

# 高速ゲノム解析システムの開発と遺伝的背景を支配する遺伝子の解析

理化学研究所 ゲノム科学総合研究センター プロジェクトディレクター 林崎良英

ゲノム DNA 上に書きこまれている情報は、塩基配列と修飾(メチル化)が、現在知られている限り、ゲノム DNA 上にある情報の全てである。本プロジェクトでは、高速に引き出すシステムを開発し、遺伝子単離に応用するため、その各々につき最高速のシステムとして、RISA sequencer (RISA; Riken Integrated Sequence Analysis System) とRLGS(Restriction Landmark Genomic Scanning)の総合システムの開発に成功した。また、それを用いて、ヒトゲノム上から Genomic Imprinting という遺伝形式を指標にし、ヒト新生児一過性糖尿病遺伝子、偽副甲状腺機能亢進症の遺伝子の positional candidate cloning に成功したので報告する。

ゲノム DNA 上の塩基配列情報解読を高速化する為に、高速シーケンシングシステムの開発を行った。シーケンシングの速度を上げるためには、多本並列化と解読長の改善がある。本プロジェクトでは、多本並列化の方針を採用し、384 Capillary Sequencer (RISA シーケンサー) を開発した。このマシンは、384 プレートから一度に 384 キャピラリーアレイを通してサンプルが Injection でき、独自のゲル泳動系で平均 600 から 650bp を解読できる。一台により、90000bp/台/時間が実現できるようになった。また、理研で開発した完全長 cDNA 技術やプラスミド調製機などその他一連のシステムと合わせ、50000 検体/日/16 台の速度で、生体組織からデータまでの 384 穴プレートによる一貫システムが完成した。これにより、マウスの全完全長 cDNA を収集する理研マウス cDNA プロジェクトが急速に加速された。この RISA シーケンサーは、現在商品化されるに至った。また、メチル化をスクリーニングするシステムとして、*NofI* のようなメチル化感受性 GC-rich 制限酵素サイトランドマークを二次元パターン上に視覚化する RLGS を開発した。このシステムにより、ゲノム上の多くの CpG-island をスクリーニングできる。この二次元パターンを迅速かつ精度良く行うデバイスの一つとして、シンチレーションファイバーを用いた β線 2 次元画像検出器を開発し、そのパターンを自動比較するコンピュータープログラムも新規開発した。

これらを用いて、ヒトの Imprint 遺伝子のスクリーニングを行い、それを用いてヒト新生児一過性糖尿病遺伝子、偽副甲状腺機能亢進症の遺伝子を単離した。

ゲノミックインプリンティングとは、ある遺伝子が父親由来または母親由来のアリールより選択的に発現する現象をいう。多くのインプリント遺伝子の近傍から、メチル化が一方の親由来のアリールに特異的な領域が見つかっており、インプリントの成立と維持に必須な要素であると考えられている。Bonthonron らの報告したヒト単為発生キメラは、そのリンパ球が全て単為発生由来で、インプリント遺伝子の母親アリール特異的なメチレーションを保持している事が確認されている。一方、胞状奇胎は雄核発生を原因として発症し、メチレーションの状態も父親アリールの状態を反映している。そこで、これらのサンプルのメチル化状態を上記の RLGS 法で比較し、ヒトインプリント遺伝子の系統的な探索を行った。全ゲノム領域にわたってホモ接合体の細胞を使用するので、多型を示さないスポットであっても、ゲルパターン上の 2000 から 3000 の全てのスポットが検出できる。

この戦略を用い、約 4,500 の座位のメチル化状態をスクリーニングする事に成功し、インプリントの挙動を示す二つの座位、A20 と NV149 を発見した。A20 は *G-protein-α-subunit 1 (GNAS1)* 遺伝子の Alternative Splice された転写単位中にあることがわかり、この遺伝子がインプリントしていることが判明した。この転写単位は従来のインプリントしていない転写単位の上流から始まる新しい転写単位でこの新しい *G-protein-α-subunit 1 (GNAS1)* 遺伝子の転写単位こそが、インプリント疾患、偽副甲状腺機能亢進症 type Ia の原因遺伝子であることを突きとめた。

もう一方、NV149 はその症状がインプリントしている新生児一過性糖尿病 (TNDM) の領域に位置している事を発見した。TNDM は UPD(Uniparental Disomy) の患者や 6q24 領域の染色体転座の研究から、父方対立遺伝子特異的に発現している遺伝子の発現量が増加することが原因と考えられていた。NV149 近傍の転写単位を調べた結果、*ZAC / PLAGL1* が父親由来のアリールから特異的に発現していることを明らかにした。*ZAC / PLAGL1* は Zinc-Finger 蛋白質をコードしており、転写因子である事、癌抑制遺伝子である事、アポトーシスを起こす事などが既に示されていた。さらに最も興味深い点として、PACAP (pituitary adenylate cyclase activating polypeptide)-type 1 receptor を誘導することが報告されている。PACAP は、膵臓では、グルコース存在下でインスリンの分泌を促進する事が報告されている。このことから、*ZAC / PLAGL1* が新生児一過性糖尿病の原因遺伝子であると考えられる。

## < 論文 >

1. Sasaki N., et al., Transcriptional sequencing: A method for DNA sequencing using RNA polymerase, *Proc. Natl. Acad. Sci. USA.*, 95, 3455-3460 (1998)
2. Carninci P., et al., Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA, *Proc. Natl. Acad. Sci. USA.* 95, 520-524 (1998)
3. Izawa M., et al., Recognition sites of 3'-OH group by T7 RNA polymerase and its application to transcriptional sequencing, *J. Biol. Chem.* 273, 14242-14246 (1998)
4. Hayward B.E., et al., The human *GNAS1* gene is imprinted, and encodes distinct paternally and biallelically expressed G proteins, *Proc. Natl. Acad. Sci. USA.*, 95, 10038-10043 (1998)
5. Mizuno Y., et al., Increased specificity of reverse transcription priming by trehalose and oligo-blockers allows high-efficiency window separation of mRNA display, *Nucleic Acids Res.*, 27, 1345-1349 (1999)
6. Itoh M., et al., Automated filtration-based high-throughput plasmid preparation system, *Genome Res.*, 9, 463-470 (1999)
7. Fukunishi Y., et al., Prediction of human cDNA from its

homologous mouse full-length cDNA and human shotgun database, *FEBS Lett.*, 464, 129-132, 1999

8. Kamiya M., et al., The cell cycle control gene *ZAC/PLAGL1* is imprinted - a strong candidate gene for transient neonatal diabetes, *Hum. Mol. Genet.*, 9, 453-460, 2000
9. Komatsu S., et al., Methylation and downregulated expression of mac 25/insulin-like growth factor binding protein-7 is associated with liver tumorigenesis in SV40T/t antigen transgenic mice screened by Restriction Landmark Genomic Scanning for Methylation (RLGS-M), *BBRC*, 267, 109-117, 2000
10. Morimoto K., et al., Scintillating Fiber Imager for RLGS, *IEEE Transaction on Nuclear Science*, in press
11. Tateno M., et al., Identification of a novel member of the SNAG repressor family, *mlt 1*, which is methylated and repressed in mouse liver tumor, submitted
12. Shibata K., et al., RIKEN integrated sequence analysis system (RISA system) - 384-format sequencing pipeline with 384 multi-capillary sequencer, submitted

## < NEWS >

*Science*, 278, 1700, 1997, *Nature*, 397, 98-99, 1999, *Science*, 283, 455, 1999, *Science*, 286, 445, 1999, *Science*, 288, 257, 2000