

次に DEG に関して生物学的意味付けを行うために、個々の遺伝子のアノテーション(annotation)を調べる。National Center for Biotechnology Information (NCBI)のデータベース Entrez Gene を利用して、逐一 Gene Ontology (GO)を調べることも出来るが、DAVID Bioinformatics Resources (david.abcc.ncifcrf.gov)の Functional Annotation ツールを用いると、膨大な遺伝子セットのアノテーションを一括して解析可能である<sup>9)</sup>。多数のサンプルを比較解析する場合は、@GeneSpring(Agilent)などを用いて、DEG を指標に階層的クラスター解析を行うと、発現プロファイルのビジュアルな比較が出来る。

さらに DEG が構成している分子ネットワークを調べると、生物学的意味をより明確に把握することが出来る。生体内では、遺伝子でコードされたタンパク質は複雑なネットワークから成るシステムを構築している<sup>3)</sup>。タンパク質間相互作用には、直接的結合、活性化、不活化、酵素反応、運搬、複合体形成など多彩な作用様式が存在する。従って、複雑多岐のオミックスデータで、最も重要な役割を果たしている分子ネットワークを同定するためには、精査された文献情報に裏付けられた専用の解析ツールを使用する必要がある。すなわち、膨大な文献情報から様々な分子間相互作用を抽出し、信頼性が高い知識を整理して、コンテンツとして収録した知識データベース(knowledgebase)を用いて、既知のどのネットワークやパスウェイに最も高い類似性を呈しているかについて、統計学に解析する方法である<sup>4)</sup>。

無償で利用出来る代表的なデータベースには、Kyoto Encyclopedia of Genes and Genomes (KEGG) (www.kegg.jp), Protein Analysis Through Evolutionary Relationships (PANTHER) classification system (www.pantherdb.org), Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (string.embl.de)がある。2011 年 8 月現在、KEGG には 407 種類のレファレンスを含む 146,624 種類のパスウェイが収録されている。マイクロアレイ解析から得られた DEG のリストを、DAVID Functional Annotation ツールに入力すると、自動的に統計学的検定を行って、最も密接に関連している KEGG のパスウェイを同定することが出来る。有償ツールとしては、@Ingenuity Pathways Analysis (IPA) (Ingenuity Systems)や@KeyMolnet (Institute of Medicinal Molecular Design)がある。どちらも医学生物学などの専門家が精選された文献を精読して、信頼性の高い情報を収集しており、定期的にアップデートされている。

KeyMolnet は種々の疾患のメディエート分子を網羅的に収録しており、日本語にも対応している。解析法として、結合・発現制御・複合体形成を包括的に調べる周辺検索、発現制御に関与する転写因子群を調べる共通上流検索、始点と終点間のネットワークを調べる始点終点検索を選択することが出来る(図 2)<sup>4)</sup>。

解析ツールで調べた分子ネットワークから、創薬標的分子を探索する場合は、多数の分子からのリレーションが集中しているハブ(hub)と呼ばれる中心分子を同定することが重要である。ハブの抑制薬や活性化薬は、ネットワークのロバストネスの維持に重大な影響を及ぼし、治療効果や副作用の発現が大きい<sup>3)</sup>。

### ゲノム解析からみた MS の疾患感受性遺伝子

従来の MS の双生児や兄弟例の遺伝的研究より、何らかの遺伝因子が MS 疾患感受性に影響していると考えられ、候補遺伝子解析により、主要適合性複合体(major histocompatibility complex: MHC) 遺伝子座の一塩基多型(single nucleotide polymorphism; SNP)との関連が示唆されていた。近年、マイクロアレイを用いた網羅的なゲノムワイド関連解析(genome-wide association study: GWAS)により、いくつかの MS リスクアレルが同定された<sup>7)</sup>。この研究では、931 家系のトリオの全ゲノム上の SNP をスクリーニングし、別の 609 家系のトリオ、2,322 例の孤発性 MS、789 例のコントロールと 2 種類の外部データベースコントロールで再現性を検証した。最終的に、IL2RA, IL7RA の SNP が MS 発症と関連していることがわかった。その後追従する研究により、20 以上の MS リスクアレルが同定され、最近、15 ヶ国 23 研究グループによって収集された 9,772 症例を含む大規模共同研究により、従来の報告の再現性が確認され、さらに 29 の新規のリスクアレルが追加された<sup>8)</sup>。しかしながら、GWAS で同定された MS リスクアレルのオッズ比はおおむね 2 以下で、疾患発症に対する貢献度は小さい。そのため、次世代シーケンサーを用いて全ゲノムや全エクソン(エクソーム)を網羅的に解析し、効果の大きい遺伝子バリエーションを同定する試みが始まっている。Baranzini らは、片方が MS、他方が健常者である一卵性双生児 3 組の CD4 陽性 T 細胞を用いて、全ゲノム、トランスクリプトーム、メチル化部位配列を@GAIIX シーケンサー(Illumina)で網羅的に解析した<sup>9)</sup>。しかしながら、MS 発症に関与する遺伝子上の差異を検出出来

