

的化学療法を受けたAML患者66例の骨髓AC133陽性細胞について遺伝子発現プロファイルデータを得た。これら66例の患者中初回化学療法によって完全寛解(complete remission: CR)に到達した症例は51例あり、残り15例は寛解導入に失敗した。後者15例は一年以内に全員死亡しており極めて予後不良なグループといえる。この初回化学療法に成功した患者と失敗した患者の間では遺伝子発現プロファイルは違うのであろうか?それを検討するために4万4千種類のプローブセット(3万3千種類の遺伝子に相当)の発現データ中、両グループ間で発現が有意に異なる遺伝子を選び出し、さらに少なくともどちらかのグループでは高い発現量を示す遺伝子をスクリーニングしたところ、最終的に約30種類の遺伝子が得られた。得られた遺伝子の発現プロファイルはそれぞれが互いに独立なわけではなくよく似たものも多い。そこでcorrespondence analysis法¹²⁾によって、これら30種類の発現量パターンから全体を最も代表する仮想発現プロファイルを3種類抽出した(すなわち代表的な仮想発現パターンそれぞれを作る仮想遺伝子3種類を作成したことになる)。これら仮想遺伝子3種類を軸とした仮想空間に各サンプルを投射したのが図3Aである。その結果興味深いことに治療奏功症例と失敗症例とは空間上不連続な位置に独立して存在することが判った。言い換えると両グループは遺伝子発現プロファイルの面では異なった集団に属するのだ。

このような空間上の位置は患者予後にリンクしているのであろうか?図3Aをよく見てみると、患者サンプルはZ軸上で最も強く分離されていることがわかる。そこで症例の仮想空間上でのZ軸上の値が-0.3未満か-0.3以上かによって大きく2群に分け、その長期生命予後をKaplan-Meier法によって解析してみた。図3Bで明らかなように両群は大きく予後が異なることが判る。しかも様々な方法で予後関連遺伝子を絞り込んだ結果数種類の遺伝子の発現量を測定するだけで患

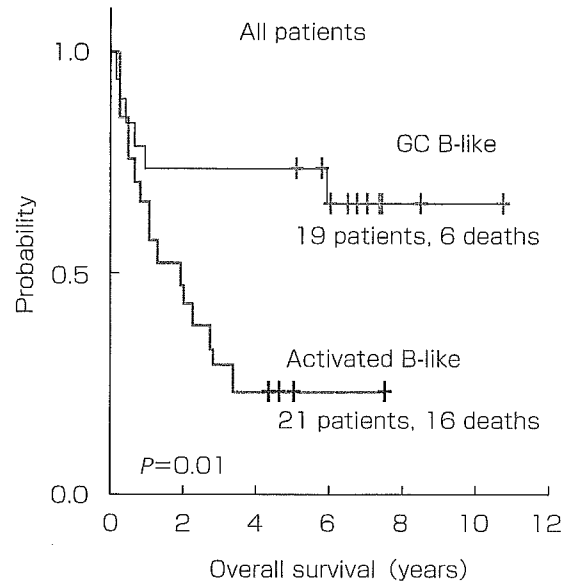


図4. びまん性大細胞型リンパ腫の予後(文献14より改変)

びまん性大細胞型悪性リンパ腫患者を、リンパ節の遺伝子発現プロファイルから「濾胞中心Bリンパ球に発現パターンが似たサンプル群(GC B-like)」と「活性化型Bリンパ球に似た群(Activated B-like)」に分け、両者の生命予後をグラフにした。後者が有意に予後不良群であることが判る。

者の長期予後を予測可能なことまで明らかになった¹³⁾。これらの予後関連遺伝子数は、以前の骨髓単核球による解析で同定された100種類以上の予後関連遺伝子数に比べて明らかに少ない。純化した細胞を標的として解析することで分化傾向などによるノイズが低減し、少数の遺伝子で予後予測可能になったのではないかと考えられる。

4. 悪性リンパ腫の予後予測

多くのリンパ腫は未だなお原因不明であり、原因遺伝子の解析に基づいて患者の治療方針決定・予後予測を行う事は不可能である。そこでDNAチップを用いた発現解析が近年盛んに試みられるようになった。

Bリンパ球由来のびまん性大細胞型リンパ腫は日常の臨床で最も高頻度に遭遇する非Hodgkin

リンパ腫のサブタイプである。同疾患はしばしばCHOP療法などアントラサイクリン系薬剤を中心とした化学療法が有効であるが、寛解となった症例の中には再発を繰り返すものも多い。AlizadehらはDNAチップを用いてびまん性大細胞リンパ腫の中で再発が少ない予後良好群と予後不良群とを鑑別する可能性を検討した¹⁴⁾。この目的のためにまず濾胞中心Bリンパ球および各種悪性リンパ腫のcDNAライブラリーから得られたcDNAなど計17,856遺伝子をスポットしたカスタムチップを用意し、びまん性大細胞型リンパ腫、濾胞性リンパ腫および慢性リンパ性白血病の患者サンプルを用いてDNAチップ解析を行った。また比較のために健常人より得られた各種リンパ組織サンプル、および刺激を加え活性化させた末梢B、Tリンパ球などについても検討している。

その結果、びまん性大細胞リンパ腫には濾胞中心Bリンパ球に遺伝子発現パターンが似ている群と活性化Bリンパ球に似ている群が存在することが示され、しかも両群間で予後に有意な差が認められることが明らかになった。すなわち活性化Bリンパ球に似た細胞からなるリンパ腫患者の5年生存率(16%)は濾胞中心Bリンパ球に似た細胞からなるリンパ腫患者のそれ(76%)に比べて有意に低いのである(図4)。このことはDNAチップによる解析で非Hodgkin悪性リンパ腫の新たなサブグループが定義可能なこと、しかもその分類が予後判定に有意な情報を与えることを示唆しており、今後の臨床の場におけるDNAチップの新たな可能性を示したものとして意義深い。またRosenwaldらも基本的にAlizadehらと一致する解析結果を報告している。

ところがShippらが行った解析では異なった結果が得られている¹⁵⁾。彼らはやはりDLBCL (diffuse large B-cell lymphoma)患者58例の腫瘍リンパ節よりmRNAを抽出し、Affymetrix社HU6800DNAチップによる遺伝子発現プロファイルを得た。全症例を「完全寛解を維持している群」

32例と「腫瘍死あるいは再発中の群」26例に分け、両群を最もよく選別する遺伝子セットを抽出した。さらにこれら遺伝子の発現量から各患者予後を予測した。その結果予後良好群と不良群とを選別する13種類の遺伝子が同定された。これら13種類の遺伝子の発現量からweighted-vote法によって患者予後を予測した際、「良好群」と診断された症例の5年生存率は70%であり、不良群のそれ(12%)より有意($P < 0.00004$)に良好であった。ShippらのデータにおいてもRosenwaldらの提唱するグループでの層別化法は有効に患者を分けることができたが、驚くべきことにその階層化は患者予後にリンクしていなかった。この結果は、DNAチップによる階層化が未だ再現性の悪いものであることを明瞭に示している。この不一致の原因は、異なったDNAチップシステムを用いている点および解析症例数の少なさにあると予想される。

おわりに

DNAチップによる予後予測法の開発は、いわば発癌原因遺伝子の細胞内における最終的な表現形である遺伝子発現を「網羅的発現解析データ」というフィルターを通して類推しパターン分類している、といえるのではないだろうか？ 現行のDNAチップ実験は手技も煩雑であり費用もかかるが、パターン認識の精度が上がるに伴い少ない遺伝子セットで分類できるようになると予想される。そのためには何よりも大量の検体の解析データが重要であろう。

文 献

- 1) International Human Genome Sequencing Consortium: Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945, 2004.
- 2) Dhanasekaran SM, et al: Delineation of prognostic biomarkers in prostate cancer. *Nature* 412: 822-826, 2001.
- 3) van't Veer LJ, et al: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415: 530-536,

- 2002.
- 4) Duggan DJ, et al : Expression profiling using cDNA microarrays. *Nat Genet* 21 : 10-14, 1999.
 - 5) Grimwade D, et al : The importance of diagnostic cytogenetics on outcome in AML : analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* 92 : 2322-2333, 1998.
 - 6) Bennett JM, et al : Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann Intern Med* 103 : 620-625, 1985.
 - 7) Yagi T, et al : Identification of a gene expression signature associated with pediatric AML prognosis. *Blood* 102 : 1849-1856, 2003.
 - 8) Bullinger L, et al : Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 350 : 1605-1616, 2004.
 - 9) Valk PJ, et al : Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* 350 : 1617-1628, 2004.
 - 10) Miyazato A, et al : Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. *Blood* 98 : 422-427, 2001.
 - 11) Tsutsumi C, et al : DNA microarray analysis of dysplastic morphology associated with acute myeloid leukemia. *Exp Hematol* 32 : 828-835, 2004.
 - 12) Fellenberg K, et al : Correspondence analysis applied to microarray data. *Proc Natl Acad Sci USA* 98 : 10781-10786, 2001.
 - 13) Mano H : Stratification of acute myeloid leukemia based on gene expression profiles. *Int J Hematol* 80 : 389-394, 2004.
 - 14) Alizadeh AA, et al : Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403 : 503-511, 2000.
 - 15) Shipp MA, et al : Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 8 : 68-74, 2002.
-

分子細胞治療

別刷

発行：株式会社 先端医学社
〒103-0007 東京都中央区日本橋浜町2-17-8 浜町花長ビル

マイクロアレイによる網羅的遺伝子発現解析でもたらされた新知見

間野博行*

ヒトゲノムプロジェクトが終了しわれわれのもつ蛋白をコードする遺伝子の全体像が明らかになりつつある。DNA マイクロアレイを用いることで、これら数万種類の遺伝子の発現量を任意の細胞・組織間で簡便にくらべることが可能になった。現在、急性骨髄性白血病や骨髄異形成症候群などさまざまな造血器悪性腫瘍の患者サンプルをマイクロアレイで解析し、得られた遺伝子発現プロファイルのなかから患者の長期予後にリンクするものをスクリーニングする試みが精力的におこなわれている。同定された予後関連遺伝子群の発現量をもとに各患者の予後を予測することがこれらのプロジェクトの目標であり、今後は遺伝子発現プロファイルにもとづいた造血器悪性腫瘍の新しい分類法が提案されると期待される。

はじめに

約 30 億塩基対におよぶヒトゲノムの核酸配列を決定する大規模事業である「ヒトゲノムプロジェクト」がついに 2003 年 4 月に終了宣言をおこない、ヒト染色体の euchromatin 領域のほぼ完全な塩基配列が決定された (<http://www.ncbi.nlm.nih.gov/genome/seq/>)。ヒトのもつ蛋白をコードする総遺伝子数はおそらく 2 万～2 万 5 千種類前後になると予想されている¹⁾。かつて夢であったヒトの全遺伝子プールの全貌がいよいよ明らかになるう

