

## ヒトのゲノムと病気

国立長寿医療研究センター  
運動器疾患研究部長  
池田 恭治

我々のゲノムの“ドラフト”配列(とにかく一通りという意味で、読めないギャップも多かった)が発表されたのが2001年。あれから11年経った。あのとき、ゲノムの全配列をとにかく解読するという数十年にわたる big project 達成の興奮とともに、ヒトの遺伝子の数がたかだか2万個程度で下等動物とたいして変わらないことがわかり、どうして我々の体や心の複雑性がその程度の数で説明できるのか、という大きな疑問が残った。また30億塩基対からなるDNA配列のうち、タンパクを作り機能的に意味のあるのは1%以下で、ゲノムの大部分は意味のない“ジャンク”(ごみ)DNAということになった。

この11年で、ゲノム・エピゲノム研究はさらに発展し、我々の生物を見る見方も大きく進化した。ひとつは、DNA→RNA→タンパクというセントラルドグマとは別の、タンパクをコードしないRNAに関する情報が急ピッチで蓄積されてきた。また、ACGTからなるDNAの塩基配列以外の情報、とりわけDNAのメチル化やDNAがまきつくヒストンというタンパクの化学修飾によって伝達される遺伝情報、すなわちエピゲノム研究が爆発した。実際エピゲノムを標的としたくすりや、癌や白血病の治療に応用されつつある。

DNA/RNAの配列を読む技術と低価格化も飛躍

的に進んだ。2003年にスタートしたENCODE (EN-Cycropedia Of Dna Elements) という大きなプロジェクトの最初の成果がついこの間発表され (nature 9月6日号)、ゲノムの80%以上はなんらかの機能をもっていることがわかってきた。NGS (Next Generation Sequencer, 次世代シーケンサー) を用いた deep sequencing (大規模配列解析、“deep”とはひとつの塩基が読まれる回数が多いこと) によって、ゲノムから転写されたRNA配列を解析すると、ゲノムのかなりの部分はRNAには読まれており、タンパクはコードしないがRNAとして機能していることが想定される。

近年、ヒトの病気、とくに癌や糖尿病、膠原病など多因子疾患の成因を解析する手段としてGWAS (ジーワス、と読む、Genome-Wide Association Studyの略、ゲノム全体にわたってDNA配列の個人差と疾患感受性との関連を調べる研究) がさかに行われているが、大部分の場合、病気と関連するゲノム変化は、タンパクをコードする部分以外の多型におちている。今回 ENCODE プロジェクトによって発表され、あるいは今後蓄積されていく莫大な量のゲノム情報によって、これら複雑な疾患の成り立ちが近い将来もっと明確な形で我々の前に提示されることになるかもしれない。