なかった。

### トランスクリプトーム解析からみた MS の免疫病態

現在までに、多くの研究者が主として末梢血リンパ球 (peripheral blood mononuclear cells: PBMC) を用いて、トランスクリプトームを網羅的に解析し、MS と健常者、再発期と寛解期、IFNB レスポンダーとノンレスポンダーの病態を比較している<sup>4)</sup>。

Achiron らは、26 例の RRMS 患者と 18 名の健常者の PBMC の遺伝子発現プロフィールを比較解析した<sup>10)</sup>。両群間で、1,109 遺伝子の発現差異を認め、MS における T 細胞活性化関連遺伝子 LEF1, TCF3, SLAM, ITGB2, CTSB の上昇を報告した。しかし 14 例の患者では、採血時に IFNB や Copaxone、イムノグロブリンを投与中であり、治療薬が遺伝子発現に直接影響した可能性がある。Stürzebecher らは、10 例の RRMS 患者で IFNB 治療前後に末梢血リンパ球を解析し、レスポンダーで 25 遺伝子の発現変動 (IFI17, OAS, Stat1 などの上昇) を確認した<sup>11)</sup>。しかし、彼らは一度凍結保存したリンパ球を解凍後に解析しており、実験操作が遺伝子発現に影響した可能性を否定出来ない。

われわれは、72 例の IFNB 未治療の活動性 MS 患者 (そのうち 46 例は初回採血直後から 2 年間 IFNB 治療を開始) と、年齢と性を一致させた 22 名の健常者の末梢血リンパ球から、CD3 陽性 T 細胞と CD3 陰性 non-T 細胞を分離し、各細胞分画の遺伝子発現プロフィールを解析した<sup>12)</sup>。両群間で発現差異を認めた上位 30 遺伝子では、T 細胞で 25 遺伝子 (NR4A2, TCF8 の上昇と MAPK1, SMARCA3, HSPA1A, TRAIL, TOP1, CCR5, BAG1, DAXX, TSC22, PARP の低下), non-T 細胞で 27 遺伝子 (ICAM1, CDC42, RIPK2, SODD, TOP2A の上昇と BCL2, RPA1, NFATC3, HSPA1L, RBBP4, PRKDC の低下) が、アポトーシス制御遺伝子に分類された。すなわち患者の末梢血リンパ球では、アポトーシスに関して促進遺伝子と抑制遺伝子の拮抗的バランスが存在していると考えられた。

またわれわれは、MS と健常者の両群間の多重検定と Bonferroni 補正を行い、T 細胞で有意な発現差異を示す 286 遺伝子を同定した<sup>13)</sup>。KeyMolnet によるネットワー