としており、ポストゲノム時代が訪れたといえる。

この膨大な遺伝子情報を有効に使う技術であるゲノミクス・プロテオミクスが、造血器悪性腫瘍の解析においても重要な役割を担うことは間違いない。今後の医学研究においては、「膨大ではあるがあくまで有限なヒト遺伝子プール」のなかから特定の特徴を備えた遺伝子を効率よく同定することが重要になる。ゲノミクス技術のなかでも DNA チップはすでに複数の疾患で有用な解析ツールとなることが示されてきた²⁾³⁾。

DNA チップはスライドガラスなどの小さな担体上に、解析する遺伝子由来の cDNA あるいはオリゴヌクレオチドを高密度に配置したものであり、いわば超高密度ドットプロット法といえよう⁴⁾。一枚の DNA チップ上には数千～数万種類の DNA 断片が配置されており、一回の実験でこれら数多くの遺伝子の発現変化を包括的に解析できる。

Key Word

DNA マイクロアレイ
急性骨髄性白血病
骨髄異形成症候群
遺伝子発現プロファイル

* MANO Hiroyuki/自治医科大学ゲノム機能研究部

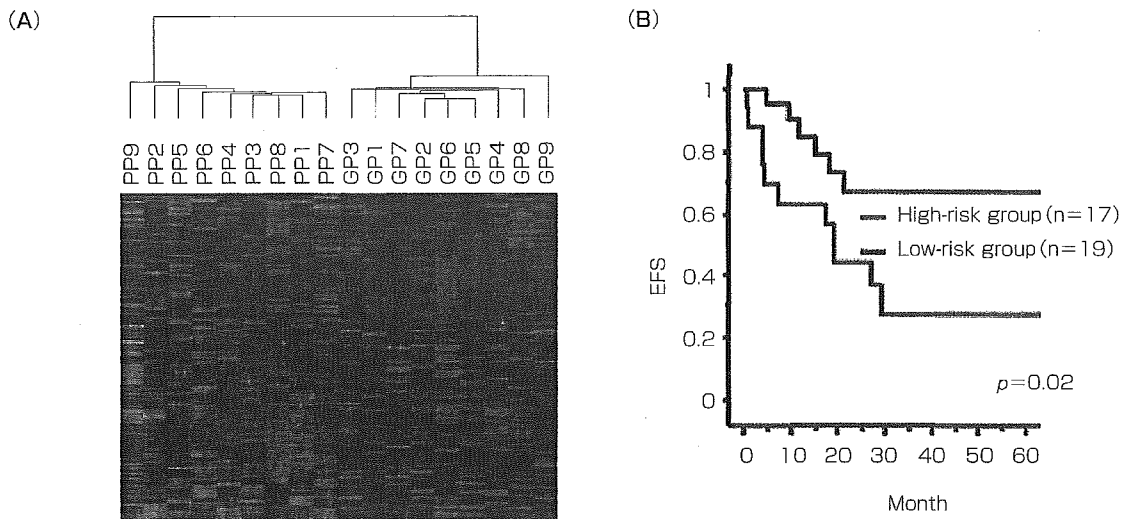


図1 骨髄単核球の発現プロファイルによる予後予測 (Yagi *et al.*, 2003²⁾ より改変引用)
 (A) 長期生存症例(GP)と予後不良症例(PP)間で発現が異なる遺伝子を取り出したところ 133 種類同定された。これら遺伝子の発現量で二次元クラスタリングをおこなった。(B) 上記遺伝子の発現量でテストサンプルを二次元クラスタリングし、大きく分かれた二群間で予後を比較した。

このように DNA チップを用いることで、これまでとは比較にならないような膨大な遺伝子発現データを解析することができるようになった。こうして得られた特定の遺伝子群の発現様式、発現パターンのことを「遺伝子発現プロファイル (gene expression profile)」とよぶ。このプロファイルをくらべることで疾患の特性・病因に迫れるのではないかと考えられる。

1. 急性骨髄性白血病(acute myeloid leukemia: AML)の予後予測

AMLは未熟な血液細胞ががん化して生じる疾患であり、これまでさまざまなパラメーターが治療予後にリンクすることが報告されてきた⁹⁾。現在臨床の場で白血病の診断および予後予測をおこなう場合、患者骨髄細胞に関して①ペルオキシダーゼ、エステラーゼなどの特殊染色、②FACSを用いた細胞表面マーカーの解析、③遺伝子異常の有無の解析(免疫グロブリン遺伝子やT細胞表面レセプター遺伝子の再構成、*BCR-ABL* 遺伝子、*PML-RAR α* 遺伝子などの有無)、④染色体型の解析、などの検査が利用される⁹⁾。これらを総合的に解析することで予後にある程度リンクした診断がおこなえるが、上記の解析をすべておこなうのは煩雑であり多くの専門

技術も必要となる。一方、例えば一枚のDNAチップを用いることで旧来の方法で得られる情報がすべて解析可能となれば、診断はより簡便になりかつ異なった施設間での診断の一致率も上昇するであろう。骨髄中の血球の多くは単核球とよばれる分画に相当し、これまでAML患者骨髄単核球を用いたDNAチップ実験により患者予後を予測する試みがおこなわれてきた。

Yagiら⁷⁾は小児AML患者54例についてAffymetrix社製GeneChip® Human Genome U95Aチップを用いた遺伝子発現解析をおこない、小児AML内での予後良好群と不良群とを予測する試みをおこなっている。彼らのサンプル内で3年以上完全寛解を維持している症例9例と初回緩解導入に失敗した症例9例の間で発現量が統計的に異なる遺伝子133種類を抽出し(図1A)、これら予後関連遺伝子の発現量をもととして、患者全体を二次元クラスタリング法あるいはsupported vector machine (SVM)法によって層別化した。その結果どちらの方法によってもAML内で生命予後が異なるサブグループを検出することに成功している(図1B)。すなわち遺伝子発現プロファイルをもとにした新しい疾患分類法の可能性が提示されたことになる。

同様に Bullingerら⁸⁾はAML患者116例の骨髄単核

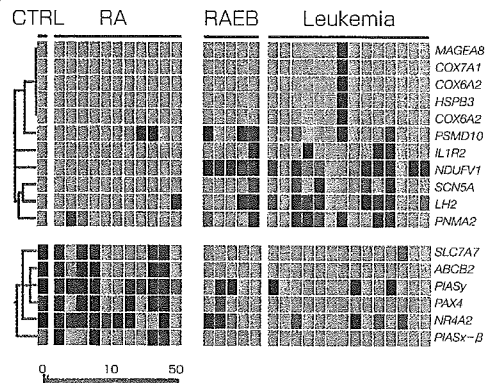


図2 MDS の病期進行に伴い発現量が変化する遺伝子 (Ueda M *et al.*, 2003¹¹⁾より改変引用)

「健常人+ RA」期と「RAEB + 白血病」期とを比較した結果、後者で発現誘導する遺伝子 11 種類(上図)、および低下する遺伝子 6 種類(下図)が同定された。それぞれの発現パターンを遺伝子系統樹(gene tree)の形で表す。遺伝子名は右に示してあり、また発現量は下にあるカラーバーに沿って色で示される。

球より mRNA を調整し、約 3 万 9 千種類の cDNA を配置した DNA チップで解析をおこなっている。彼らは患者を training set と test set に分け、前者のデータを用いて患者予後に発現量がリンクする遺伝子をスクリーニングした。その結果 100 種類以上の予後予測遺伝子セットを同定することに成功している。さらにこれら遺伝子の発現量をもとに患者を 2 群に分け、その予後と比較したところ、2 群の生命予後は有意に異なっていた ($p < 0.001$)。またあらかじめ取っておいた test set において検証したところ、やはり 2 群の予後は大きく違っている ($p = 0.006$) ことが示された。これらの予後分類法が現行の核型にもとづく分類法とは独立したものであることを示すために、彼らは核型分類では中間群に分類される「正常核型」をもった患者のみを選択し、その予後予測も試みている。彼らの分類法を適応したところ、正常核型のなかでも予後良好群と不良群とを区別することに成功した ($p = 0.046$)。彼らの解析は旧来の核型分類法とは異なった新しい予後予測法の可能性を示唆しており興味深い。

また Valk ら⁹⁾も同様の目的から 285 例の AML 患者骨髄単核球を用いて約 13,000 種類の遺伝子に関する発現量を測定した。彼らもやはり予後にリンクした遺伝子発現プロファイルがあることを検出している。

2. 骨髄異形成症候群(MDS)の病期進展機構

MDS については、その発症機構、再生不良性貧血との鑑別、病期進展機構、また MDS 由来白血病と特発性 AML との鑑別など、現在なお不明な点が数多く残されている。マイクロアレイによる解析はこれらの疑問に答えるうえで有用だと思われるが、しかしながら MDS の

ように骨髄中の悪性芽球が患者ごとに大きく異なる疾患の場合は偽陽性データを減らすよう工夫する必要がある。たとえば MDS 初期の不応性貧血(RA)と MDS 由来白血病とを比較する場合と考えると、前者患者骨髄中の悪性芽球は 5%以下であり、後者では少なくとも悪性芽球が 20%以上を占める。単純に両骨髄をアレイによって比較すると、未分化な血球に特異的な遺伝子は白血病患者骨髄中では必ず発現が誘導されているように結論づけられてしまう。しかしこれは、骨髄中での幼弱な白血病芽球の増加を反映した結果であり、ほとんどの場合実際の芽球一細胞あたりの mRNA 量に変化は認められない。

われわれは造血幹細胞分画に特異的に発現する膜蛋白 CD133(AC133)¹⁰⁾に着目し、CD133 に対するアフィニティカラムを用いて広く血液悪性疾患患者より造血幹細胞相当分画を純化保存するバンク事業「Blast Bank」を開始した。2005 年 9 月現在ですでに Blast Bank のサンプル数は 600 例を越えており、純化ヒト疾患細胞のゲノムクスプロジェクトとしては世界最大級のものである。

Ueda ら¹¹⁾は本バンクに属する MDS の各病期サンプルをマイクロアレイで比較することで、病期進展機構の解明を目指した。健常人骨髄 1 例、RA 期 11 例、RAEB 期 5 例、および MDS 由来白血病 15 例のバンクサンプルより mRNA 分画を抽出し、ヒト転写因子、膜蛋白質、およびシグナル伝達分子などの 2304 種類の遺伝子を配置した DNA マイクロアレイによる解析をおこなっている。その結果、複数の病期特異的遺伝子を同定することに成功している。一方、病初期から、RAEB/AML への移行メカニズムをこのデータセットから解析

