、メーカーのプロトコールにしたがった。すなわち、total RNA から 2 本鎖 cDNA を合成し、T7 RNA Polymerase による in vitro transcription(IVT)反応を行う際に、ビオチンラベル dUTP を取り込ませて標識を行った。そして、Bioarray 上でハイブリダイゼーションを行った後、Cy5-ストレプトアビジンと反応させた。洗浄後、Agilent G2565 マイクロアレイスキャナー（Agilent Technologies）で蛍光シグナルを読み取った。蛍光シグナルの数値化には、CodeLink Expression Analysis Version 4.0 を用い、Duplicate のデータは平均化した。また、Gene Ontology 解析には Onto-express（<u>http://vortex.cs.wayne.edu/index.htm</u>）、階層的クラスター解析には GeneSpring 7.2（Silicon Genetics）を用いた。さらに転写因子解析には TRANSFAC®（BIOBASE GmbH）を用いた。

（倫理面への配慮）

本研究ではアレルの識別のためにゲノタイプ情報を使用するが、全て連結不可能匿名化されたゲノム DNA を用いた解析であり、東京大学先端科学技術研究センター研究倫理審査委員会において承認されている。

C. 研究結果

<u>染色体コピー数の多様性解析</u>

マイクロアレイにおけるサンプルとコントロールのシグナル比を測定することにより、染色体上の最大約 10 万箇所のポイントにおいて 1 コピーの変化が見られる解像度で見ることができるようになった。染色体上の約 1 万箇所を見ることの出来るオリゴヌクレオチドマイクロアレイを用いて男女のゲノムより得られるシグナルを比較したところ X 染色体の本数に応じた変化が検出され 1 コピーの差でも検出可能であっ

た（図1）。また従来の CGH アレイとの比較では、X 染色体とそれ以外の常染色体でのシグナル強度比(median±SD)を求めると、GIM 法では、0.99±0.09/0.63±0.08 であるのに対して、CGH array での測定では、1.00±0.09/0.71±0.14 になる。したがって従来の CGH アレイよりも、明確に遺伝子コピー数を識別することができた（図1)。

さらにより高密度な染色体上の約 10 万箇所を見ることの出来るオリゴヌクレオチドマイクロアレイの解析でも同様に X 染色体の本数の差が明確に検出できると同時に、一方の検体を株化されたリンパ球にすることで免疫グロブリン遺伝子の再構成による局所的なコピー数の変化も検出することができた（図2）。またそれ以外の染色体でも 100Kb 以上の領域において複数のプローブにおけるシグナル比の有意な変化が捉えられるところが数箇所みられ、コピー数の多型と考えられた。そのうちの一つは血球における発生、分化に関わるとされる prominin1(CD133)という遺伝子を含む100Kb 超の領域であった。これらの手法によって CNP という小さなコピー数変化でも、高密度オリゴヌクレオチドマイクロアレイを用いたシグナル評価法によって十分解析可能であることが示された。

慢性疲労症候群の発現プロファイル解析

解析を行った 26 検体全てにおいて Positive と判定された 15965 プローブを用いて階層的クラスタ解析を行った結果、CFS 患者群と健常者群が分類することができた。しかし、Ⅰ群とⅢ群を完全に分類することはできなかった。また、3 群間で差がある遺伝子を抽出するために、ANOVA と Kruskal-Wallis 解析を行った。その結果、両解析法で共通して P<0.01 であった 1650 遺伝子を得た。これらの遺伝子を Gene Ontology ごとに分類したところ、細胞接着、免疫応答、炎症反応関連の遺伝子が有意に多かった。

次に、CFS 群と健常者で差がある遺伝子を抽出するために、Ⅰ群と健常者、Ⅲ群と健常者の Signal-to-noise 解析を行い、それぞれ、上位 400 遺伝子を選択した。そして、Ⅰ群とⅢ群で共通して健常者と差がある 52 遺伝子を抽出した。IFNAR1 や TNFSF13B 等の炎症関連遺伝子のほか、ストレス応答に関与する MAPK14 が CFS 群で有意に高発現していた。また、ケモカイン（CCL28)、自然免疫や 2 価イオンの輸送に関わる NRAMP2 等が健常者の群で高発現していた。しかし、今回の我々の解析では、CFS で高発現していると報告されている TGF beta において、有意な差は認められなかった。