ク解析で、共通上流として NF- $\kappa$ B を介する遺伝子発現制御系を検出した。NF- $\kappa$ B はサイトカインやケモカインの発現を正に調節する転写因子で、炎症の増幅と遷延化に働く。すなわち MS の T 細胞では NF- $\kappa$ B を介する遺伝子発現制御系に恒常的な異常が存在すると考えられる。さらに 286 遺伝子を指標遺伝子として階層的クラスター解析を行い、臨床データとの関連性を評価した(図 3)<sup>13)</sup>。286 遺伝子は 5 クラスに分類され、患者群は健常者群から識別され、さらに 4 グループ A, B, C, D に分類された。A 群は遺伝子発現プロファイルが最も健常者群に類似しており、軽症者が多く、B 群は臨床的活動性が最も高く、多数のケモカイン遺伝子が集積している class #5 遺伝子群の発現レベルが最も高く、C 群は脳限局病変を呈する患者が多く、D 群は神経学的機能障害度(Expanded Disability Status Scale: EDSS)スコアが最も高値であった。IFNB 治療を開始した 46 例に関して 2 年間追跡し、治療前後 2 年間の再発回数、ステロイドパルス日数、入院日数、EDSS、MRI T2 強調画像病巣数、患者の治療に対する満足度をスコア化して、IFNB レスポンダーとノンレスポンダーに分類すると、レスポンダーは A 群と B 群に集積していた。またレスポンダーでは IFN 応答遺伝子群(IFN-responsive genes: IRG)の発現レベルが、治療開始 6 ヶ月の時点でも高レベルに維持されていたが、ノンレスポンダーでは低下傾向を呈した。すなわち階層的クラスター解析による病型分類と IRG の経時的発現変化を組み合わせると、IFNB 治療反応性をある程度予測出来ることがわかった。最近 Comabella らは、IFNB ノンレスポンダーは治療前から I 型 IFN の産生が恒常的に高い可能性を指摘している<sup>14)</sup>。

さらにわれわれは、末梢血リンパ球を IFNB で刺激した場合、3 時間以内に CXCR3 リガンドケモカインである CXCL11, CXCL10, CXCL9 と CCR2 リガンドケモカインである MCP1, MCP2 の発現が、100 倍以上上昇することを見出した<sup>15)</sup>。前者はエフェクター Th1 細胞、後者は単球やマクロファージの遊走を促進し、炎症を増強するケモカインである。すなわち多くのケモカインが、早期 IRG として検出された。このようなケモカインバーストは、MS で IFNB 治療開始後早期に出現する発熱、皮膚潰瘍形成、肝障害の発現と関連している可能性がある。

最近われわれは、ハンガリー人一卵性双生児 MS ペア 4 組(MS/MS, MS/MS, MS/MS, MS/健常者)より分離した末梢血 CD3 陽性 T 細胞を解析し、MS 病態特異的

20 遺伝子を同定した<sup>16)</sup>。KeyMolnet によるネットワーク解析で、共通上流として T 細胞の生存と分化を制御する転写因子 Ets を介する遺伝子発現制御系を検出した。

またわれわれは、6 例の RRMS 患者を追跡して再発期と寛解期に採血し、CD3 陽性 T 細胞を分離して比較解析し、再発期特異的 43 遺伝子を同定した<sup>17)</sup>。KeyMolnet によるネットワーク解析で、共通上流として NF- $\kappa$ B を介する遺伝子発現制御系を検出した。以上の知見は、NF- $\kappa$ B は MS と健常者を識別し、MS の再発期と寛解期を識別する分子ネットワークを制御するハブ的な転写因子として働き、MS 治療薬の創薬ターゲットと成り得ることを示唆している。実際に NF- $\kappa$ B 阻害薬は、MS モデル動物系である自己免疫性脳脊髄炎(experimental autoimmune encephalomyelitis: EAE)において有効性が確認されている。

この他の重要な研究として、Lock らは、MS 脳組織の急性炎症性病巣と慢性非活動性病巣の遺伝子発現プロファイルを比較し、前者における G-CSF の上昇と後者における IgG Fc receptor, IgE receptor, histamine receptor type 1 の上昇を認め、G-CSF 投与で EAE が軽症化し、FcR  $\gamma$ -chain 遺伝子欠損マウスで EAE の慢性化が抑制されることを証明した<sup>18)</sup>。

### プロテオーム解析からみたMSの脳分子病態

2008 年に、Han らは 6 例の MS 患者の凍結脳を用いて、病巣から病理学的ステージを確認後にレーザーマイクロダイセクションでサンプルを採取し、電気泳動で分離後にタンパク質を抽出しペプチドを精製して、質量分析装置を用いて網羅的に解析した<sup>19)</sup>。病理学的ステージに関しては、炎症性細胞浸潤と浮腫を主徴とする急性脱髄巣(active plaque: AP), 炎症が脱髄巣辺縁部に限局している慢性活動性脱髄巣(chronic active plaque: CAP), 炎症所見に乏しくグリア瘢痕を主徴とする慢性非活動性脱髄巣(chronic plaque: CP)に分類した。また同時に 2 例の健常脳も解析した。健常脳では検出されず、かつ各ステージ特異的なタンパク質を同定し、AP 158, CAP 416, CP 236 種類のプロテオームデータを公開した。彼らは CAP において 5 種類の血液凝固系タンパク質を同定した。この所見に基づき、抗凝固薬で EAE を治療することに成功し、血液凝固系タンパク質が新規 MS 創薬標的分子となることを示した。しかしなが

