

文部科学記者会・科学記者会

厚生労働記者会・厚生日比谷クラブ 同時発表

2022 年 4 月 22 日

横浜市立大学

神経発達障害の原因遺伝子の大規模な同定に成功

横浜市立大学大学院医学研究科 遺伝学の濱中耕平助教、水口剛准教授、松本直通教授、国立国際医療研究センター研究所 疾患ゲノム研究部 三宅紀子部長、理化学研究所 脳神経科学研究センター 分子精神病理研究チーム 高田篤チームリーダーらの研究グループは、全エクソーム解析を用いて多数の神経発達障害症例で突然変異を解析することにより、新規の原因遺伝子の網羅的な同定に成功しました。

この研究成果は、昭和大学医学部小児科学講座 加藤光広教授、大阪母子医療センター 岡本伸彦研究所長らとの共同研究によるものです。

本研究成果は、科学雑誌「*Genome Medicine*」に掲載されます。（日本時間 2022 年 4 月 26 日 午前 9 時）

研究成果のポイント

- 神経発達障害を持つ 41,165 家系において、全エクソーム解析^{*1}を用いてコピー数変化^{*2}を含む突然変異^{*3}を網羅的に解析した。
- 突然変異が統計的に有意に多く見られる 380 遺伝子を同定し（偽発見率 5%）、内 52 個は新規の遺伝子であった。
- この 52 個の内、3 個はコピー数変化を考慮したことが、その統計的有意性に貢献していた。
- この 52 個の内、既知の原因遺伝子との類似性を指標にベイズ推定^{*4}を用いて、真の原因である確率が 90%以上の遺伝子が 11 個抽出された。

研究背景

神経発達障害は、知的障害群や自閉スペクトラム症などを含む疾患群です。神経発達障害の原因遺伝子は、患者集団で一塩基置換^{*5}などの突然変異がその確率から予想されるより多く見られる遺伝子を探索することで、これまでに多数同定されてきました。しかし、症例数を増すことで発見できる原因遺伝子がまだ多数残っていることも予測されていました。また、遺伝子を巻き込むコピー数変化についてはその突然変異確率が知られていないため、過去の研究ではコピー数変化が考慮されておらず、原因遺伝子を効率よく発見できていない可能性がありました。

研究内容

本研究では、神経発達障害のトリオ^{*6}の突然変異のデータを可能な限り集めることを試み、一塩基置換や小さな挿入欠失^{*7}の解析用データとして、グループ独自のデータ（1,317

報道解禁：日本時間 2022 年 4 月 26 日 午前 9 時・26 日夕刊

家系)、DDD31k^{*8} データ (31,058 家系)、denovo-db^{*9} データ (8,790 家系) (合わせて 41,165 家系)、またコピー数変化の解析用データとして、グループ独自のデータ (1,298 家系)、Simons Simplex Collection^{*10} データ (2,377 家系) (合わせて 3,675 家系) が集積しました。次に、遺伝子ごとのコピー数突然変異の確率を明らかにしようと試みました (図 1)。この確率は各遺伝子において集団で観察されるコピー数変化の種類の数と理論的に比例することが知られています。そこで、全遺伝子におけるコピー数突然変異の総確率を、各遺伝子におけるコピー数変化の数で除算することで、各遺伝子におけるコピー数突然変異の確率を明らかにしました。