また、Signal-to-noise 解析で差があった上位 400 遺伝子について、転写開始点の上流 1kb にどのような転写因子の結合モチーフがあるか検索を行ったところ、細胞周期に関わる E2F1 がⅠ群、Ⅲ群で共通して有意に高かった（Ⅰ群: P=0.001、Ⅲ群: P=0.012)。

D. 考察

発現プロファイル解析から染色体コピー数変異が遺伝子発現量に相関することを明らかにし、報告してきた（緑川、2004)。コピー数の多様性がゲノム全体にどの程度存在しており、疾患と関連するかについては不明であるが、網羅的な解析手法がまさに求められており、本研究においてゲノタイピング用に開発された短鎖オリゴヌクレオチドアレイを用いて 1 アレルの違いを定量的に検出できるアルゴリズム GIM を開発できた。CNP の頻度について多くの人種についてそれぞれ解析を進めて、HapMap プロジェクトの SNP タイピングデータとの関連についても検証していく必要があると考えられた。また CNP の同定には複数

のプローブが連続的に変動することが求められるので、さらに高密度なアレイによる解析が求められる。

慢性疲労症候群と健常者の遺伝子発現プロファイルは有意に異なっており、CFS I群とⅢ群および健常群を識別する遺伝子として細胞接着、免疫応答、炎症反応関連の遺伝子が有意に多かった。CFS 群と健常群を識別する遺伝子には転写因子として E2F1 の結合配列を有するものが多数認められた。

E. 結論

高密度タイピングアレイを用いて染色体コピー数をアレル別に定量するアルゴリズム GIM を開発し、X 染色体数の異常や免疫グロブリン遺伝子座の再構成を検出することが出来た。

末梢血由来の細胞 RNA の発現プロファイル解析により、慢性疲労性症候群患者と対照健常者を識別することが出来た。

F. 健康危険情報

なし

G.研究発表

1． 論文発表

1. Ge X, Yamamoto Shogo, Tsutsumi Shuichi, Midorikawa Yutaka, Ihara Sigeo, Wang SM, Aburatani H. Interpreting expression profiles of cancers by genome-wide survey of breadth-of-expression in normal tissues. **Genomics** in press
2. Kano M, Tsutsumi S, Kawahara N, Wang Y, Mukasa A, Kirino T, Aburatani H. A meta-clustering analysis indicates distinct pattern alteration between two series of Gene Expression profiles for induced ischemic tolerance in rats. **Physiological Genomics.** in press
3. Liu HJ, Moroi S, Yasumoto S, Koga T, Masuda T, Chen QJ, Tu YT, Fure M, Aburatani H, Urabe K. Expression of elafin in extramammary Paget's disease. **Br J Dermatol.** 152(3):578-9. 2005
4. Hanada T, Tanaka K, Matsumura Y, Yamauchi M, Nishinakamura H, Aburatani H, Mashima R, Kubo M, Kobayashi T, Yoshimura A. Induction of hyper Th1 cell-type immune responses by dendritic cells lacking the suppressor of cytokine signaling-1 gene. **J Immunol.** 174(7):4325-32. 2005
5. Komura D, Nakamura H, Tsutsumi S, Aburatani H, Ihara S. Multidimensional support vector machines for visualization of gene expression data. **Bioinformatics.** 21(4):439-44. 2005
6. Inozume T, Matsuzaki Y, Kurihara S, Fujita T, Yamamoto A, Aburatani H, Shimada S, Kawakami Y. Novel melanoma antigen, FCRL/FREB, identified by cDNA profile comparison using DNA chip are immunogenic in multiple melanoma patients. **Int J Cancer.** 114(2):283-90. 2005
7. Midorikawa Y, Tsutsumi S, Nishimura K, Kamimura N, Kano M, Sakamoto H, Makuuchi M, Aburatani H. Distinct chromosomal bias of gene expression signatures in the progression of hepatocellular carcinoma. **Cancer Res.** 64(20): 7263-70. 2004
8. Minami T, Horiuchi K, Miura M, Abid R, Takabe W, Kohro T, Ge X, Aburatani H, Hamakubo T, Kodama T, Aird WC. VEGF- and thrombin-induced termination factor, down syndrome critical region-1, attenuates endothelial cell proliferation, and angiogenesis. **J Biol Chem.** 279(48):