ら、大多数を占める凝固系以外のタンパク質に関しては、MS 脳分子病態における意義は明らかされなかった。

われわれは、Hanらのプロテオームデータを KEGG, PANTHER, IPA, KeyMolnet に入力し、ステージ特異的プロテオームを最も良く反映している分子ネットワークを同定した<sup>20)</sup>。これら 4 種類のツールは独自の分子ネットワークを検出したが、共通して、CAP, CP プロテームにおける細胞外基質(extracellular matrix: ECM)-インテグリンシグナル伝達系の中心的役割が示唆された(図 4)<sup>20)</sup>。インテグリンは複数の $\alpha$ ,  $\beta$ サブユニットから構成される 24 種類のヘテロダイマー分子で、細胞膜上に発現し、ECM のリガンドとして働く。 $\beta 1$  インテグリンファミリーはコラーゲン、フィブロネクチン、ラミニンと結合し、 $\alpha v$  インテグリンファミリーはビトロネクチンと結合する。ECM-インテグリン系は細胞接着、遊走、分化、増殖に必須なシグナルを伝達する。MS 慢性病巣においては、髄鞘や軸索の再生が著しく乏しい。その理由として、グリア瘢痕に含まれている ECM が軸索再生阻害因子として働く可能性や、活性化マクロファージやミクログリアが産生するタンパク質分解酵素が ECM に結合して長期に保持され、その結果、髄鞘崩壊が遷延化している可能性がある。現在、欧米では MS 再発抑制目的で、抗 $\alpha 4 \beta 1$  integrin モノクローナル抗体®Natalizumab が用いられている。しかしながら、Natalizumab は進行性多巣性白質脳症(progressive multifocal leukoencephalopathy: PML)を惹起する危険性があり、より安全な薬の開発が必要である。ECM-インテグリンシグナル伝達系では、focal adhesion kinase (FAK)がハブとして働く(図 4)<sup>20)</sup>。従って FAK を標的とする治療薬は、慢性炎症性脱髄抑制薬の候補となる可能性があり、今後の開発が待たれる。

## Key Words

- ・ 次世代シーケンサー

最近数年のうちに急速に進歩した電気泳動を行わずに、高速で並列して塩基配列を解析出来る装置。通常は、30 塩基程度の短い配列(ショートリード)を 1 度のランで 10 億塩基(1Gb)以上決定出来る。取得した膨大なデータに関しては、参照配列にマッピングしたり、連結(アセンブリ)することにより連続配列を構築する。全ゲノム、全エクソン、RNA(トランスクリプトーム)、メチル化部位、転写因子結合部位(ChIP-seq)などの配列を網羅的に解析出来る。

- ・ オミックス

細胞における遺伝子 RNA (トランスクリプトーム)、タンパク質(プロテオーム)、代謝物質(メタボローム)の網羅的発現情報を統合して、個々の細胞の生理学的・病理学的全体像を把握する研究の流れ。

- ・ テーラメイド医療

ヒトゲノムが解読され、個人で異なる塩基配列の多様性が明らかにされた。このような遺伝子多型情報やDNAマイクロアレイや次世代シーケンサーによる網羅的遺伝子発現情報に基づいて、薬物応答性の個人差を予測可能となり、個々の患者に最適な治療方法を計画出来るようになった。

- ・ 正規化

特定のルールに従ってデータを変換して利用し易くすることを言う。  
®GeneChip(Affymetrix)では、各チップのデータを MAS5 や RMA というアルゴリズムに従って正規化してから、チップ間の比較解析を行うことが多い。

- アノテーション

あるデータに対して注釈として付与された関連する情報のこと。ここでは遺伝子の機能に関連する情報を指す。Gene Ontology (GO)には、Cellular Function, Cellular Process, Cellular Component の分類がある。

