

論文内容の要約

氏名：田村 豪良

論文題名：Long-read sequence analysis for clustered genomic copy number aberrations revealed architectures of intricately intertwined rearrangements

(ロングリード解析による集簇したコピー数異常を呈した染色体の複雑構造異常の解明)

【背景】

染色体構造異常のほとんどは染色体の 1 か所のみ異常を認める単純な構造異常である。それらは、染色体の部分欠失、部分重複、逆位、均衡型転座、不均衡型転座である。このような構造異常により特定の遺伝子が障害を受けると様々な疾患の原因となる。例えば、22 番染色体長腕の 22q11.2 領域が欠失すると、精神発達遅滞、特徴的顔貌、胸腺低形成、心血管異常などを主兆候とする 22q11.2 欠失症候群 (DiGeorge 症候群)の原因となる。血液悪性疾患においては均衡型転座がしばしば疾患原因となる。例えば、慢性骨髄性白血病では腫瘍細胞に生じた t(9;22)相互転座によって *BCR-ABL* 融合遺伝子が形成され、チロシンキナーゼ活性が亢進することで腫瘍細胞が増殖する。このような染色体構造異常の検出は、長らく染色体構造異常解析の主な目的となってきた。1970 年代以降に G-band 法が開発されて以降多くの疾患が見出され、1990 年代に FISH 法が開発されてからはより関心のある染色体領域を精査することができるようになった。染色体上の様々な領域において、コピー数の増減に起因する種々の微細欠失/重複症候群が見出された。2000 年代にマイクロアレイ法 (CMA; Chromosomal microarray analysis) が臨床応用されてからは、それらが網羅的に、より解像度高く検出可能となった。そのため、現在 CMA は原因不明の先天性疾患に対する遺伝学的な第一選択の検査として推奨されている。

その折に、より複雑な染色体構造異常も認知されるようになった。複雑な構造異常とは、具体的には染色体の切断点が 3 か所以上に及び、単純に 1 か所の構造異常だけではなく複数の構造異常が同時に認められるものである。これらは FISH 法などの従来の手法を用いても解析されてきたが、技術的な限界、及び検査解像度の限界などから全容解明は困難であった。CMA も、コピー数異常のある領域は詳細に検出可能だが、均衡型転座や逆位などのコピー数が正常 (N=2) の領域は特定できない。しかし近年、次世代シーケンシング (NGS; Next-generation sequencing) を始めとする解析技術が発展し、1 個人のゲノム DNA の塩基配列決定がより簡便に施行可能となった。それにより従来は解析が困難であった複雑な染色体構造異常も詳細に構造決定が可能となり、現在ではその形成メカニズムについて関心が向けられている。

染色体構造異常が形成される際の根源的なメカニズムについて述べる。これらは主に、染色体構造異常が特定の領域に反復して認められる場合 (Recurrent rearrangement) と非特定の領域に認められる場合 (Non-recurrent rearrangement) に大別される。Recurrent rearrangement は、いわゆる構造異常の Hot spot であり、構造異常の始点と終点に存在するゲノム反復配列が発端となる。これは主に非対立遺伝子相同組み替え (NAHR; non-allelic homologous recombination) が関与する。NAHR は DNA 修復メカニズムの一つである Homologous recombination (HR) の一形態であり、HR には Rad51 や Replication protein A が関与している。一方、Non-recurrent rearrangement においてはゲノム反復配列は関与せず、染色体の不特定の箇所で起こる、① DNA 二重鎖切断 (DSB; double-strand break) の再結合 (Recombination-based mechanism)、もしくは② DNA 複製におけるエラー (Replication-based mechanism) が関与する。Recombination-based mechanism には、主として DNA 2 本鎖切断の修復メカニズムの一つである非相

同末端結合 (NHEJ; non-homologous end joining)が関与している。NHEJ においては Ku70/80 ヘテロ二量体, DNA リガーゼIVなどが関与する。Replication-based mechanism には FoSTeS (Fork stalling and template switching)や MMBIR (Microhomology-mediated break induced replication)などが機序として知られている。これらのメカニズムが基盤となり、染色体構造異常は発生する。