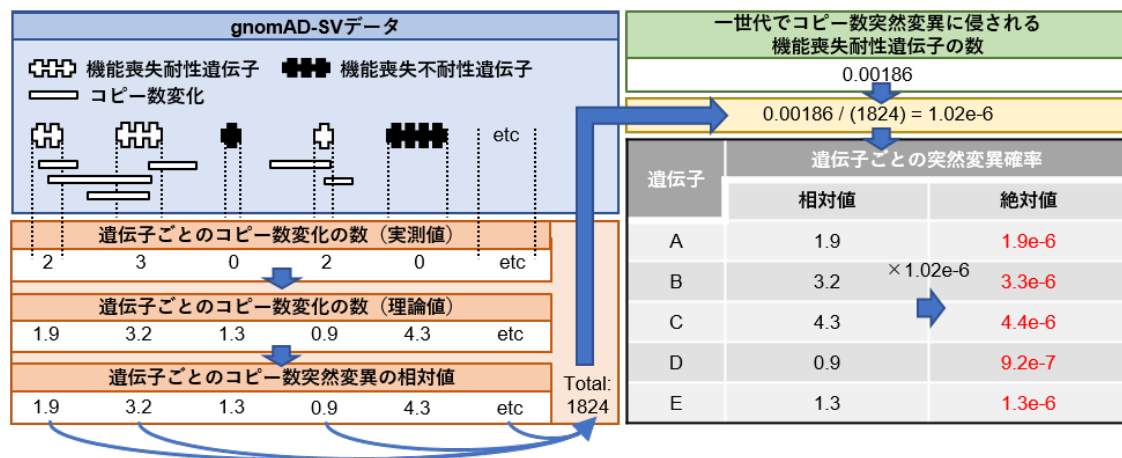


図 1 コピー数突然変異の確率の計算

一般集団で見られたコピー数変化を収載している gnomAD-SV^{*11} を基にして、各遺伝子のコピー数変化の数からコピー数突然変異の相対的な確率を計算する。次に、一世代でコピー数突然変異に侵される遺伝子の数をこれらの相対値を基に分配することで、各遺伝子のコピー数突然変異の確率が求まる。

次に、この確率と以前の研究で計算された一塩基置換などの突然変異確率とを合計することで予想される突然変異の数と、上記の神経発達障害の家系で見られた突然変異の数を比較しました。結果、突然変異が統計的に有意に多く見られる 380 遺伝子が同定されました (誤検出率 5%)。この 380 個の内の 52 個は過去に神経発達障害との関連が報告されておらず、新規の原因遺伝子候補と考えられました。この 52 個のうち、3 個 (*GLTSCR1*、*MARK2*、*UBR3*) はコピー数変化がその統計的有意性に貢献しており、コピー数変化を解析に含めることの重要性が示唆されました (図 2)。