50537-54. 2004
9. Miyachi H, Ikebukuro K, Yano K, Aburatani H, Karube I. Single nucleotide polymorphism typing on DNA array with hydrophobic surface fabricated by plasma-polymerization technique. **Biosens Bioelectron.** 20(2): 184-9. 2004
10. Kamei Y, Miura S, Suzuki M, Kai Y, Mizukami J, Taniguchi T, Mochida K, Hata T, Matsuda J, Aburatani H, Nishino I, Ezaki O. Skeletal muscle FOXO1 (FKHR)-transgenic mice have less skeletal muscle mass, down-regulated type I (slow twitch / red muscle) fiber genes, and impaired glycemic control. **J Biol Chem.** 279(39): 41114-23. 2004
11. Watanabe T, Akishita M, Nakaoka T, He H, Miyahara Y, Yamashita N, Wada Y, Aburatani H, Yoshizumi M, Kozaki K, Ouchi Y. Caveolin-1, Id3a and two LIM protein genes are upregulated by estrogen in vascular smooth muscle cells. **Life Sci.** 75(10): 1219-29. 2004
12. Takasato M, Osafune K, Matsumoto Y, Kataoka Y, Yoshida N, Meguro H, Aburatani H, Asashima M, Nishinakamura R. Identification of kidney mesenchymal genes by a combination of microarray analysis and Sall1-GFP knockin mice. **Mech Dev.** 121(6): 547-57. 2004
13. Kohro T, Tanaka T, Murakami T, Wada Y, Aburatani H, Hamakubo T, Kodama T. A Comparison of Differences in the Gene Expression Profiles of Phorbol 12-myristate 13-acetate Differentiated THP-1 Cells and Human Monocyte-derived Macrophage. **J Atheroscler Thromb.** 11(2): 88-97. 2004
14. Ogihara T, Asano T, Katagiri H, Sakoda H, Anai M, Shojima N, Ono H, Fujishiro M, Kushiyama A, Fukushima Y, Kikuchi M, Noguchi N, Aburatani H, Gotoh Y, Komuro I, Fujita T. Oxidative stress induces insulin resistance by activating the nuclear factor-kappaB pathway and disrupting normal subcellular distribution of phosphatidylinositol 3-kinase. **Diabetologia.** 47(5): 794-805, 2004
15. Takita J, Ishii M, Tsutsumi S, Tanaka Y, Kato K, Toyoda Y, Hanada R, Yamamoto K, Hayashi Y, Aburatani H. Gene expression profiling and identification of novel prognostic marker genes in neuroblastoma. **Genes Chromosomes Cancer.** 40(2): 120-32. 2004
16. Matsuo K, Galson DL, Zhao C, Peng L, Laplace C, Wang KZ, Bachler MA, Amano H, Aburatani H, Ishikawa H, Wagner EF. NFAT rescues osteoclastogenesis in precursors lacking c-Fos. **J Biol Chem.** 279(25): 26475-80. 2004
17. Fujiwara K, Ochiai M, Ohta T, Ohki M, Aburatani H, Nagao M, Sugimura T, Nakagama H. Global gene expression analysis of rat colon cancers induced by a food-borne carcinogen, 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine. **Carcinogenesis.** 25(8): 1495-505. 2004
18. Hippo Y, Watanabe K, Watanabe A, Midorikawa Y, Yamamoto S, Ihara S, Tokita S, Iwanari H, Ito Y, Nakano K, Nezu J, Tsunoda H, Yoshino T, Ohizumi I, Tsuchiya M, Ohnishi S, Makuuchi M, Hamakubo T, Kodama T, Aburatani H. Identification of Soluble Amino Terminal Fragment of Glypican-3 as a Serological Marker for Early Stage Hepatocellular Carcinoma. **Cancer Research** 64(7): 2418-2423. 2004
19. Nakatani N, Aburatani H, Nishimura K,