複雑な構造異常は、上述したメカニズムが複合して形成される。例えば、*PLP1* や *MECP2* の周辺領域においては、重複領域に内在するよう (凸型)に 3 倍体領域が存在する異常が報告されている。これらの 3 倍体領域は逆位になっており、duplication-inverted triplication-duplication (DUP-TRP/INV-DUP)と称されている。この構造異常を有する複数症例の染色体構造解析により、逆位の 3 倍体領域のテロメア側の断端の座位は同一 (Recurrent)であることが示されている。そしてその領域には重複領域のテロメア側の終点に存在するゲノム反復配列が対になるように存在していることが判明し、結果として① 対となる反復配列間での複製フォークの乗り換え、そしてその後は② NHEJ (再結合)、もしくは FoSTeS/MMBIR (複製鎖の乗換)のいずれか、の 2-Step mechanism により発生することが解明された。この他にも様々なパターンの複雑構造異常が報告されているが、最も複雑な構造異常として Chromoanagenesis が近年トピックスである。Chromoanagenesis は「Chromosome rebirth」を示す造語であり 2012 年に提唱された、より大規模で複雑、無秩序な構造異常である。これらは悪性腫瘍や原発性発達障害の患者の染色体から検出され、主に次世代シーケンス技術の進歩により解明できるようになった。主に 2 種類の発生機構が考えられており、それぞれ① Chromothripsis、② Chromoanasythesis と呼ばれている。Chromothripsis は、減数分裂中に単一/複数の染色体が核小体へと移行し複数の断片へと破碎 (Chromosome shattering)を受け、NHEJ などにより再結合することで形成される。Chromoanasythesis は複製エラーであり、FoSTeS/MMBIR により主に単一染色体の局所が複雑な再配置を来す。このように、近年ではより複雑な構造異常が悪性腫瘍や神経疾患を始めとする様々な疾患で検出され、その重要性が認識されている。そしてその発生メカニズムに対する検討も重ねられている。

今回、研究を実施した施設において、染色体同腕にコピー数異常が集簇して縦列に 2 か所認められた、コピー数異常に共通の特徴を持つ希少な症例が 4 例認められた。本解析を実施した施設においては、長年に渡り染色体構造異常症が疑われる患者に対して CMA を実施しており、全国からの症例相談を受けており約 500 症例を超える CMA での検査経験がある。4 症例のコピー数異常はその中から抽出され、他に類似症例の経験はなかった。また文献においても、このような縦列したコピー数異常についての染色体構造解析が詳細に記載されている報告はない。そのため、これらの染色体構造を詳細に解析し、2 か所の欠失または重複が単純に存在しているか、それとも複雑構造異常の一端となっているかを明らかにし、その形成機序について共通したメカニズムが見出せるか、また患者の症候を説明しうる遺伝子が認められるか、という点について検討するため本研究を実施した。

【目的】

- ① 主要評価項目; CMA で縦列した 2 か所のコピー数異常が認められた 4 症例の染色体構造を詳細に解析する。
- ② 副次的評価項目; 染色体離断接続点の塩基配列を確認し、染色体構造異常を来したメカニズムを検討する。染色体離断切断点周囲のゲノム構造を UCSC Genome browser を用いて確認し、染色体構造異常の誘因となるような反復配列が認められるか確認する。染色体構造異常によって影響を受けた遺伝子を精査し、患者の疾患表現型との対比を検討する (論文では Supplemental information にまとめた)。

【対象と方法】

① 対象及び研究デザイン; 本研究は珍しいコピー数異常が認められた染色体の詳細な構造異常解析, 及びそれらの形成メカニズムの検討のため施行された。生下時より原因不明の神経発達障害を来し, 遺伝学的原因究明のため紹介された患者のうち, 染色体同腕に縦列し 2 か所以上のコピー数異常が認められた 4 症例を対象とした。本研究は東京女子医科大学の臨床研究審査委員会 (承認番号: 384C) と日本大学医学部附属板橋病院の臨床研究倫理審査委員会 (承認番号: RK-220308-2) の承認の下に行なわれた。