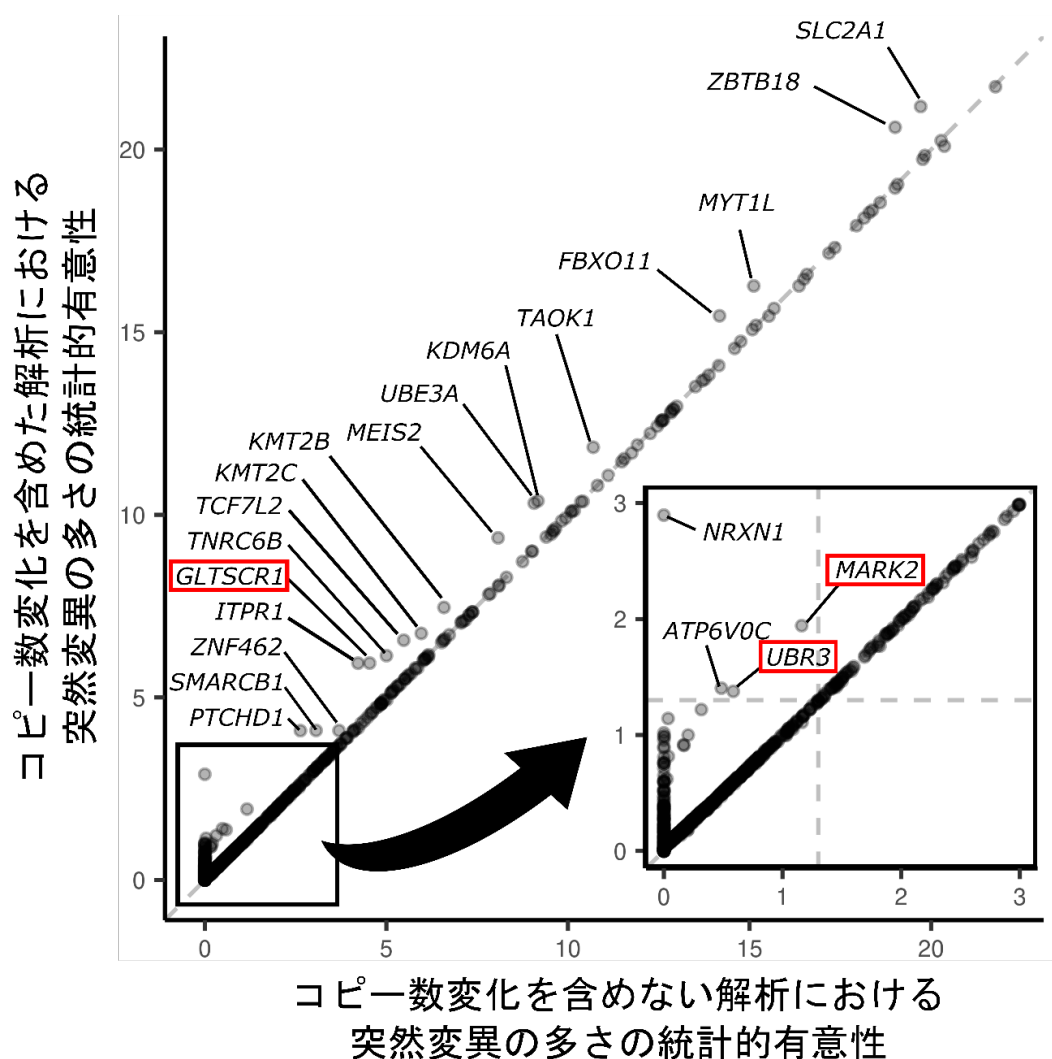


図2 コピー数変化を含めた場合と含めない場合の統計的有意性の比較

コピー数変化を含めた場合（縦軸）と含めない（横軸）場合の、各遺伝子における突然変異の多さの統計的有意性の比較。丸印が各遺伝子を示す。対角線より左上に位置する遺伝子は、コピー数変化が同定されたため、コピー数変化を考慮した解析の方が、統計的有意性が高くなっている。特に赤枠で囲った3遺伝子は新規原因遺伝子候補であることを示す。

次に、これらの新規遺伝子候補について、神経発達障害の既知の原因遺伝子との機能的な類似性をもとに、さらに絞り込むことを試みました。まず、既知の原因遺伝子とそれ以外の遺伝子を教師データとして、ある遺伝子が神経発達障害の原因遺伝子に似ているかどうかを pLI ^{*12} や遺伝子オントロジーターム^{*13} などの特徴量^{*14} から計算する深層学習^{*15} モデルを訓練しました。このモデルを我々が同定した新規遺伝子候補に適用すると、陰性対照（神経発達障害の原因でない遺伝子群）よりも有意に高いスコアを示し、新規遺伝子候補が既知

報道解禁：日本時間 2022 年 4 月 26 日午前 9 時・26 日夕刊

の原因遺伝子と機能的に類似していることが確認されました。次に、我々の新規候補遺伝子が真の原因遺伝子である確率を、このスコアを指標にして事前確率（ $63\% = (52 - 380 \times 0.05) / 52$ ）からベイズ推定を用いて計算しました。その結果、真の原因遺伝子である事後確率が 90%以上を示す 11 個の有力な原因遺伝子候補を抽出することに成功しました（*HDAC2*、*SUPT16H*、*HECTD4*、*CHD5*、*XPO1*、*GSK3B*、*NLGN2*、*ADGRB1*、*CTR9*、*BRD3*、*MARK2*）。

今後の展開

本研究により、神経発達障害の原因解明には一塩基置換や小さな挿入欠失に加えてコピー数変化を解析対象にする重要性が明らかになりました。この手法を用いて明らかにした原因候補遺伝子群は、神経発達障害の遺伝子診断や病態メカニズムの解明、医学的管理法や治療法の開発に近い将来大きく寄与することが期待されます。