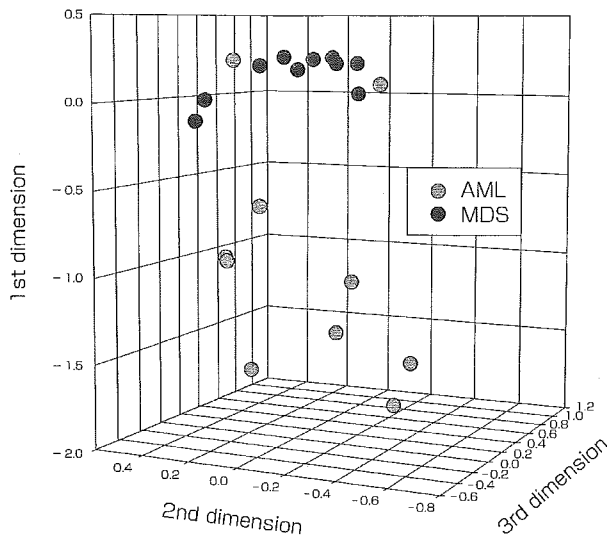


図4 MDS由来白血病と *de novo* AML サンプルの仮想空間への投射(筆者作成)

両疾患間で発現量が有意に異なる遺伝子を抽出し、これらの遺伝子発現パターンの主成分を3種類抽出した。得られたパターンを仮想軸とする三次元空間にサンプルを投射したところ、両疾患群が遺伝子発現パターンの上から、異なる疾患単位であることが示唆された。

することも可能である。たとえば全体を予後良好群である「正常+RA」期と予後不良群である「RAEB+AML」期に大きく分け、両者間で発現量が大きく異なる遺伝子を同定することも試みた。その結果、図2に示すように11種類の遺伝子が病期の進展に伴い発現量が増加することが明らかになり、一方6種類の遺伝子の発現量が減少することも示された。前者は病期進展に関与するがん遺伝子の、また後者はがん抑制遺伝子のよい候補といえる。

たとえば後者の遺伝子群に属するPIASyは当初転写因子STAT1の機能を抑制する細胞内分子としてクローニングされ、現在はユビキチン類似分子であるSUMO (small ubiquitin-like modifier)の結合酵素そのものであると考えられている。細胞内蛋白分解系専用の「蛋白タグ」であるユビキチンとは異なり、SUMOの細胞内機能は多彩であり、PIASyの機能もSTAT1, p53, アンドロゲンレセプターの活性調節など多岐にわたる。また、Uedaら¹⁴⁾はPIASyの発現低下がMDSにおける病期進行メカニズムに直接関与している可能性を検証する目的で、G-CSF依存性に顆粒球へと分化するマウス骨髓系細胞株32Dへ同遺伝子を導入した。その結果PIASyの発現誘導に伴って細胞の著明なアポトーシスが認められることがわかった。これらの結果はPIASyの「がん抑制遺伝子」としての機能を強く示唆しており、PIASyの発現低下がMDS病期進行の新たなメカニズムである可能

性が注目される。

3. MDSとAMLの鑑別診断

一方、鑑別診断マーカーのスクリーニングにもマイクロアレイは有用である。MDS由来のAMLは*de novo* AMLにくらべ治療抵抗性であることが多く、その鑑別診断は重要であるが、両者を明確に区別することはしばしば困難である。Miyazatoら¹²⁾はMDS由来AMLおよび*de novo* AML両群のBlast Bankサンプルそれぞれ5例を選び、1152遺伝子の発現量を比較した。その結果膜蛋白DlkがMDS由来群特異的に高発現していることを確認し、またその疾患特異的発現は症例を追加した定量的real-time RT-PCR法によっても確認された。

Oshimaら¹³⁾は、はたしてMDS由来AMLと*de novo* AMLが遺伝子発現の面から鑑別可能か否かを正確に検証する目的で、Blast Bankに属する両疾患群のうち、FABサブタイプでM2に分類されるサンプルのみ各10例を選び、12,000種類以上の遺伝子発現量をマイクロアレイで比較した。この解析により分化傾向(FABサブタイプ)もそろったCD133陽性細胞間での詳細な比較が可能になる。まずt検定により両群間で発現量が統計的に有意に異なる遺伝子を抽出し、さらに各群における発現量の平均値が群間で大きく異なる遺伝子を選び出した。その結果計73種類の遺伝子が最終的に同定

された。実際には、これら遺伝子の発現量のパターンは、73 遺伝子がそれぞれ独立しているわけではなく、よく似たものも多い。そこで correspondence analysis 法によって発現量の変化パターン 73 種類から代表的なもの 3 種類を抽出した(すなわち代表的な仮想発現パターンそれぞれ作る仮想遺伝子 3 種類を作成したことになる)。これら仮想遺伝子 3 種類を軸とした仮想空間に各サンプルを投射したのが図 4 である。その結果 MDS 由来 AML と *de novo* AML のサンプルの多くが仮想空間上で異なった位置に存在することがわかった。すなわち両疾患は異なった単位と考えるとよいことを示唆している。一方図 3 において、一部の症例が遺伝子発現パターンと臨床診断が合致していないことも明らかである。これらの結果は遺伝子発現プロファイルにもとづく新しい診断法が可能であることを示唆しているが、その正当性は長期的な患者生命予後とのリンクが証明されてはじめて明らかになるであろう。

おわりに

DNA マイクロアレイがその新規性のみで注目・検討された時期は終わりつつあり、比較的高額な実験費用に見合うだけの有用性が本当にあるのかが問われる時代になった。ヒトゲノムプロジェクトの成果は、医学にとってきわめて貴重(かつ膨大)な情報リソースであり、この有効活用なくして 21 世紀の医療は成り立たないであろう。必ずしも DNA マイクロアレイが理想のシステムとはいえないが、現段階でヒトゲノム情報の網羅的利用に関して DNA マイクロアレイを超える技術は存在しない。実際の実験にあたっては「網羅的発現比較によって何を明らかにしたいのか？」を明確に念頭に置き、そのための最も効率よいアプローチを検討することが重要であろう。DNA マイクロアレイシステムは慎重にデザインされた実験においてのみ、そのきわめて高いスクリーニング能力を発揮すると考えられる。

文 献

- 1) International Human Genome Sequencing Consortium : Finishing the euchromatic sequence of the human genome. *Nature* 431 : 931-945, 2004
- 2) Dhanasekaran SM *et al* : Delineation of prognostic biomarkers in prostate cancer. *Nature* 412 : 822-826, 2001
- 3) van't Veer LJ *et al* : Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415 : 530-536, 2002
- 4) Duggan DJ *et al* : Expression profiling using cDNA microarrays. *Nat Genet* 21 : 10-14, 1999
- 5) Grimwade D *et al* : The importance of diagnostic cytogenetics on outcome in AML : analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* 92 : 2322-2333, 1998
- 6) Bennett JM *et al* : Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann Intern Med* 103 : 620-625, 1985
- 7) Yagi T *et al* : Identification of a gene expression signature associated with pediatric AML prognosis. *Blood* 102 : 1849-1856, 2003
- 8) Bullinger L *et al* : Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 350 : 1605-1616, 2004
- 9) Valk PJ *et al* : Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* 350 : 1617-1628, 2004
- 10) Hin AH *et al* : AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* 90 : 5002-5012, 1997
- 11) Ueda M *et al* : DNA microarray analysis of stage progression mechanism in myelodysplastic syndrome. *Br J Haematol* 123 : 288-296, 2003
- 12) Miyazato A *et al* : Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. *Blood* 98 : 422-427, 2001
- 13) Oshima Y *et al* : DNA microarray analysis of hematopoietic stem cell-like fractions from individuals with the M2 subtype of acute myeloid leukemia. *Leukemia* 17 : 1990-1997, 2003

MANO Hiroyuki

まの・ひろゆき
 1959 年、岡山県生まれ
 1984 年、東京大学医学部医学科卒
 1986 年、東京大学医学部第三内科入局
 1989 年、米国テネシー州 St. Jude 小児研究病院生化学部門 留学
 1991 年、東京大学医学部第三内科文部教官 助手
 1993 年、自治医科大学分子生物学講座 講師
 2001 年より自治医科大学ゲノム機能研究部 教授(現職)
 専門は血液内科、ゲノム医学
 研究テーマはゲノミクス技術を用いた造血器悪性腫瘍の病態解明

特集●個別化医療の実現を目指して

遺伝子発現プロファイリングによる 急性骨髄性白血病の予後予測

間野 博行

遺伝子発現プロファイリングによる急性骨髄性白血病の予後予測

間野博行

自治医科大学ゲノム機能研究部

DNA マイクロアレイを用いることで、数万種類の遺伝子の発現量を任意の細胞・組織間で簡便に比べることが可能になった。現在、急性骨髄性白血病患者サンプルをマイクロアレイで解析し、得られた遺伝子発現プロファイルのなかから患者の長期予後にリンクするものをスクリーニングする試みが精力的に行われている。また最近では、患者骨髄の単核球全体を解析するのではなく、きわめて幼弱な造血幹細胞分画のみをあらかじめ純化した後にマイクロアレイ解析を行うことも試みられている。同定された予後関連遺伝子群の発現量をもとに各患者の予後を予測することがこれらのプロジェクトの目標であり、今後は遺伝子発現プロファイルに基づいた急性白血病の新しい分類法が提案されると期待される。

Key words DNA microarray, genomics, AC133, karyotype

遺伝子発現プロファイル

約 30 億塩基対に及ぶヒトゲノムの核酸配列を決定する大規模事業である「ヒトゲノムプロジェクト」がついに 2003 年 4 月に終了宣言を行い、ヒト染色体のユークロマチン (euchromatin) 領域のほぼ完全な塩基配列が決定された (<http://www.ncbi.nlm.nih.gov/genome/seq/>)。ヒトのもつタンパク質をコードする総遺伝子数はおそらく 20,000~25,000 種類前後になると予想されている¹⁾。かつて夢であったヒトの全遺伝子プールの全貌がいよいよ明らかになるうとしており、ポストゲノム時代が訪れたといえる。

この膨大な遺伝子情報を有効に使う技術である

ゲノミクス、プロテオミクスが、白血病の解析においても重要な役割を担うことは間違いないであろう。今後の医学研究においては、「膨大ではあるがあくまで有限なヒト遺伝子プール」のなかから、特定の特徴を備えた遺伝子を効率よく同定することが重要になる。ゲノミクス技術のなかでも DNA マイクロアレイ、DNA チップ (以下 DNA マイクロアレイ) はすでに複数の疾患で有用な解析ツールとなることが示されてきた^{2,3)}。

DNA マイクロアレイはスライドガラスなどの小さな担体上に、解析する遺伝子由来の cDNA あるいはオリゴヌクレオチドを高密度に配置したものであり、いわば超高密度ドットプロット法といえよう⁴⁾。たとえば今、組織 A 由来の cDNA を蛍光色素 Cy3 で標識し、組織 B 由来の cDNA を蛍

Prediction of prognosis for AML based on gene expression profile

Hiroyuki Mano

まの・ひろゆき 1984 年東京大学医学部卒業、86 年同第三内科入局、医学博士、91 年同助手、93 年自治医科大学医学部分子生物学講座講師、2001 年同大学医学部ゲノム機能研究部教授 (現在に至る)。研究テーマ: ゲノミクス技術を用いた造血器悪性腫瘍の病態解明。