② ゲノムコピー数解析及び全ゲノム解析; 家族よりインフォームド・コンセントを得たのち, 患者の血液を採取し DNA を抽出した。Agilent 社のマイクロアレイシステムを用いてコピー数解析を行い, 上述のコピー数異常を認めた 4 症例を抽出した。Pacific Biosciences 社のロングリードシーケンサー (LRS) Sequel を用いて全ゲノム解析を行なった。全ゲノム解析を施行した理由としては, 染色体接続点の移行先が想定外の染色体である場合, 解析が不可能になる可能性があるためである。得られたデータは Integrative Genomics Viewer を用いて視覚的に解析し, 患者の染色体構造を想定した。

③ 離断接続点 (Breakpoint junction; BJ) の解析; 染色体同士が移行している部分 (離断接続点) の正確な座位同定, 及び接続点の塩基配列を調べるため Sanger 法を実施した。Sanger 法を施行した理由としては, LRS 解析では塩基単位での配列決定に誤差が生じることが多いため, 詳細な確認のためである。また接続点の塩基配列が構造異常発生のメカニズム推定に重要なためである。

[補足説明: 離断接続点の塩基配列は発生メカニズム推定に重要である。即ち, 移行した染色体同士に微細な相同配列が認められる場合, その配列は Microhomology と呼称されており NHEJ, または FoSTeS/MMBIR のいずれかが示唆される。移行した染色体のどちらにもない塩基挿入が認められる場合, FoSTeS/MMBIR が主に示唆される。何の介在配列もなくそのまま接続している場合 (Blunt End) は, NHEJ が示唆される。]

④ 患者の臨床症状との対比; 染色体構造異常によって欠失や重複, 離断などを受けた遺伝子を UCSC Genome browser を用いて検索・評価した。染色体構造異常が認められた領域について過去文献を精査し評価も行った。患者の臨床症状については質問票を作成し, 紹介元の医療機関に聴取した。

【結果】

① 症例 1 の臨床経過及び解析結果:

症例 1 は 3 歳男児, 在胎 38 週 2 日で異常なく出生した。3 歳時に主に低身長と精神発達遅滞のため紹介され, 身長は 84.6 cm (-3.8 SD) と低身長であり発語は 2 語文のみで軽度の知的障害を認めた。体表奇形は目立たないが, 左眼に弱視, 乱視を認め眼科診察では視神経コロポーマを認めた。頭部 MRI においては視神経萎縮が認められた。G-band 法では t(11;22)(q23.3;q11.2) 均衡型転座を認めたが, 生来健常な父にも同じ転座が認められ病的意義は低いと考えられていた。CMA では 10 番染色体長腕に縦列して 2 か所 [del(10q24.2q24.31) と del(10q24.33q25.1)] の微細欠失を認めた。

LRS 解析では, 10 番染色体長腕は計 3 つの BJ で接続された 4 segments に分けられていることが判った。2 つの欠失の中間領域は逆位になっていた。BJ の塩基配列は BJ1, 3 は Microhomology (2~3 塩基対), BJ2 は Blunt End であった。

本症例の染色体欠失領域には約 50 個の遺伝子が存在した。OMIM (Online Mendelian Inheritance In Man) を参照すると, うち 16 個は疾患原因遺伝子と報告されていた。うち PAX2 (Paired Box 2) 遺伝子は腎, 眼, 耳などの発生に重要な転写因子をコードし, 片アレルの機能喪失変異により Papillorenal syndrome の原因と報告されていた。視神経症状は PAX2 欠失が原因と考えられた。CNVM2 (Cyclin

M2)遺伝子は中枢神経発達に重要であり、機能喪失によって脳の発達が阻害されることがゼブラフィッシュで確かめられている。また片アレルの病的変異により神経発達障害を来した症例報告も認められた。本症例の神経発達障害の原因として *CNVM2* 欠失の関与が示唆された。低身長の原因と報告される遺伝子は認められなかった。なお、G-band 法で認められた均衡型転座の座位に遺伝子は認められなかった。

② 症例2の臨床経過及び解析結果:

症例2は6歳女児、在胎40週5日で出生した。生下時より小頭症、丸顔、眼間乖離、瞼裂斜下、耳介低位、口蓋裂、小顎症、巻き毛などの身体的特徴を認めた。幼少期は摂食障害のためチューブ栄養だった。自律歩行は4歳と遅れ、6歳で数個の単語しか話せず、現在は自傷行為も認められる。3歳時の頭部MRIでは異常を認めなかった。G-band 法では $t(1;2)(q31;q36), inv(2)(q31q36)$ であり1番/2番染色体の均衡型転座と2番染色体の逆位領域が認められた。健常な両親には均衡型転座は認められなかった。CMAでは2番染色体長腕に縦列して欠失 [$del(2q31.2q32.1)$ と $del(2q32.3q33.1)$]が認められた。

LRS解析では、CMAで認められた2番染色体の2か所の欠失よりテロメア側に更に小さな欠失領域が1か所確認された。2番目と3番目の欠失の中間領域は逆位であった。G-band法で認められた逆位はこの領域と考えられる。その先は1番染色体長腕(1q31.2)へと転座し、これもG-band法と一致した。欠失、逆位、転座を含む複雑構造異常であった。BJの塩基配列は全て Microhomology (1~5塩基対)であった。

これら3か所の染色体欠失領域には合計して42個の遺伝子が存在した。そのうち12個は疾患原因遺伝子として OMIM に記載されていた。2番目の欠失領域に認められた *SATB2* (Special AT-rich sequence-binding protein 2) 遺伝子は、転写制御に関与する核マトリクス結合蛋白であり、片アレル欠失により Glass 症候群の原因となることが知られている。本症例は臨床症候も合致し Glass 症候群と診断できた。

③ 症例3の臨床経過及び解析結果:

症例3は生後6か月の女児、在胎週数40週2日で異常なく出生した。生後6か月で体重5.4kg (-2.6SD)と体重増加不良を認めた。生下時より、隆起した前頭部、外側眉毛の増生、眼間乖離、広い鼻梁、広い鼻尖、広い人中、口角下垂、下顎後退などの特徴的顔貌を認めた。心エコーで心房中隔欠損症、心室中隔欠損症を、また第5指に斜指症も認めた。生後6か月でようやく顎定が確認でき、軽度の運動発達遅滞が示唆されている。G-band法では正常であった。CMAでは11番染色体長腕に微細重複 [$dup(11q13.2q13.3)$]と微細欠失 [$del(11q13.3q13.5)$]が縦列して認められた。

LRS解析では、微細重複の領域は折り返すよう逆位であった。そして欠失を挟むようにテロメア側へと接続していた。BJには長い塩基挿入(5~65塩基対)が認められ、ヒトゲノムに存在しない配列であった。

本症例ではコピー数異常領域に計73個の遺伝子が存在し、うち20個が疾患原因遺伝子と報告されていた。本例とOverlapする領域に欠失を認める症例が過去に2例報告されていた。いずれも神経発達障害を認め、広い鼻梁、眼間乖離、口角下垂、下顎後退、斜指症などの共通する症候を認めた。本症例の欠失領域には *SHANK2* (SH3 and multiple ankyrin repeat domains 2) 遺伝子が認められた。SHANK2はグルタミン酸作動性シナプスの足場蛋白の1つであり、片アレル性の欠失や機能喪失変異によって自閉症や知的障害きたす報告がある。ファミリー蛋白の一つである *SHANK3* の欠失は重度の発達障害を来すと報告されている。本症例の発達遅滞は *SHANK2* 欠失に起因した可能性が考えられた。また、この染色体領域は嗅覚受容体遺伝子ファミリーが集合しており、それらの相同配列を介した NAHR の好発領域

とする報告もあった。しかし、本例のBJはいずれもこれらの配列に位置せず患者固有であった。

④ 症例4の臨床経過及び解析結果:

症例4は8歳女児、在胎週数39週3日で異常なく出生した。生下時から肺動脈低形成、心室中隔欠損症があり兩大血管右室起始症と診断され、生後1か月、6か月時にBlalock-Taussigシャント術、2歳時に三尖弁輪縫縮術を施行された。生下時より、短頭症、眼角隔離、広い鼻梁などの特徴的顔貌を認めた。精神運動発達遅滞は顕著であり、ひとり座りは18か月、自律歩行は5歳、現在も発語は確認できていない。また自傷傾向や自閉傾向も認められている。CMAでは9番染色体長腕に微細重複 [dup(9q34.3)]と微細欠失[del(9q34.3)]が縦列して認められた。

LRS解析では、縦列した重複・欠失の更にテロメア側に微細重複が1か所確認された。セントロメア側の重複は単純な繰り返しであったが、テロメア側の重複は逆位になっていた。BJの塩基配列では、BJ1にヒトゲノムにはない長い挿入(11塩基対)が認められ、BJ2は1塩基対の挿入、BJ3はBlunt endであった。

コピー数異常領域には計49個の遺伝子が存在し、うち計8個は疾患原因遺伝子として報告されていた。欠失領域に認められた*NOTCH1*(Notch receptor 1)遺伝子は、膜貫通受容体の一つであり血管形成や心臓発生などの重要な制御因子と知られている。そして片アレル変異によりFallot四徴症などを含む種々の原発性心疾患の原因となると報告されている。同時に、この遺伝子のハプロ不全によりAdams-Oliver症候群の原因になることが報告されている。Adams-Oliver症候群は、頭皮欠損、四肢末端欠損を主兆候とするが種々の原発性心疾患のほか、発達障害も合併する。そのため本症例の一部の臨床症状は*NOTCH1*欠失が原因とも考えられた。また本例で異常が認められた9q34領域は、欠失により9q34微細欠失症候群(Kleefstra症候群)が発症することが知られる。本症候群の臨床像は、重度の発達障害、短頭症や眼角隔離などの特徴的顔貌、低緊張症、行動障害、そして種々の先天性心疾患であり、本症例の臨床兆候はむしろ本症候群とよく合致していた。しかし本症候群は*EHMT1*(Euchromatic histone methyltransferase 1)遺伝子のハプロ不全が原因であり、本症例の欠失領域は*EHMT1*の0.6 Mb上流に位置していた。可能性として、位置効果によって遺伝子発現に影響を与えたとも考えられた。

【考察】

1. 複雑染色体構造異常の解明

本検討では、CMAで特徴的なコピー数異常が見出された4症例の染色体構造を詳細に解明することができた。4症例はコピー数異常が近位に集簇し縦列している点で類似の特徴を有していた。同様のコピー数異常に焦点を充てて染色体構造の検討を試みた報告はない。4症例で認められた、集簇したコピー数異常領域はいずれも単純に存在していたのではなく、全てが複雑構造異常の一端を成していた。周囲には全例で逆位領域が認められ、一部の症例は均衡型転座も含有した大規模な構造異常であった。これらから言えることは、2か所以上のコピー数異常が集簇して認められる場合、より複雑な構造異常を考慮すべきということである。近年、このように詳細にゲノム解析が可能になったことで、従来手法においては判別できなかった異常までを確認できるようになった。文献では、従来手法では単純かつ均衡型の構造異常が検出されていた患者からも、より複雑な構造異常が想定よりも多く認められることが報告されている。そして複雑構造異常は、遺伝子の欠失・重複・離断といった直接の修飾を越え、より多くの遺伝子の発現調節領域を複雑に再構成させ、位置効果を介してより広範囲の遺伝子発現に影響を与えることが示唆されている。実際に、本検討の症例4の臨床症状は直接障害を受けていない*EHMT1*遺伝子からの影響がより考えられた。また、自験例(副論文として提出した報告症例)ではG-band法で均衡型

転座と考えられていた症例の転座切断点に複雑構造異常が明らかになり、患者の臨床症状はその下流に存在する *MEF2C* の機能不全に合致していた。従って、染色体構造異常を詳細に解析することで、未だ詳細に判明していない遺伝子発現調節のネットワークを明らかにし、患者の疾患原因の更なる究明に繋がられる可能性がある。そして病態の深い理解は、新たな治療法開発に発展する可能性も内在している。