- ゲノムワイド関連解析

大規模な集団のDNAサンプルをジェノタイピングアレイで解析し、全ゲノムの遺伝子多型の頻度を評価することで、疾患や形質の原因となるアリルを探索する解析方法。疾患や形質をもたないコントロール集団と比較して、特異的な多型が高頻度に集積していれば、その多型は疾患に関連していると判断出来る。



## 文献

1. Oksenberg JR, Baranzini SE. Multiple sclerosis genetics--is the glass half full, or half empty? *Nat Rev Neurol* 2010; 6: 429-437.
2. Lassmann H, et al. The immunopathology of multiple sclerosis: an overview. *Brain Pathol* 2007; 17: 210-218.
3. Kitano H. A robustness-based approach to systems-oriented drug design. *Nat Rev Drug Discov* 2007; 6: 202-210.
4. Satoh J. Bioinformatics approach to identifying molecular biomarkers and networks in multiple sclerosis. *Clin Exp Neuroimmunol* 2010; 1: 127-140.
5. 佐藤準一. アレイインフォーマティクスの進展. *YAKUGAKU ZASSHI* 2008; 128: 1537-1545.
6. Huang da W, et al. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; 4: 44-57.
7. The International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007; 357: 851-862.
8. The International Multiple Sclerosis Genetics Consortium, the Wellcome Trust Case Control Consortium. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011; 476: 214-219.
9. Baranzini SE, et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature* 2010; 464: 1351-1356.
10. Achiron A, et al. Blood transcriptional signatures of multiple sclerosis: Unique gene expression of disease activity. *Ann Neurol* 2004; 55:410-417.
11. Strüzebecher S, et al. Expression profiling identifies responder and non-responder phenotypes to interferon- $\beta$  in multiple sclerosis. *Brain* 2003; 126: 1419-1429.
12. Satoh J, et al. Microarray analysis identifies an aberrant expression of apoptosis and DNA damage-regulatory genes in multiple sclerosis. *Neurobiol Dis* 2005; 18: 537-550.
13. Satoh J, et al. T cell gene expression profiling identifies distinct subgroups of

- Japanese multiple sclerosis patients. *J Neuroimmunol* 2006; 174: 108-118.
14. Comabella M, et al. A type I interferon signature in monocytes is associated with poor response to interferon- $\beta$  in multiple sclerosis. *Brain* 2009; 132: 3353-3365.
  15. Satoh J, et al. Microarray analysis identifies a set of CXCR3 and CCR2 ligand chemokines as early IFN $\beta$ -responsive genes in peripheral blood lymphocytes: an implication for IFN $\beta$ -related adverse effects in multiple sclerosis. *BMC Neurology* 2006; 6: 18.
  16. Satoh J, et al. Aberrant transcriptional regulatory network in T cells of multiple sclerosis. *Neurosci Lett* 2007; 422: 30-33.
  17. Satoh J, et al. Molecular network analysis of T-cell transcriptome suggests aberrant regulation of gene expression by NF- $\kappa$ B as a biomarker for relapse of multiple sclerosis. *Dis Markers* 2008; 25: 27-35.
  18. Lock C, et al. Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nature Med* 2002; 8: 500-508.
  19. Han MH, et al. Proteomic analysis of active multiple sclerosis lesions reveals therapeutic targets. *Nature* 2008; 451: 1076-1081.
  20. Satoh J, et al. Molecular network of the comprehensive multiple sclerosis brain-lesion proteome. *Mult Scler* 2009; 15: 531-541.

### Further reading

- ・ 藤渕 航, 堀本勝久(編). マイクロアレイデータ統計解析プロトコール. 東京: 羊土社; 2008.  
マイクロアレイデータ解析入門書としてお勧め
- ・ 小田吉哉, 長野光司(編). 創薬・タンパク質研究のためのプロテオミクス解析. 東京: 羊土社; 2010.  
プロテオーム解析入門書としてお勧め

## 図の説明

### 図 1. 網羅的発現解析から分子ネットワーク解析へ.

比較対象となる遺伝子発現レベルが異なる数種類以上の細胞や組織から RNA を精製し、蛍光標識して、アレイとハイブリダイゼーションを行う。スキャン後に、シグナル強度を正規化し、サンプル間の遺伝子発現プロファイルを統計学的に比較解析し、有意な発現差異を呈する遺伝子群(DEG)を抽出し、定量的 PCR で検証する。生物学的意味付けのため、Gene Ontology (GO)のアノテーションを調べ、階層的クラスター解析を行い、KEGG, PANTHER, STRING, IPA, KeyMolnet を利用して分子ネットワークを解析する。図は文献 4 より引用改変。

