

## 別紙9 (補足資料)

<p>サブテーマ名：共同研究全般にわたるデータベース構築・管理          小テーマ名：バイオインフォマティクスの手法を用いたデータマイニング (1)</p>
<p>サブテマリーダー (所属、役職、氏名) 千葉県産業振興センター (研究員) 甲賀弘、村上 雅利          サブテマリーダー (所属、役職、氏名) 千葉県産業振興センター (研究員) 三沢 計治          研究従事者 (所属、役職、氏名) 千葉県産業振興センター (研究員) 三沢 計治</p>
<p>研究の概要、新規性及び目標</p> <p>①研究の概要          ヒトおよびマウスの脳を中心とした領域で発現するmRNAの配列を比較解析し、配列と発現量の関連を詳細に調べ、さらに進んで配列から機能を解明する手がかりとすることを目指す。そのために、まず、mRNAの配列を比較する方法の理論的研究を行い、もっともふさわしい比較解析法を開発した。その後、ヒトおよびマウスのmRNAの配列解析を行い、配列上のどこにどのような違いがあるのか調べた。その結果を発現部位や発現量と比較し、機能が配列へ与える影響を調べた。</p> <p>②研究の独自性・新規性          大量の配列データを使うことにより、精度の高い解析が可能となった。また発現量についても、同一の基準を用いて測定することで、公平な比較が可能となった</p> <p>③研究の目標 (フェーズ毎に数値目標等をあげ、具体的に)          フェーズ I (H14-H15) / H15年度より従事: mRNA解析法の開発、および、先行研究で得られていたmRNA配列の解析を行う。          フェーズ II (H16-H18) : ヒトおよびマウスのmRNAの発現データを用い、mRNAの上のたんぱく質をコードしている領域(CDS)ならびにその前後(UTR)の配列と発現量を比較する。また、ヒトとマウスの突然変異率を調べ、配列の変化がどのように生じたか調べる。          フェーズ III (H18. 11. 25事業完了日以降) :          配列の変化が遺伝子の機能の変化にどのように影響しているか解析を続ける。</p>
<p>研究の進め方及び進捗状況 (目標と対比して)</p> <p>ヒトおよびマウスの脳を中心とした領域で発現するmRNAの配列を比較解析した。材料としてはかずさDNA研究所との共同研究で得られたKIAA/mKIAAを用いた。解析にあたってはmRNAを2つの領域に分けた。たんぱく質をコードしている領域 (以下CDS) としていない領域(UTR)である。</p> <p>まず、UTRについて、配列と発現量の関連を詳細に調べた。すると、3' UTR領域の長さや発現量に相関があることがわかった。この結果はDNA researchに発表された(Okazaki et al. 2003)。3' UTRの中に発現量をコントロールしている配列があるかどうか探索したが、見つけることができなかった。</p> <p>次に、CDSについて、ヒトおよびマウスのmRNAの発現量との比較を行った。ひとつのアミノ酸をcodeするDNAの3つ組 (codon) は多くの場合複数あり、同じアミノ酸配列を作る遺伝子でもDNA配列にわずかな違いが許されていて、どのcodonを使うか、頻度が遺伝子ごとに違っている。先行研究から、大腸菌やショウジョウバエなどの実験用モデル生物では、発現量とcodonの使用頻度に相関があることが知られている。なぜならば、使用されるcodonによって、RNA配列からアミノ酸配列へと翻訳されるときに効率が違うためである。大腸菌やショウジョウバエなどでは、このことを逆に利用し、配列から発現量を推定することが行われることもある。しかしながら、発現とcodon頻度の相関は、ヒトやマウスでは見つかっていなかった。そこでわれわれは、codon頻度と発現量の相関に注目して解析を行った。しかしながら、先行研究と同じく、相関は見られなかった。</p> <p>相関が出てこない理由のひとつとして、突然変異の偏りがあると考えられる。ヒトやマウスのcodon頻度は突然変異の偏りに影響を受けている。われわれは、特にCpG hypermutabilityに注目した。CpG hypermutabilityとは、DNA配列中のCGのCが特異的にメチル化され、そのmethyl-Cが脱アミノ化してTへと変化することによって、CからTへの突然変異が高い率で生じる現象である。このタイプの突然変異は、隣の塩基の影響を受け、時間軸に対称でないという特徴がある。我々は、ヒト・チンパンジー・マウスのコード領域を比較し、コドン置換速度を調べた所、CpG hypermutabilityが引き起こす置換速度は他のものよりも約10倍大きいことがわかった。また、全体のコード置換のうち、およそ2割がCpG hypermutabilityに引き起こされていた。CpG hypermutabilityのために、codonの中でもCとGが並んでいるようなものは数が少ない。</p>

CpG hypermutabilityを考慮に入れた上で、codon頻度と発現量の相関を調べたところ、優位な相関が見られた。このことから2つのことがわかった。(1) ヒトやマウスでも、codonによって発現の効率が違う。(2) しかしその影響は、突然変異率の影響よりも小さい。先行研究で相関が見つかっていなかったのは突然変異率を考慮に入れていなかったためらしい。このことを利用すると、突然変異率の大きさと比較することで、codon頻度が翻訳効率に及ぼす影響を定量化することができる。

また、この研究の副産物として、たんぱく質を構成するアミノ酸が、長い時間の中で少しずつ置換されていく傾向をモデル化することができた。これもヒトとマウスの間での突然変異率を考慮に入れることで可能となった。

#### 主な成果

具体的な成果内容：

特許件数： 0

論文数： 1

口頭発表件数： 9

#### 研究成果に関する評価

##### 1 国内外における水準との対比

ヒトとマウスの発現量とcodon頻度の相関よりも突然変異率の効果のほうが大きいことを示したのは初めての研究である。

##### 2 実用化に向けた波及効果

codon使用頻度から発現量を推定することへつながるかもしれない

#### 残された課題と対応方針について

3' UTR領域の長さで発現量に相関があることがわかったが、3' UTRの中に発現量をコントロールしている配列があるかどうか探索したが、見つけることができなかった。これからより多くのデータをもちいることで、発現量をコントロールする要素を見つけ出したい。これが見つかれば、ヒトの遺伝子解析への応用が期待され、遺伝病の治療へとつながる可能性がある。

#### 代表的な設備名と仕様 [既存 (事業開始前) の設備含む]

JST負担による設備：大容量ワークステーションなど

地域負担による設備：大型コンピューターシステムなど