2. 複雑構造異常に至ったメカニズムの検討

本検討では、サブ解析として離断切断点周囲のゲノム構造を UCSC Genome browser を用いて精査し、いわゆる Recurrent rearrangement を惹起するようなゲノム反復配列が周囲に認められるかを精査した。結果、一部の切断点は LINE (Long interspersed element) や SINE (Short interspersed element) などのトランスポゾン反復配列中に認められたものの、いずれも配列の相同性は低く、反復配列を介した形成機序は否定的であった。症例 3 の異常領域は NAHR の好発部位と報告されていたが本例では考えづかった。4 症例とも構造異常の一部が逆位となっていたことから DUP-TRP/INV-DUP と類似であることも検討したが、DUP-TRP/INV-DUP の原因となるような、対に配置されたゲノム反復配列も認めなかった。よって全例が Non-recurrent rearrangement で、不特定の領域に生じた構造異常と考えられた。

他の複雑染色体構造異常として報告されているパターンに合致したのものもなく、Chromoanagenesis を考えた。症例 1, 2 の構造異常は、コピー数は欠失 (N=1) または正常 (N=2) のみで構成されており、重複は含まれなかった。離断接続点は Microhomology か Blunt end で接続されており、これらは NHEJ により DNA の再結合によって接続されたことを示唆させた。よってこれらの全体像は Chromothripsis と合致した。Microhomology は FoSTeS/MMBIR などの DNA 複製エラーにおいても切断点に認められ、その場合 Chromoanagenesis も否定できなかったが、同機序では重複 (N=3) 領域も多く認められることから、より Chromothripsis が考えられた。

症例 3, 4 では、離断接続点にヒトゲノムには認められない長い塩基挿入 (最長で 65 塩基対) が認められたことが特徴的であった。文献検索を行ったところ、長い塩基挿入は Polymerase θ (Pol θ) を介する Alternative-NHEJ (Alt-NHEJ) に起因するとの報告を複数認めた。Alt-NHEJ は error-prone であり、DSB 修復においては NHEJ より下位の機序である。種々の影響により HR や NHEJ などの中心的な修復機構が作用できない場合に動員される。Pol θ は、DSB 修復の際に切断された箇所の ss-DNA 断端を橋渡し、helicase domain の作用により鋳型に依存せずニヌクレオチドの付加を行いながら切断箇所を修復する。そのためにヒトゲノムには認められない塩基挿入が付加される。近年では悪性腫瘍だけではなく、神経発達障害患者の染色体構造異常からも Pol θ の報告がされている。また神経疾患の患者において Pol θ による Chromoanagenesis の症例報告も認められている。今回の検討では、症例 3, 4 の 2 例において Pol θ が関与した可能性が考えられ、実際に悪性新生物のみならず原発性神経発達障害の患者の生殖細胞系列においても認められることが裏付けられた。様々な悪性腫瘍においては、HR や NHEJ に必要な蛋白に変異が蓄積し、Pol θ への依存性が高まることで発癌に寄与していることが報告されている。例えば、HR において重要な遺伝子として *BRCA1*, *BRCA2* が知られており、*BRCA1*, *BRCA2* の機能喪失は遺伝性乳がん卵巣がん症候群の原因であることが知られている。HR は鋳型 DNA を用いる修復機構であり、最も正確に DSB 修復を行う。HR と Alt-NHEJ は antagonistic な関係性があり、Alt-NHEJ を不活化させることで HR の動員効率が増加するとも報告されている。具体的には、*BRCA1* 変異のある細胞系列で Pol θ を失活させると、細胞の生存に有利であることが報告されている。このように悪性疾患においては、Pol θ を抑制することによる治療への発展が示唆されている。染色体構造異常の発生メカニズムを探索することで、新たな治療プロセスへの道標となっていく可能性があると考えられた。

【結論】

本検討では, LRS や CMA, Sanger 法を組み合わせることにより, 4 例の染色体構造異常を詳細に確認することができた。CMA にて集簇して 2 か所以上のコピー数異常が確認される場合, 更なる複雑構造異常を考慮する必要がある。