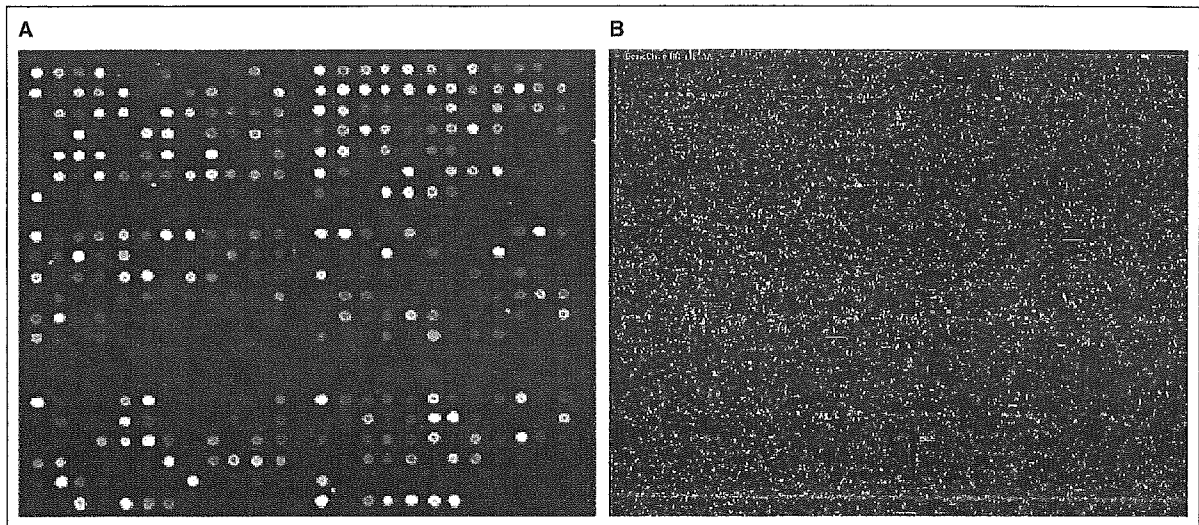


図1 DNA マイクロアレイのスキャン画像

cDNA タイプマイクロアレイ (A) および Affymetrix 社の GeneChip (B) をスキャンした画像をそれぞれ示す。

A: cDNA アレイの場合は2種類のサンプルから調製したcDNAをそれぞれCy3とCy5の蛍光色素で標識し、混合してハイブリダイズした。Cy3標識cDNAが主に結合したスポットは緑色、Cy5標識cDNAが主に結合したスポットは赤色、両者がほぼ等量結合したスポットは黄色で表される。

B: 単一のサンプルより調製した標識cRNAをハイブリダイズさせてあり、実際は単色のグレースケールで表されるが、発現強度をわかりやすくするためにpseudo-colorで表してある。

光色素Cy5で標識したとする。両者の混合物をcDNAを高密度に配置したマイクロアレイとハイブリダイズさせると、標識cDNAは各スポット上の遺伝子の組織AとBにおける発現量の比に応じた形でスポットに結合する(図1A)。レーザーで励起するとCy3は緑色、Cy5は赤色の蛍光を発するため、あるスポット上の遺伝子が組織Aのみに発現していれば緑色のスポットとなり、Bのみに発現していれば赤色のスポットとなる。また両組織においてほぼ等量に発現している遺伝子の場合には黄色のスポットとして見える。このようにして両色素の蛍光強度を測定することで、それぞれのスポット上の遺伝子の測定サンプル間における発現比が定量できるのである。1枚のDNAマイクロアレイ上には数千~数万種類のDNA断片が配置されており、1回の実験でこれら数多くの遺伝子の発現変化を包括的に解析できる。

なお、担体上にオリゴヌクレオチドを配置したオリゴヌクレオチドアレイの場合には、上述のようなCy3とCy5の二重染色による2検体間発現比

較ではなく、1種類の検体を1種類の蛍光染色で解析することのほうが多い。したがって、1枚のアレイからは単一の組織の発現プロファイルが得られるのみであり、他の組織と比べる場合は、それぞれ異なったアレイで得られたデータをなんらかの方法でノーマライズさせた後にコンピュータ上で比較することになる(図1B)。また、オリゴヌクレオチドアレイの場合にはスポットの蛍光強度を増強するために、mRNAからいったん二本鎖cDNAを作製し、さらにT7RNAポリメラーゼなどを用いてRNAを増幅するとともに蛍光色素を取り込ませることが多い。

このようにマイクロアレイを用いることで、これまでとは比較にならないような膨大な遺伝子発現データを解析することができるようになった。こうして得られた遺伝子群の発現様式、発現パターンのことを「遺伝子発現プロファイル (gene expression profile)」と呼ぶ。このプロファイルを比べることで白血病の特性に迫れるのではないかと考えられる。

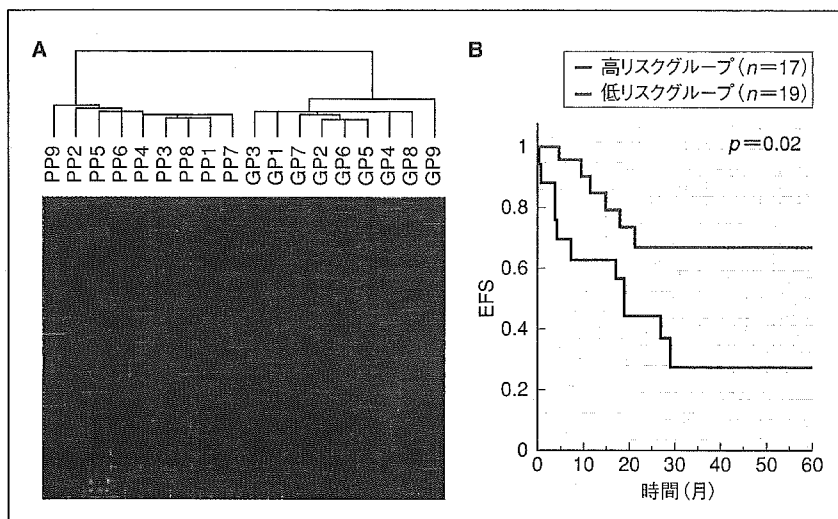


図2 骨髄単核球の発現プロファイルによる予後予測

A: 長期生存症例 (GP) と予後不良症例 (PP) 間で発現が異なる遺伝子を取り出したところ 133 種類同定された。これら遺伝子の発現量で 2 次元クラスタリングを行った。

B: 上記遺伝子の発現量でテストサンプルを 2 次元クラスタリングし、大きく分かれた 2 群間で予後を比較した。EFS: event-free survival

(Yagi T, et al. Blood 2003 ; 102 : 1849-56⁷⁾より)

DNA マイクロアレイを用いた急性骨髄性白血病の新しい予後予測

急性骨髄性白血病 (AML) は未熟な血液細胞が癌化して生じる疾患であり、これまでさまざまなパラメータが治療予後にリンクすることが報告されてきた⁵⁾。現在、臨床の場で白血病の診断および予後予測を行う場合、患者骨髄細胞に関して①ペルオキシダーゼ、エステラーゼなどの特殊染色、②FACSを用いた細胞表面マーカーの解析、③遺伝子異常の有無の解析 (免疫グロブリン遺伝子やT細胞表面レセプター遺伝子の再構成、*BCR-ABL* 遺伝子、*PML-RAR α* 遺伝子などの有無)、④染色体型の解析、などの検査が利用される⁶⁾。これらを総合的に解析することで予後にある程度リンクした診断が行えるが、上記の解析をすべて行うのは煩雑であり多くの専門技術も必要となる。一方、たとえば1枚のDNAマイクロアレイを用いることで旧来の方法で得られる情報がすべて解析可能となれば、診断はより簡便になり、かつ異なった施設間での診断の一致率も上昇するであろう。骨髄中の血球の多くは単核球 (mononuclear cell) と呼ばれる分画に相当し、これまでAML患者骨髄単核球を用いたDNAマイクロアレイ実験により患者予後を予測する試み

が行われてきた。

Yagiらは、小児AML患者54例についてAffymetrix社GeneChip HGU95Aチップを用いた遺伝子発現解析を行い、小児AML内での予後良好群と不良群とを予測する試みを行っている⁷⁾。彼らのサンプル内で3年以上完全寛解 (complete remission ; CR) を維持している症例9例と初回寛解導入に失敗した症例9例の間で発現量が統計的に異なる遺伝子133種類を抽出し (図2A)、これら予後関連遺伝子の発現量をもとにして、患者全体を2次元クラスタリング法、あるいはsupported vector machine (SVM) 法によって層別化した。その結果、どちらの方法によってもAML内で生命予後が異なるサブグループを検出することに成功している (図2B)。すなわち、遺伝子発現プロファイルをもとにした新しい疾患分類法の可能性が提示されたことになる。

同様にBullingerらは、AML患者116例の骨髄単核球よりmRNAを調製し、約39,000種類のcDNAを配置したマイクロアレイで解析を行っている⁸⁾。彼らは患者をtraining setとtest setに分け、前者のデータを用いて患者予後に発現量がリンクする遺伝子をスクリーニングした。その結果、100種類以上の予後予測遺伝子セットを同定することに成功している。さらに、これら遺伝子の発

現量をもとに患者を2群に分け、その予後を比較したところ、2群の生命予後は有意に異なっていた ($p < 0.001$)。また、あらかじめ取っておいた test set において検証したところ、やはり2群の予後は大きく違っている ($p = 0.006$) ことが示された。これらの予後分類法が現行の核型に基づく分類法とは独立したものであることを示すために、彼らは核型分類では中間群に分類される「正常核型」をもった患者のみを選択し、その予後予測も試みている。彼らの分類法を適用したところ、正常核型のなかでも予後良好群と不良群とを区別することに成功した ($p = 0.046$)。彼らの解析は旧来の核型分類法とは異なった新しい予後予測法の可能性を示唆しており興味深い。

また Valk らも同様の目的から、285 例の AML 患者骨髄単核球を用いて約 13,000 種類の遺伝子に関する発現量を測定した。彼らもやはり予後にリンクした遺伝子発現プロファイルがあることを検出している⁹⁾。

純化細胞を用いた 急性骨髄性白血病の予後予測

しかし、実際に DNA チップを用いた実験を行うと、本法が意外にも偽陽性データが多く効率の悪い手技であることにも気付く。その理由を、たとえば健康人と白血病患者の骨髄を比べる場合で考えてみよう (図 3)。正常骨髄はさまざまな系統の細胞群から成る「ヘテロ」な集団である。もし今、この骨髄中の多能性幹細胞の 1 クローンが悪性転化して白血病になったと仮定すると、白血病患者の骨髄は幹細胞由来の白血病細胞が多くを占めるようになり、それ以外の細胞の骨髄における割合は相対的に低下する。したがって、両者の骨髄細胞 (あるいは単核球) 全体を単純に比較するような DNA チップ解析を行うと、造血幹細胞分画以外の細胞 (たとえば図 3 の青色細胞) 特異的に発現する遺伝子などはすべて減少して見えてしまう。この変化は、それぞれの遺伝子の細胞内における mRNA のコピー数を反映しておらず、