研究費

本研究は、国立研究開発法人日本医療研究開発機構（AMED）の難治性疾患実用化研究事業「新技術を用いた難治性疾患の高精度診断法の開発（研究代表者：松本直通）」、厚生労働省、日本学術振興会、武田科学振興財団、横浜市立大学の支援により実施されました。

論文情報

タイトル： Large-scale discovery of novel neurodevelopmental disorder-related genes through a unified analysis of single-nucleotide and copy number variants

著者： Kohei Hamanaka, Noriko Miyake, Takeshi Mizuguchi, Satoko Miyatake, Yuri Uchiyama, Naomi Tsuchida, Futoshi Sekiguchi, Satomi Mitsuhashi, Yoshinori Tsurusaki, Mitsuko Nakashima, Hirotomo Saitsu, Kohei Yamada, Masamune Sakamoto, Hiromi Fukuda, Sachiko Otori, Ken Saida, Toshiyuki Itai, Yoshiteru Azuma, Eriko Koshimizu, Atsushi Fujita, Biray Erturk, Yoko Hiraki, Gaik-Siew Ch'ng, Mitsuhiro Kato, Nobuhiko Okamoto, Atsushi Takata, and Naomichi Matsumoto

掲載雑誌： *Genome Medicine*

DOI： 10.1186/s13073-022-01042-w



横浜市立大学は、様々な取り組みを通じてSDGsの達成を目指します。



用語説明

- *1 全エクソーム解析：遺伝子のエクソン領域（タンパク質の配列を決定する領域）を網羅的に濃縮後、次世代シーケンサー（高速にゲノムの配列を読むことができる機器）を用いて塩基配列を決定する方法。
- *2 コピー数変化：各遺伝子はゲノム上に 2 コピーずつ（父・母からそれぞれ 1 コピーずつ由来）存在するが、その全体または一部の数が変化（増減）すること。
- *3 突然変異：両親には見られず子供にのみ新規に生じた変異。
- *4 ベイズ推定：観測結果をもとに、推定したい事柄の確率を推論すること。医療検査にも応用され、例えば人口におけるある疾患を持つ人の割合を 0.1%とする（事前確率）。その疾患に対して感度 99%、特異度 97%の検査をある人に行い陽性と判定された場合、その人が疾患を持つ確率は 3.2%（ $= \text{感度 } 99\% \times \text{事前確率 } 0.1\% \div (\text{感度 } 99\% \times \text{事前確率 } 0.1\% + (1 - \text{事前確率 } 0.1\%) \times (1 - \text{特異度 } 97\%))$ ）で判定される（事後確率）。
- *5 一塩基置換：ゲノム上のある塩基が他の塩基に変わること。
- *6 トリオ：患児とその両親の 3 人。
- *7 小さな挿入欠失：ゲノム上のある位置に 1~数十程度の塩基が挿入されたり欠失したりすること。
- *8 DDD31k：イギリスで行われた Deciphering Developmental Disorders（DDD）と呼ばれる研究であり、神経発達障害を持つ約 31,000 家系における突然変異が公開されている。
- *9 denovo-db：論文で公開された突然変異を収載しているデータベース。
- *10 Simons Simplex Collection：自閉スペクトラム症を持つ 2,000 以上の家系の次世代シーケンサー解析の公開データ。
- *11 gnomAD-SV：一万人規模の一般集団で観察された、コピー数変化を含むゲノム構造異常を網羅したデータベース。
- *12 pLI：probability of being loss-of-function intolerant の略で、ある遺伝子の変異が一般集団でどの程度自然淘汰されているか（認められないか）を示す指標。
- *13 遺伝子オントロジーターム：遺伝子の機能を生物学的プロセスや分子機能などの観点から記述した語彙。
- *14 特徴量：対象の特徴を表し、予測の手掛かりとなる変数。
- *15 深層学習：人間の神経細胞の仕組みを再現した多層構造を用いた機械学習の手法。