Semba J, Yoshikawa T. Comprehensive expression analysis of a rat depression model. **Pharmacogenomics J**. 4(2):114-26. 2004

20. Mukasa A, Ueki K, Ge X, Ishikawa S, Ide T, Fujimaki T, Nishikawa R, Asai A, Kirino T, Aburatani H. Selective expression of a subset of neuronal genes in oligodendroglioma with chromosome 1p loss. **Brain Pathology** 14(1): 34-42. 2004

2. 学会発表

1. 国公私立大学病院医療技術関係職員研修（東京）5/21　DNAチップと病理診断
2. 第14回難病治療研究会（東京）7/8　アレイを用いた機能ゲノム解析
3. LSBM国際シンポジウム（東京）7/13 2nd International Symposium on New Frontiers of Systems Biology and Medicine Novel Biomarker discovery through cancer genomics
4. マイクロアレイ研究会「次世代トランスクリプトーム解析」（東京）8/31 マイクロアレイ技術によるハイスループットバイオロジー
5. 第55回日本皮膚科学会中部支部学術大会（金沢）9/12 特別講演「ゲノム研究は医療をどう変えるか」
6. BioJapan2004シンポジウム「ゲノム・プロテオーム解析の癌診療へのインパクト」（東京）9/28 ゲノム解析の癌研究への応用
7. 日本癌学会ランチョンセミナー（福岡）9/30 Molecular Karyotyping: allelic gene dosage analysis
8. 三菱化学ヘルスケアフォーラム（東京）11/13 癌の分子診断と治療への展開
9. 国際疲労学会（軽井沢）2/11 International Conference on Fatigue Science 2005 Gene Expression Signatures in CFS patients
10. 東京大学国際シンポジウム（東京）2/18 The University of Tokyo International Symposium – Frontiers in Drug Development Genomic Technology in Drug Development

H. 知的財産権の出願・登録状況（予定も含む）

1. 特許取得

2004.9.7「遺伝子コピーの解析方法及び装置」特願 2004-260328 東京大学 TLO （発明者）油谷浩幸、石川俊平、西村邦裕

2. 実用新案登録

なし

3. その他

なし

図1　GIM法による染色体コピー数解析

染色体上で1万箇所を評価できる高密度オリゴヌクレオチドマイクロアレイを用いた我々のコピー数解析（GIM 法）(上)と従来の CGH array(下)による男女間のX染色体（矢印）におけるコピー数変化の解析結果。各アレイの probe のシグナル比を染色体順に並べてある。