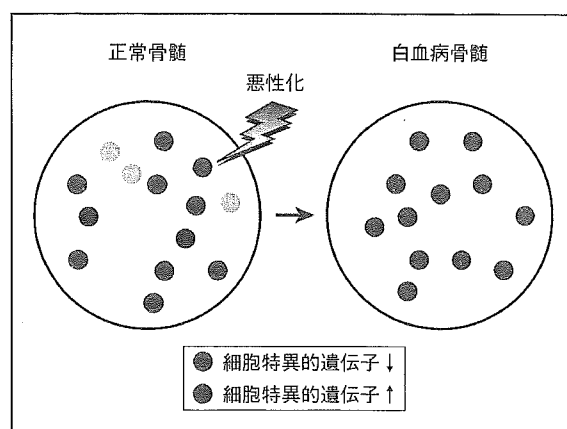


図 3 「Population shift 効果」による偽陽性データ

正常骨髄はさまざまな細胞から成るヘテロな集団である。今そのなかから緑色の細胞が悪性化し白血病が生じたとすると、患者骨髄内は結果的に緑色由来の細胞が大半を占めることになり、それ以外の細胞の割合は相対的に減少する。したがって、骨髄全体を正常と白血病患者で比べるような単純なマイクロアレイ解析を行えば、青色細胞特異的な遺伝子は必ず減少しているように見え、逆に緑色細胞特異的な遺伝子の発現は必ず増加しているように見えてしまう。このような変化は実際の細胞内での発現変化を反映しているわけではなく、単に骨髄内でのポピュレーションの変化を見ているに過ぎない。この「population shift 効果」による見せかけの発現変化が、大量の偽陽性データを生み出すもととなる。

誤って「発現減少」と判断されたわけである。一方、幹細胞特異的に発現していた遺伝子は、細胞あたりの発現量の変化の有無にかかわらず骨髄中における割合が増加するために、誤って「発現誘導」と判断されることになる¹⁰⁾。これらのさまざまな偽陽性や偽陰性データのノイズのなかから真に「癌化に伴って細胞あたりの発現量が変化する」遺伝子を選択するのはきわめて困難であろう。

では、どのようにすればこれら偽陽性データを回避する効率のよいゲノム解析が可能となるのであろうか。白血病を含む多くの血液悪性疾患が、造血幹細胞あるいはそのごく近傍の幼若細胞における悪性転化に起因することが知られる。そこで、患者骨髄より造血幹細胞相当分画のみを純化しこれを用いて DNA チップ解析を行えば、患者骨髄中の芽球の割合や芽球の分化傾向の違いに影響されない精度の高い解析が可能になると期待される。筆者らは、造血幹細胞分画に特異的に発現する膜

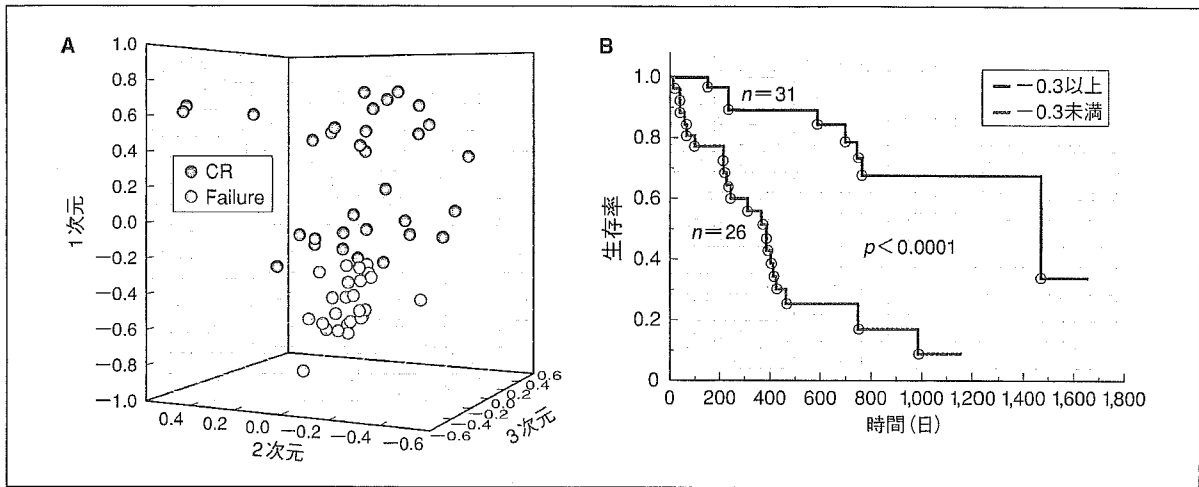


図4 AC133陽性細胞を用いた予後予測

A: 化学療法による初回寛解導入が成功した群(CR)と失敗した群(Failure)との間で有意に発現量が異なる遺伝子をスクリーニングし、その発現プロファイルから代表的なものを3種類抽出した。これらプロファイル上の仮想発現量に基づいてサンプルを空間に投射したところ、CR群とFailure群とは異なった位置に配置された。
 B: 仮想空間のZ軸の値が-0.3未満(灰色)か以上(赤色)かによって患者を二分し、その予後をKaplan-Meier法で比較した。

タンパク質 AC133¹¹⁾に着目し、AC133に対するアフィニティカラムを用いて広く血液悪性疾患患者より造血幹細胞相当分画を純化・保存するバンク事業「Blast Bank」を開始した。平成16年12月現在ですでにBlast Bankのサンプル数は600例を超えており、純化細胞のゲノミクスプロジェクトとしては世界最大級といえる¹²⁻¹⁴⁾。

では具体的に、Blast Bankを用いて白血病の予後予測が可能か否か検討してみよう。筆者らは、ヒト全遺伝子が配置されたDNAマイクロアレイを用いて標準的化学療法を受けたAML患者66例の骨髓AC133陽性細胞について遺伝子発現プロファイルデータを得た。これら66例の患者中初回化学療法によってCRに到達した症例は51例あり、残り15例は寛解導入に失敗した。後者の15例は1年以内に全員死亡しており、きわめて予後不良なグループといえる。この初回化学療法に成功した患者と失敗した患者の間では遺伝子発現プロファイルは違うのであろうか。それを検討するために44,000種類のプローブセット(33,000種類の遺伝子に相当)の発現データ中、両グループ間で発現が有意に異なる遺伝子を選び

出し、さらに少なくともどちらかのグループでは高い発現量を示す遺伝子スクリーニングしたところ、最終的に約30種類の遺伝子が得られた。得られた遺伝子の発現プロファイルはそれぞれが互いに独立なわけではなく、よく似たものも多い。そこでcorrespondence analysis法¹⁵⁾によって、これら30種類の発現量パターンから全体を最も代表する仮想発現プロファイルを3種類抽出した(すなわち、代表的な仮想発現パターンそれぞれを作る仮想遺伝子3種類を作製したことになる)。これら仮想遺伝子3種類を軸とした仮想空間に各サンプルを投射したのが図4Aである。その結果興味深いことに、治療奏効症例と失敗症例とは空間上不連続な位置に独立して存在することがわかった。言い換えると、両グループは遺伝子発現プロファイルの面では異なった集団に属することになる。

このような空間上の位置は患者予後にリンクしているのであろうか。図4Aをよく見てみると、患者サンプルはZ軸上で最も強く分離されていることがわかる。そこで、症例の仮想空間上でのZ軸上の値が-0.3未満か-0.3以上かによって大

きく2群に分け、その長期生命予後をKaplan-Meier法によって解析してみた。図4Bで明らかのように、両群は大きく予後が異なることがわかる。しかも、さまざまな方法で予後関連遺伝子を絞り込んだ結果、数種類の遺伝子の発現量を測定するだけで患者の長期予後を予測可能なことまで明らかになった(Yamashitaら、未発表データ)。これらの予後関連遺伝子数は、以前の骨髓単核球による解析で同定された100種類以上の予後関連遺伝子数に比べて明らかに少ない。純化した細胞を標的として解析することで分化傾向などによるノイズが低減し、少数の遺伝子で予後予測可能になったのではないかと考えられる。

DNAマイクロアレイがその新規性のみで注目・検討された時期は終わりつつあり、比較的高額な

実験費用に見合うだけの有用性が本当にあるのかが問われる時代になった。ヒトゲノムプロジェクトの成果は、医学にとってきわめて貴重(かつ膨大)な情報リソースであり、この有効活用なくして21世紀の医療は成り立たないであろう。必ずしもDNAマイクロアレイが理想のシステムとはいえないが、現段階でヒトゲノム情報の網羅的利用に関してDNAマイクロアレイを超える技術は存在しない。実際の実験にあたっては「網羅的発現比較によって何を明らかにしたいのか」を明確に念頭に置き、そのための最も効率よいアプローチを検討することが重要であろう。DNAマイクロアレイシステムは慎重にデザインされた実験においてのみ、そのきわめて高いスクリーニング能力を発揮すると考えられる。

文献

- 1) International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004 ; 431 : 931-45.
- 2) Dhanasekaran SM, Barrette TR, Ghosh D, et al. Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001 ; 412 : 822-6.
- 3) van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002 ; 415 : 530-6.
- 4) Duggan DJ, Bittner M, Chen Y, et al. Expression profiling using cDNA microarrays. *Nat Genet* 1999 ; 21 : 10-4.
- 5) Grimwade D, Walker H, Oliver F, et al. The importance of diagnostic cytogenetics on outcome in AML : Analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* 1998 ; 92 : 2322-33.
- 6) Bennett JM, Catovsky D, Daniel MT, et al. Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann Intern Med* 1985 ; 103 : 620-5.
- 7) Yagi T, Morimoto A, Eguchi M, et al. Identification of a gene expression signature associated with pediatric AML prognosis. *Blood* 2003 ; 102 : 1849-56.
- 8) Bullinger L, Dohner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 2004 ; 350 : 1605-16.
- 9) Valk PJ, Verhaak RG, Beijen MA, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* 2004 ; 350 : 1617-28.
- 10) Miyazato A, Ueno S, Ohmine K, et al. Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. *Blood* 2001 ; 98 : 422-7.
- 11) Hin AH, Miraglia S, Zanjani ED, et al. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* 1997 ; 90 : 5002-12.
- 12) Tsutsumi C, Ueda M, Miyazaki Y, et al. DNA microarray analysis of dysplastic morphology associated with acute myeloid leukemia. *Exp Hematol* 2004 ; 32 : 828-35.
- 13) Ota J, Yamashita Y, Okawa K, et al. Proteomic analysis of hematopoietic stem cell-like fractions in leukemic disorders. *Oncogene* 2003 ; 22 : 5720-8.
- 14) Oshima Y, Ueda M, Yamashita Y, et al. DNA microarray analysis of hematopoietic stem cell-like fractions from individuals with the M2 subtype of acute myeloid leukemia. *Leukemia* 2003 ; 17 : 1990-7.
- 15) Fellenberg K, Hauser NC, Brors B, et al. Correspondence analysis applied to microarray data. *Proc Natl Acad Sci USA* 2001 ; 98 : 10781-6.