### 図 2. KeyMolnet による MS 疾患メディエート分子のネットワーク解析.

KeyMolnet に収録されている 91 種類の多発性硬化症疾患メディエート分子(赤)を入力して、上下流 1 パス周辺検索法で、ネットワークを解析した。913 分子と 1,005 分子リレーションから構成される複雑なネットワークが検出された。ハブ分子として働く vitamine D receptor (VDR) (青楕円)による発現調節との関連性が最も強く示唆された。図は文献 4 より引用改変。

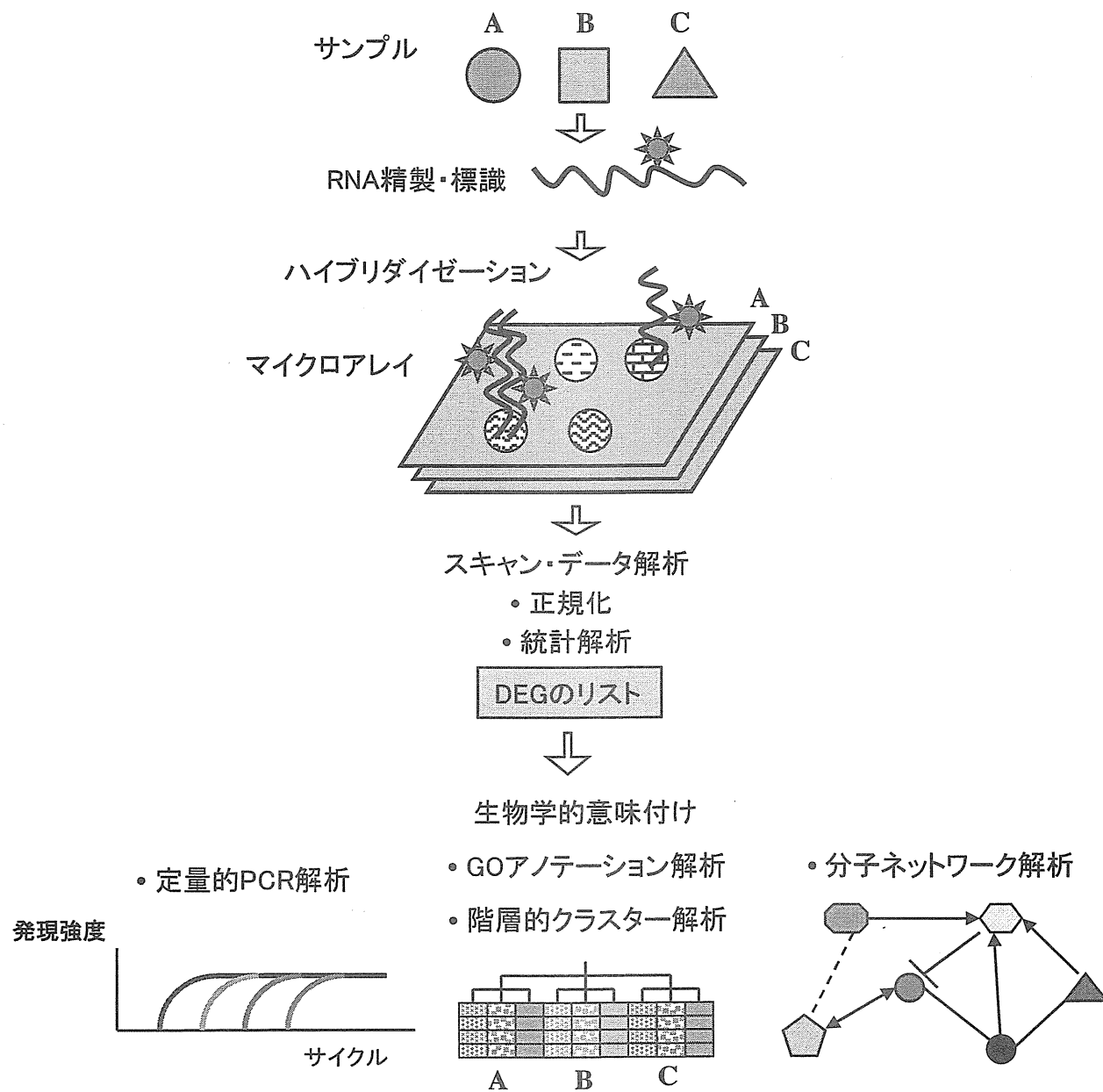
### 図 3. 階層的クラスター解析による MS の病型分類.