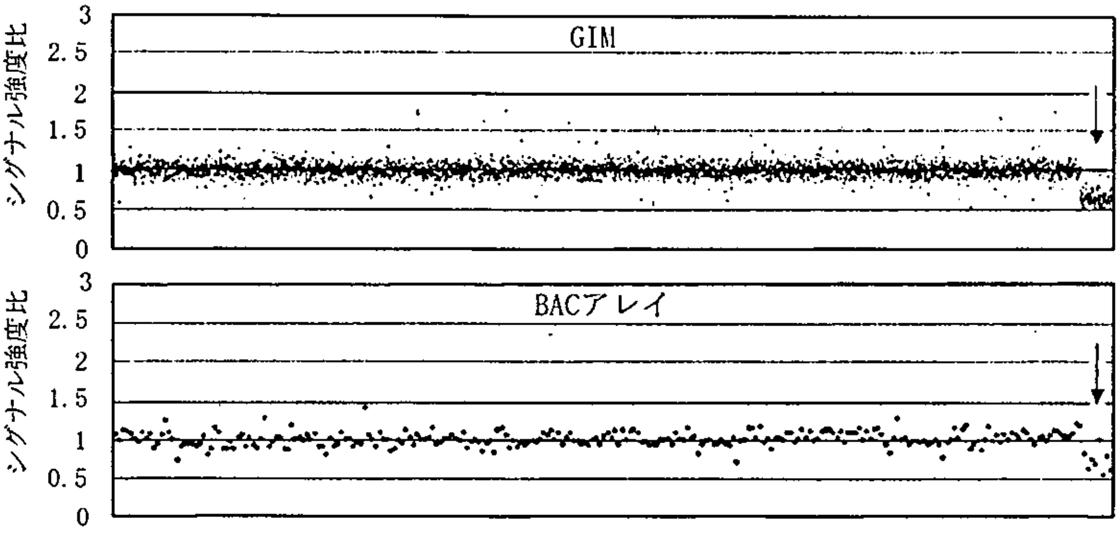
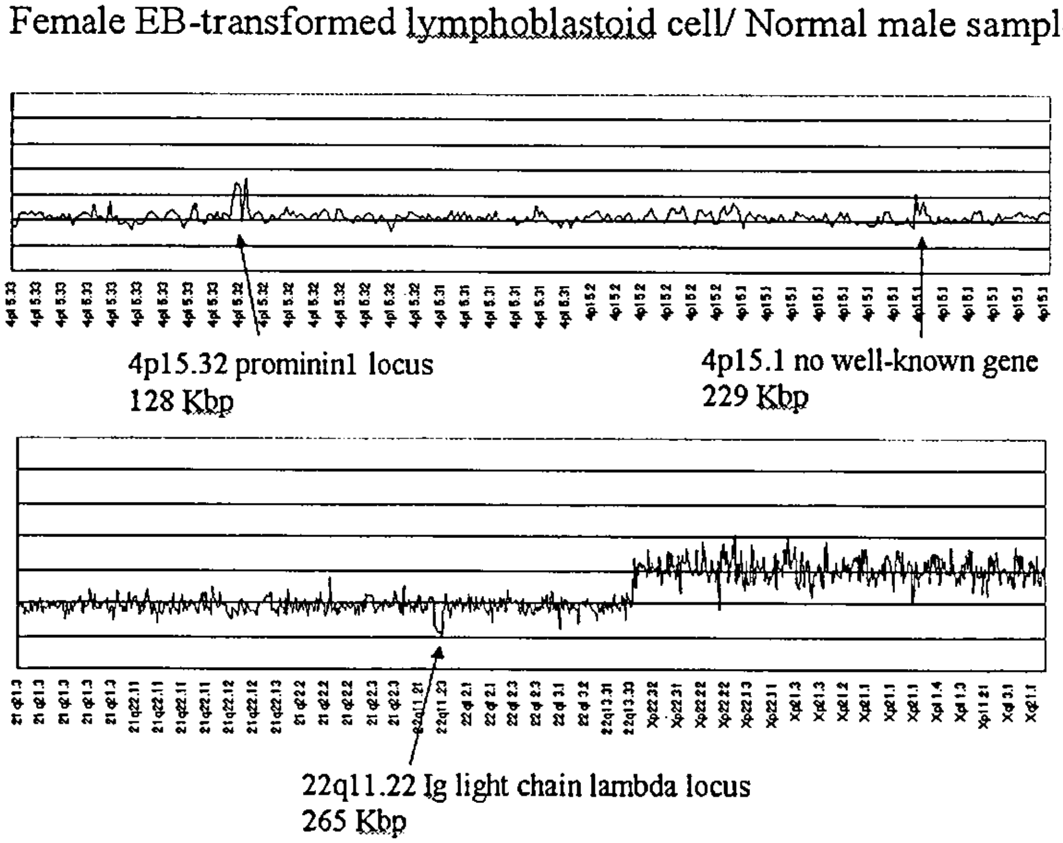


図2　染色体上で 10 万箇所を評価できる高密度オリゴヌクレオチドマイクロアレイの解析を用いたＥＢ株化リンパ球からの女性のゲノムと、正常男性の末梢血から採取したゲノムとの比較。

X 染色体の本数の変化や株化リンパ球のクローナルな免疫グロブリンの再構成によるコピー数変化が捕らえられると同時に（下)、それ以外の領域でも CNP と思われる部分がいくつか検出され、その一つは prominin1(CD133)という遺伝子を含む 100Kb 超の領域であった（上)。