Scalable Generation of High-Titer Recombinant Adeno-Associated Virus Type 5 in Insect Cells

Masashi Urabe,^{1*} Takayo Nakakura,¹ Ke-Qin Xin,² Yoko Obara,¹ Hiroaki Mizukami,¹
Akihiro Kume,¹ Robert M. Kotin,³ and Keiya Ozawa¹

Division of Genetic Therapeutics, Jichi Medical School, Tochigi 329-0498, Japan¹; Department of Molecular Biodefense Research, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan²; and Laboratory of Biochemical Genetics, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland³

Received 14 June 2005/Accepted 27 November 2005

We established a method for production of recombinant adeno-associated virus type 5 (rAAV5) in insect cells by use of baculovirus expression vectors. One baculovirus harbors a transgene between the inverted terminal repeat sequences of type 5, and the second expresses Rep78 and Rep52. Interestingly, the replacement of type 5 Rep52 with type 1 Rep52 generated four times more rAAV5 particles. We replaced the N-terminal portion of type 5 VP1 with the equivalent portion of type 2 to generate infectious AAV5 particles. The rAAV5 with the modified VP1 required α 2-3 sialic acid for transduction, as revealed by a competition experiment with an analog of α 2-3 sialic acid. rAAV5-GFP/Neo with a 4.4-kb vector genome produced in HEK293 cells or Sf9 cells transduced COS cells with similar efficiencies. Surprisingly, Sf9-produced humanized *Renilla* green fluorescent protein (hGFP) vector with a 2.4-kb vector genome induced stronger GFP expression than the 293-produced one. Transduction of murine skeletal muscles with Sf9-generated rAAV5 with a 3.4-kb vector genome carrying a human secreted alkaline phosphatase (SEAP) expression cassette induced levels of SEAP more than 30 times higher than those for 293-produced vector 1 week after injection. Analysis of virion DNA revealed that in addition to a 2.4- or 3.4-kb single-stranded vector genome, Sf9-rAAV5 had more-abundant forms of approximately 4.7 kb, which appeared to correspond to the monomer duplex form of hGFP vector or truncated monomer duplex SEAP vector DNA. These results indicated that rAAV5 can be generated in insect cells, although the difference in incorporated virion DNA may induce different expression patterns of the transgene.

Recombinant adeno-associated virus (rAAV) is being developed as a gene transfer vector. rAAV based on serotype 2 (rAAV2) successfully transduces nondividing cells, including muscle, liver, and brain cells (29). Conventional rAAV production requires packaging of rAAV DNA into type 2 capsids by transient transfection of HEK293 cells with two or three plasmids: an AAV helper plasmid encoding *rep* and *cap* genes devoid of inverted terminal repeat (ITR) sequences, a vector plasmid harboring the therapeutic gene between ITRs, and an adenovirus helper plasmid expressing E2A, virus-associated (VA) RNA, and E4orf6. Transient cotransfection is the major limitation for scale-up of rAAV production. Since rAAV can be purified using column chromatography, which can result in highly purified rAAV while eliminating other contaminating viruses, some efforts were made to develop rAAV production systems by using recombinant mammalian viruses such as adenovirus (10) or herpes virus (4) which do not rely on the plasmid transfection and therefore may be amenable to scale-up production.

Recombinant baculoviruses based on the *Autographa californica* nuclear polyhedrosis virus are widely employed for production of heterologous proteins in cultured insect cells. The highly active, late *A. californica* nuclear polyhedrosis virus promoters, such as polyhedrin and p10 promoters, regulate the expression of heterologous proteins, resulting in large amounts

of foreign proteins. Insect cells may be grown in suspension cultures in volumes ranging from shake flasks of sizes from, e.g., 50 to 400 ml, up to commercial-size bioreactors, e.g., 1,000 liters and larger. Recently, we described a highly scalable and efficient method for packaging rAAV2 in insect cells by use of baculovirus expression vectors (31). The ease of scale-up production is perhaps the most attractive feature of this production system. Infection of insect cells in suspension culture with recombinant baculoviruses eliminates the transfection process. Standard downstream processing to recover rAAV, such as tangential flow filtration and column chromatography, is readily applied.

In addition to vectors derived from serotype 2, other serotypes, utilizing different cell surface receptors, constitute a vector set from which an appropriate vector can be selected for a specific application. AAV5 is the most divergent dependo-virus characterized (2), and type 5 AAV vectors have desirable properties that differ from other serotype vectors. AAV5 utilizes different receptors from other serotypes (14, 30), and rAAV5 has demonstrated different tropism from AAV2 (5), thus making it worthwhile to establish a method to produce rAAV5 in insect cells.

AAV is a member of the family *Parvoviridae*. The genome of AAV is a linear, single-stranded DNA of 4.7 kb in length. The ITRs flank the unique coding sequences for the nonstructural replication initiator proteins, Rep, and the structural capsid proteins, VP. The ITRs serve as origins of DNA replication and may also function as the packaging signal. Type 2 Rep78 is generated by the p5 promoter, while Rep68 is translated from spliced mRNA from the p5 promoter. The small Rep polypep-

* Corresponding author. Mailing address: Division of Genetic Therapeutics, Jichi Medical School, 3311-1 Yakushiji, Minami-kawachi, Tochigi 329-0498, Japan. Phone: 81-285-58-7402. Fax: 81-285-44-8675. E-mail: murabe@jichi.ac.jp.

types Rep52 and Rep40 are expressed by the p19 promoter with nonspliced or spliced mRNA. The p40 promoter regulates expression of capsid proteins VP1, VP2, and VP3. Alternate usage of two splice sites and translation of VP2 at a non-AUG codon results in a stoichiometry of 1:1:10 of VP1, VP2, and VP3. Both p5 proteins Rep78 and Rep68 are AAV origin binding proteins, and the presence of either is required for AAV DNA replication and processing replicative intermediates of the virus DNA (13). Also, either Rep52 or Rep40 is necessary for packaging the single-stranded, linear virion genome into preformed empty capsids (17). The transcriptional map of type 5 AAV differs from that of type 2; the p7 promoter or p19 promoter transcribes mRNA for Rep78 or Rep52. Type 5 AAV does not encode the spliced form of Rep polypeptides Rep68 and Rep40 (25). Structural protein VP1 is a minor constituent in the AAV capsid. But the VP1-unique portion of approximately 140 amino acid residues is highly conserved among different serotypes and has a phospholipase A₂ motif. The YXGGX and HDXXY motifs (where X is any amino acid residue) in phospholipase A₂ indicate the catalytic site and Ca²⁺ binding loop, respectively (see Fig. 3A). Enzymes classified into the secretory phospholipase A₂ family hydrolyze the ester bond at the 2-acyl ester position of glycerophospholipids in the presence of Ca²⁺ and are involved in many aspects of cellular pathways, such as lipid membrane metabolism and signal transduction pathways (1, 21). The VP1-unique portion of parvovirus is required for transfer of the virus from late endosomes to the nucleus (36). A mutant virus lacking the VP1-unique portion or the activity of phospholipase is not processed properly, and thus no virus or vector genes are expressed.

In the present study, we describe a rAAV5 production system based on recombinant baculovirus and insect cells. In order to achieve high production levels of rAAV5 particles, we replaced a portion of the VP1 polypeptide with the corresponding portion of type 2. The VP1 substitution did not alter the tropism of rAAV5, which behaved indistinguishably from rAAV5 with wild-type VP1. In an attempt to improve the yields of rAAV5 particles, we used type 1 Rep52 instead of type 5, which resulted in the production of more than 5×10^4 vector genomes (vg) per insect cell.

MATERIALS AND METHODS

Plasmid construction. A flow chart of plasmid construction is shown in Fig. 1. pSR485 is an AAV5 vector plasmid harboring green fluorescent protein (GFP) and neomycin (Neo) genes between the ITRs (27). NotI sites were introduced outside the GFP/Neo expression cassette by PCR amplification using primers 5'-GATCGTCGACGCGGCCGCTCTCAGTACAATCTGCTCTGATGCC and 5'-AGTCGTGACGCGGCCGCTGCAGGCATGCAAGCTGTGAAAAAATGC. The NotI sites (underlined) were introduced. The resulting 4-kb DNA fragment was inserted into the BglII-SalI (blunt) sites of pSR485 (pSR485 α). pFB5GFP was constructed by insertion of the 4.8-kb PagI fragment from pSR485 into the Eco105III site of pFBHTA, which was derived from pFBHTb (Invitrogen, Carlsbad, CA) after removal of the polyhedrin promoter with BstZ17I and HindIII digestion. A humanized *Renilla* GFP (hGFP) gene was excised from phrGFPII-1 (Stratagene, La Jolla, CA) by treatment with BamHI and EcoRV and subcloned into an expression plasmid regulated by the cytomegalovirus (CMV) immediate-early promoter (pCMV). The resulting plasmid, pCMVhGFP, was treated with NotI to cut out the entire hGFP expression cassette, which was inserted into the corresponding site of pSR485 α or pFB5GFP (pSR485hGFP or pFB5hGFP, respectively). A human secreted alkaline phosphatase (SEAP) gene was excised from pSEAP2-Basic (Clontech, Mountain View, CA) with NruI and SalI, and the resulting 1.8-kb fragment was blunt-