患者(MS; n =72)と健常者(CN; n = 22)の末梢血 CD3<sup>+</sup> T 細胞で有意な発現差異を示す 286 遺伝子を抽出し、指標遺伝子として階層的クラスター解析を行った。286 遺伝子は 5 クラスに分類され、MS 群は CN 群から分離され、4 群 A, B, C, D に分類された。IFNB レスポンダーは A 群と B 群に集積していた。図は文献 13 より引用改変。

### 図 4. MS 脳病巣プロテオームの分子ネットワーク解析.

PANTHER による MS 脳病巣 CAP プロテオームの分子ネットワーク解析では、インテグリンシグナル伝達系との有意な関連性が示唆された。Reference pathway 上の分子とヒットした CAP タンパク質をピンク色で示す。Focal adhesion kinase (FAK)(青楕円)が、

ネットワークのハブ分子として働いていることがわかる。図は文献 20 より引用改変。



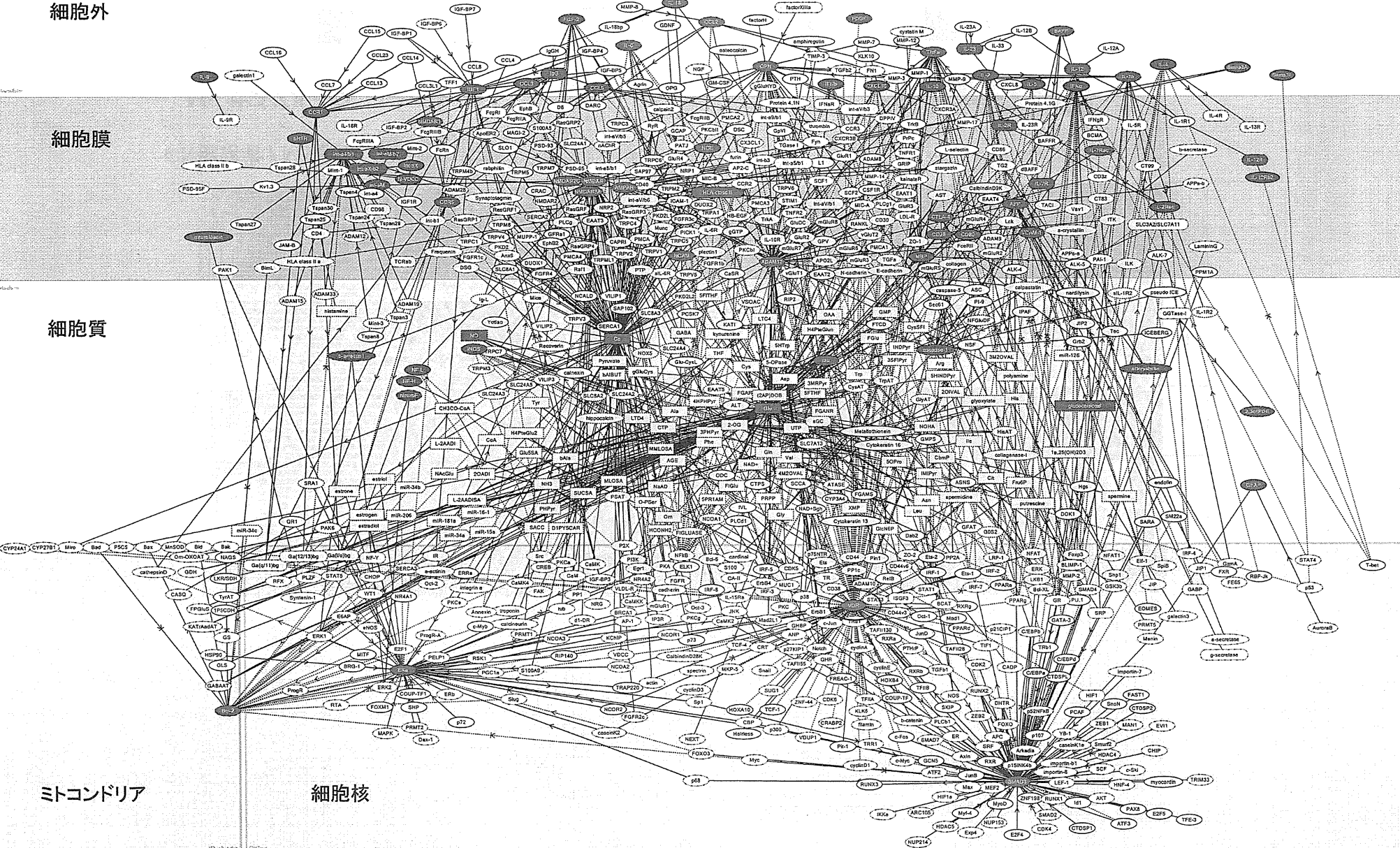
細胞外

細胞膜

細胞質

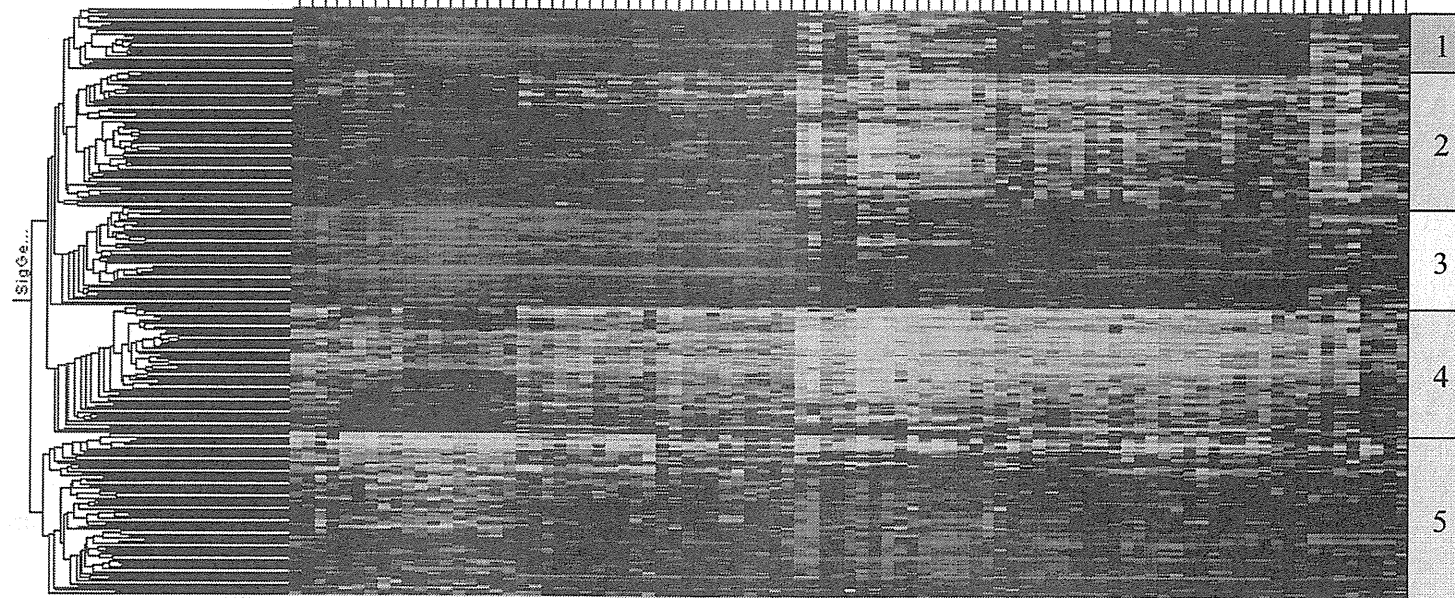
ミトコンドリア

細胞核



94サンプル  
286遺伝子

SigGene(pierre,Nc\_vs\_pre,e=0.25),Exp(Nc+pre-5),Std.Cor.



CN健常者 / MS患者

MS サブグループ