ended and inserted into pCMV. The entire SEAP cassette was then excised with NotI and inserted into the corresponding site of pAAVGFP or pFBGFP (31) between the type 2 ITRs (pAAVSEAP or pFBSEAP, respectively). The type 5 p5 Rep open reading frame (ORF) equivalent to type 2 Rep78 was PCR amplified from pAAV5-2 (2) by using primers 5'-GAAGAAGCGCGCTATGAGTTCTCGCGAGACTTC and 5'-CGATTACTGTTCTTTATTGGCATCGTCAAATC and inserted into a cloning vector. The Rep ORF was cut out by NruI and BssHII, blunt-ended, and subcloned into the NotI site (blunt) of pBACAIERep (31), which was then treated with BglII and ClaI and blunt-ended, and the resulting 2.1-kb fragment was inserted into the NotI-PstI (blunt) sites of pFBD Δ (pFBD5LR). pFBD Δ is a derivative of FastBac Dual (Invitrogen) generated by the removal of the polyhedrin and p10 promoters with NcoI and BamHI treatment. The small Rep ORF was cut out from pFBD5LR by partial digestion with Eco47III and SalI, and the resulting 1.3-kb fragment was blunt-ended and inserted into the StuI site of pFastBac Dual (pFBD5SR). pFBD5SR was then digested with BstZ17I and SalI and treated with T4 DNA polymerase, and the resulting 1.4-kb fragment was inserted into the KpnI site (blunt) of pFBD5LR (pFBD5LSR). To generate the truncated p10 promoter, complementary 5'-phosphorylated oligonucleotides 5'-TAAAATCGCGAC and 5'-CATGGTCGC GATTTAAT were annealed to each other and inserted into the PacI-NcoI sites of pFastBac Dual (p Δ 5FBD). The type 5 Rep78 gene was PCR amplified with primers 5'-GCGCTTAATTAATAATCGCTAGTATGGCTACCTTCTATGAGTCATT-3' and 5'-GATCGCTAGCTTACTGTTCTTTATTGGCATCGTCA-3' and subsequently digested with PacI and NheI and inserted into the PacI-NheI sites of p Δ 5FBD (pFBD5LR12) (the Rep78 ORF is capitalized). The type 5 Rep52 gene amplified using primers 5'-GATCGCGCCCATGGCGCTCGTCAACTGGCTCGTGGAG-3' and 5'-GATCGCTGACTTACTGTTCTTTATTGGCATCGTCA-3' was digested with BssHIII and SalI and inserted into the corresponding sites of pFBD5LR12 (pFBD5LSR12 α). To replace type 5 Rep52 on pFBD5LSR12 α with type 1, 2, 3, or 4 Rep52, PCR was conducted with sense primer 5'-gatcccATGGAGCTGGTGGTGGTGGTGGTGGTGA-3' and antisense primer 5'-gatacctagTATTGCTCAGAAACAGTCATCCA-3' (for type 1 or 3) or 5'-gatacctagTATTGTTCCATGTCCAGTCATCCA-3' (for type 4) from AAV1 (purchased from American Type Culture Collection), an AAV2 helper plasmid pHLP19 (20), p3-2 (22), or p4-2 (3) (NcoI and SpeI sites are underlined). The resulting 1.2-kb DNA was digested with NcoI and SpeI and inserted into the corresponding sites of pFBD5LSR12 α (pFBD5LSR121, pFBD5LSR122, pFBD5LSR123, and pFBD5LSR124). The resulting recombinant baculoviruses expressing type 5 Rep78 and type 1, 2, 3, 4, or 5 Rep52 are designated Rep5/1, 5/2, 5/3, 5/4, and 5/5, respectively. The type 5 VP ORF was obtained by PCR amplification from pAAV5-2 by using primers 5'-gtaagctctctgtaagACGTCTTTTGTGATCACCCTCCAGATTGGT-3' and 5'-cgaatcagaTTAAAGGGTTCGGGTAAGGTATCG-3'. The sequence corresponding to the VP ORF is capitalized, and the initiation codon was mutated to ACG to reduce its translational efficiency. The 2.2-kb PCR product was cloned into pCMV (pCMV5VPm). The plasmid was digested with Acc65I and treated with T4 DNA polymerase and subsequently with XbaI to excise the VP ORF, which was then inserted into the BamHI (blunt)-XbaI sites of pFastBac Dual (pFBD5VPm). Plasmid expressing a chimeric VP was constructed by the use of an overlapping-PCR method as follows. VP251 was generated by PCR from pAAV5-2 using primers #30 and #31 (Table 1). The resulting PCR product was treated with BamHI and HindIII and cloned into the corresponding sites of pFBD5VPm. For VP252 construction, the type 2 VP portion was PCR amplified with primers #32 and #34 from pHLP19. The type 5 VP was amplified with primers #33 and #31. After gel purification, the two PCR products were combined and subjected to the second round of PCR using primers #31 and #32. Chimeric VP253, -254, -255, and -256 were produced in the same way except for primers for the first round of PCR. For VP253, primers #32 and #36 were used to amplify the type 2 VP1 portion and #31 and #35 to amplify the type 5 VP portion (see Fig. 3A). A PCR-generated chimeric VP1 gene was digested with HindIII and BamHI and inserted into the HindIII-BamHI sites of pFBD5VPm.

Cell culture. HEK293 cells were maintained in Dulbecco's modified Eagle's medium-F-12 (1:1, vol/vol; Invitrogen) supplemented with 10% fetal calf serum (Sigma-Aldrich, St. Louis, MO). *Spodoptera frugiperda* Sf9 cells (Invitrogen) were grown at 27°C in shake flask cultures containing Sf-900 II SFM (Invitrogen) supplemented with 10% fetal calf serum.

Western blotting and silver staining. Cells were lysed in 1 \times sodium dodecyl sulfate sample buffer and resolved on a 4 to 12% NuPAGE Bis-Tris gel (Invitrogen). After electrophoresis, separated proteins were transferred to a Durapore membrane filter (Millipore, Bedford, MA) and incubated with a primary antibody, either an anti-Rep monoclonal antibody (303.9; Research Diagnostics, Flanders, NJ) at a dilution of 1:200 or a polyclonal anti-type 5 VP antibody raised against a portion of type 5 VP3 polypeptide at a dilution of 1:50,000. The blots

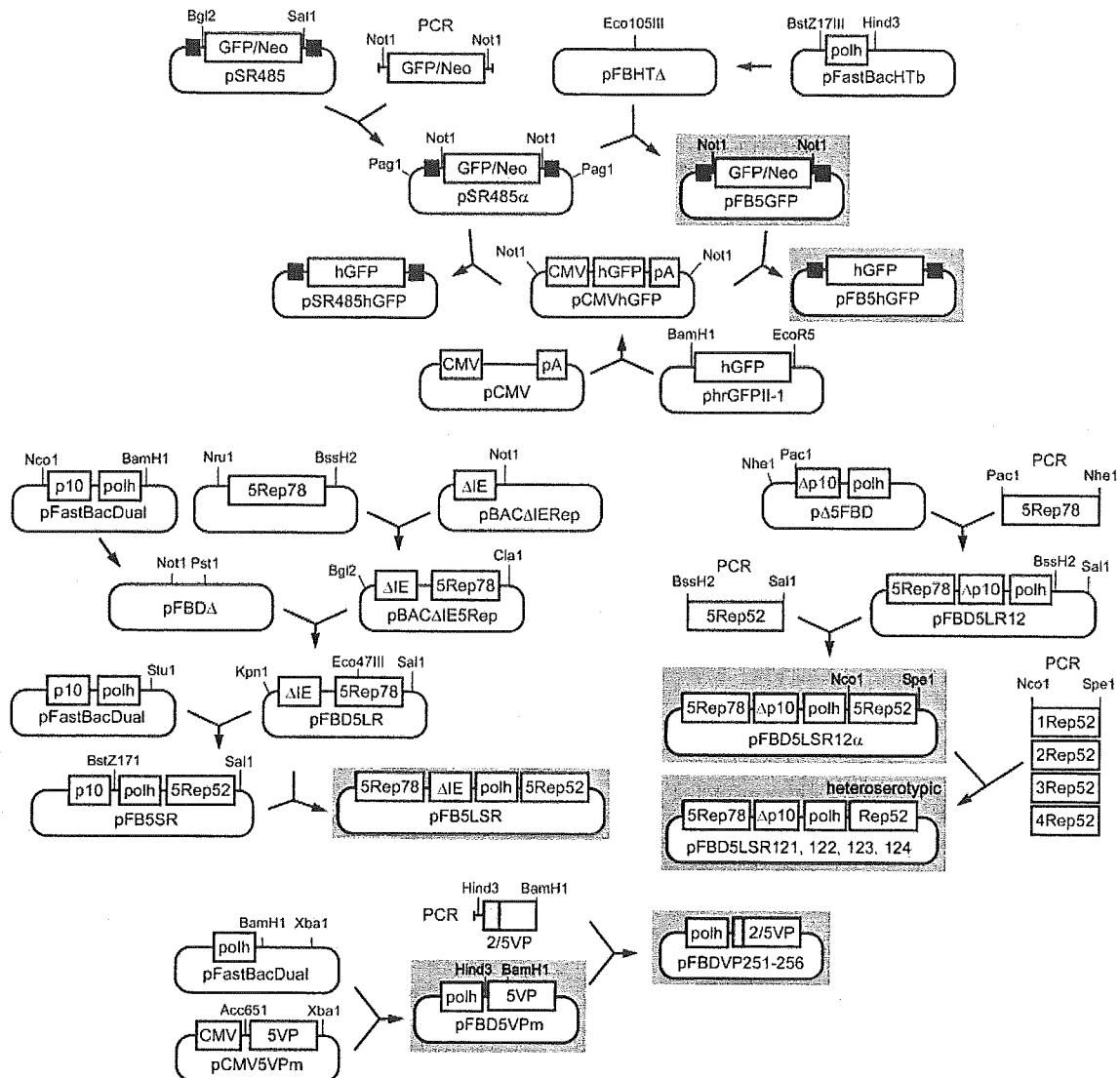


FIG. 1. Flow chart of plasmid construction. See Materials and Methods for details. Plasmids on gray backgrounds were used for generation of recombinant baculovirus vectors. Black boxes, type 5 ITR sequence; p10, p10 promoter; polh, polyhedrin promoter; pA, simian virus 40 polyadenylation sequence.

were then incubated with a secondary anti-mouse or anti-rabbit immunoglobulin G labeled with horseradish peroxidase at a dilution of 1:7,500 (Pierce, Milwaukee, WI). Membranes were incubated in Tris-buffered saline with Tween 20 (TBS-T) (10 mM Tris-HCl [pH 7.6], 0.15 M NaCl, 0.05% Tween 20, 5% nonfat dry milk). Antibodies were added to TBS-T for 1 h. After incubation, membranes were washed three times for 10 min each in TBS-T. All steps were performed at ambient temperature. The development of chemiluminescence catalyzed by horseradish peroxidase was performed according to the manufacturer's instructions (SuperSignal West Pico chemiluminescent substrate; Pierce), and the signals were detected with an X-ray film. Silver staining was performed using a SilverQuest silver staining kit (Invitrogen) according to the manufacturer's instructions.

Analysis of replicated rAAV DNA in Sf9 cells. Sf9 cells (2×10^5 cells per well) in 12-well plates were infected with GFP with or without Rep baculoviruses at a multiplicity of infection (MOI) of 3 and incubated at 27°C for 3 days. After incubation, extrachromosomal DNA was isolated by the method of Hirt (12) and a volume corresponding to 2×10^4 cells was resolved on a 0.8% agarose gel in Tris-borate buffer. Ethidium-stained gel was visualized under UV.

Production of rAAV5 in HEK293 cells. To produce rAAV5-GFP in mammalian cells, HEK293 cells at 80% confluence (approximately 10^5 cells per cm^2) in a 225- cm^2 flask were cotransfected with 27 μg of an AAV vector plasmid and 53 μg pSR487 by the calcium phosphate coprecipitation method. pSR487 harbors

type 5 *rep* and *cap* genes and adenovirus E2A, E4orf6, and VA genes (27). Two days after transfection, rAAV5 was purified as described below. For production of pseudotyped type 5 rAAV-SEAP, HEK293 cells were cotransfected with pAAVSEAP; a Rep plasmid expressing type 2 Rep78, Rep68, Rep52, and Rep40; a VP plasmid expressing VP254; and an adenovirus helper plasmid.

Production and purification of rAAV5 in Sf9 cells. Typically, 4×10^8 Sf9 cells (2×10^6 cells per ml) were infected with a Rep baculovirus (RepBac), a VP baculovirus (VPBac), and a GFP baculovirus (GFPBac) with an MOI of 1 per baculovirus construct. To generate pseudotyped 2/5 rAAV-SEAP, Sf9 cells were infected with a RepBac expressing type 2 Rep78 and Rep52, VP254Bac, and SEAPBac. Pseudotype virus refers to the ITRs of one serotype packaged into a capsid derived from a different AAV serotype. For example, rAAV2/5 consists of AAV2 ITRs packaged into an AAV5 capsid. Three days after infection, the cells were pelleted by centrifugation and lysed in a lysis buffer of 20 mM Tris-HCl (pH 8.4), 50 mM NaCl, 2 mM MgCl_2 , 0.4% deoxycholic acid, 0.5% 3-[(cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS) (Merck, Darmstadt, Germany), and 60 U per ml of Benzonase (Merck) and incubated at 37°C for 30 min. The concentration of NaCl in the cell lysate was adjusted to 150 mM and incubated for an additional 30 min. Solid CsCl was added to obtain a final density of 1.36 g/cm^3 . After centrifugation at 36,000 rpm for 24 h at 21°C using an SW40 Ti rotor (Beckman, Fullerton, CA), fractions containing rAAV5 were recovered and subjected to a second round of CsCl ultracentrifugation. For some experi-

TABLE 1. Oligonucleotides used for construction of chimeric VP genes

Primer	Sequence ^a
#30	5'- <u>gtcaagctt</u> cctgtaagAcGGCTGCCGAcGGTTATCTaCCcGA TTGGTTGGAAGAAGTTGGTGAAGGT-3'
#31	5'-GCTGGGATCCGCTGGGTCCAGCTTCGGCGT-3'
#32	5'- <u>gtcaagctt</u> cctgtaagAcGGCTGCCGAcGGTTATCTaCCcGA TTGGTTGGAAGAAGGT-3'
#33	5'-ACAGCAGGGGTCTTGTGCTGCCTGGTTATAACTA-3'
#34	5'-TAGTTATAACCAGGCAGCACAAAGCCCTGCTGT-3'
#35	5'-GACTCGACAAGGGAGAGCCTGTCAACAGGGCAGA-3'
#36	5'-TCTGCCCTGTGACAGGCTCCCTTGTGCGAGTC-3'
#37	5'-GAGACAACCCGTACTCAAGTACAACCACGGGA-3'
#38	5'-TCCCGTGGTTGTACTTGAGGTACGGGTTGTCTC-3'
#39	5'-GAGCAGTCTCCAGGCGAAGAAAAGGGTTCTCGA-3'
#40	5'-TCGAGAACCCTTTCTTCGCCTGGAAGACTGCTC-3'
#41	5'-AGGAACCTGTTAAGACGGCCCTACCGAAAGCG-3'
#42	5'-CGCTTCCGGTAGGGGCCGTCTAACAGTTCCT-3'

^a The HindIII or BamHI sites are underlined. The initiation codon for the VP1 gene was mutated to ACG. The possible splicing donor site was destroyed by introducing silent mutations. The VP ORFs are capitalized, and mutated nucleotides are indicated by lowercase letters.

ments, rAAV5 was further purified by anion-exchange column chromatography. CsCl-banded rAAV5 fractions were dialyzed against a buffer of 20 mM Tris-HCl (pH 8.4), 20 mM NaCl, 2 mM MgCl₂, and 4% glycerol and loaded onto a HiTrap Q Sepharose XL column (1-ml bed volume; Amersham Biosciences, Piscataway, NJ). Bound rAAV5 was eluted with a 20 to 500 mM linear NaCl gradient. Fractions containing rAAV5 were dialyzed against a buffer of 50 mM HEPES (pH 7.4), 150 mM NaCl, 2 mM MgCl₂, and 5% sorbitol and stored at -80°C until use. The titer of rAAV was determined by real-time PCR with CMV-specific primers 5'-TATGGAGTTCCCGTTACATAACTTACGGT-3' and 5'-GAC TAATACGTAGATGTACTGCCAAGTAGG-3' on an HT7000 genetic analyzer (Applied Biosystems, Foster City, CA). Dilutions of pSR485 were employed as a copy number standard.

Competition experiment with a type 2 or type 5 AAV receptor analog. COS cells were plated in a 12-well plate at 30% confluence 24 h prior to infection. rAAV2-GFP or rAAV5-GFP was incubated in 0 or 20 µg per ml of heparin (Sigma-Aldrich), an analog of heparan sulfate proteoglycan (HSPG), for 2 h at room temperature. The cells were infected with adenovirus (3 PFU per cell) at 37°C for 2 h. The cells were washed with medium and then infected with rAAV2-GFP at 10⁴ vg per cell or rAAV5-GFP at 10⁵ vg per cell. At 24 h postinfection, the cells were visually examined under a fluorescent microscope and the percentages of positive cells were determined by flow cytometric analysis of 10⁵ infected cells. Experiments were performed in triplicate. Competition experiments with α-2,3 sialic acid were performed as described previously (14). COS cells were plated at 30% confluence 1 day before infection in a 12-well plate. The cells were infected with adenovirus (3 PFU per cell) and incubated at 37°C for 2 h. The adenovirus-containing medium was removed, and the cells were washed with medium. The cells were then infected with rAAV2-GFP (10⁴ vg per cell) or rAAV5-GFP (10⁵ vg per cell) for 1.5 h in 0 or 0.5 mM 3'-N-acetylneuraminyl-N-acetylglucosamine (3'-SLN) (Sigma-Aldrich), an analog of α-2,3 sialic acid. The cells were washed twice with medium and further incubated for 1 day. The cells were then examined for GFP fluorescence, and the number of positive cells was measured by flow cytometry.

Treatment of cells with neuraminidase. COS cells were infected with adenovirus at 3 PFU per cell for 1 h at 37°C. The cells were treated with 0.08 U per ml of neuraminidase (*Vibrio cholerae*, type III; Sigma-Aldrich) for 1 h and infected with rAAV2-GFP at 10⁴ vg per cell or rAAV5-GFP at 10⁵ vg per cell for 2 h. The infected cells were then washed twice with medium and incubated for 1 additional day. The GFP-positive cells were counted by flow cytometry. Experiments were done in triplicate.

Muscle injection of rAAV5 in mice. A total of 10¹¹ vg of pseudotyped rAAV5-SEAP produced in either 293 cells or Sf9 cells were injected into murine tibialis anterior muscles and blood was taken at the indicated weeks after injection. The serum SEAP activity was measured by a SEAP report gene assay (Roche Diagnostics, GmbH, Penzberg, Germany). The mouse study was approved by a review board at Jichi Medical School.

RESULTS

Construction of recombinant VP and Rep baculoviruses. Production of rAAV2 in insect cells uses three baculovirus

vectors providing the following: (i) genes for three AAV structural proteins that form the virus capsid (VP1, VP2, and VP3), (ii) two of the AAV nonstructural proteins for replication and encapsidation (Rep78 and Rep52), and (iii) AAV vector DNA consisting of the gene of interest flanked by the AAV origins of replication (ITRs). In the presence of the AAV nonstructural proteins, the AAV vector DNA is "rescued" from the baculovirus genome and replicates as AAV via the ITRs (31).

Similarly to AAV type 2, the type 5 capsid proteins VP1, VP2, and VP3 are synthesized from two spliced mRNAs arising from the p41 promoter (Fig. 2A) (25). One mRNA is translated into VP1, while another transcript encodes VP2 and VP3. The initiation codon for VP2 is ACG, which is poorly utilized, resulting in the ribosome scanning through to the VP3 initiation codon AUG. The alternate usage of two acceptor sites and the poor utilization of the ACG initiation codon for VP2 are responsible for the 1:1:10 stoichiometry of VP1, VP2, and VP3. As shown in our previous report, the type 2 VP gene with an AAV intron does not express all of the VP polypeptides in insect cells (31). Mutating the VP1 AUG initiation codon to ACG resulted in production of VP1, VP2, and VP3 with a stoichiometry of approximately 1:1:10 from a single transcript without alternate splicing (31). Based on our initial success with AAV2, we constructed a similar type 5 VP baculovirus (VP5Bac) that harbored a type 5 VP gene where the initiation codon for VP1 was changed to ACG (Fig. 2B). Although this VP5Bac was able to produce type 5 capsids into which type 5 AAV vector DNA was incorporated, VP1 was poorly expressed compared to that synthesized in 293 cells (Fig. 2C). The resulting rAAV5-GFP particles poorly transduced COS cells. The calculated ratio of vector genomes to transducing units for the Sf9 cell-produced rAAV5-GFP was 10 times higher than the ratio for the 293 cell-produced counterpart. The VP1 polypeptides have phospholipase A₂ activity and are critical for efficient transfer of the viral genome from late endosomes to the nucleus (36). The efficiency with which a scanning eukaryotic ribosome recognizes an AUG codon for translational initiation is dependent on the local sequence context of the codon. The sequence ACCAUGG is optimal for initiation (18). Residue G at +4 seems particularly important for translation from a non-AUG codon where the A of the AUG codon is defined as +1 (11). In type 2 VP1, the nucleotide at +4 is G while the corresponding nucleotide at +4 in type 5 is U. To increase the efficiency of translation from an ACG codon for type 5 VP1 in insect cells, we tested some VP1 mutants that introduced a G residue at +4. However, these mutants also failed to produce infectious type 5 AAV particles (data not shown). The VP1-unique portion is conserved well among different serotypes compared to the VP3 portion that constitutes the majority of the viral capsids and is responsible for receptor binding specificity. The type 5 VP1-unique portion is approximately 70% identical to the equivalent portion of type 2 (Fig. 3A), while the type 5 VP3 portion is 60% homologous to the equivalent portion of type 2 (2). Since we successfully produced rAAV2 that was as infectious as the 293 cell-produced one, we tested a series of chimeric capsids between types 2 and 5 in which a part of the type 5 VP1-unique portion was replaced by the corresponding portion of type 2 VP1. Figure 3A shows the chimeric VP1 genes constructed. Figure 3B shows the Western analysis of type 5 VP poly-