

## 学位論文の要旨

### A pipeline for complete characterization of complex germline rearrangements from long DNA reads

(ロングリードシーケンスを用いたヒト染色体構造異常解析のパイプライン構築と複雑な構造異常の解明)

2022年3月

尾堀 佐知子

横浜市立大学大学院医学研究科 産婦人科学 生殖生育病態医学

横浜市立大学大学院医学研究科 医科学専攻 遺伝学

(研究指導教員：松本 直通 教授)

(指導教員：宮城 悦子 教授)

## 学位論文の要旨

# A pipeline for complete characterization of complex germline rearrangements from long DNA reads

(ロングリードシーケンスを用いたヒト染色体構造異常解析のパイプライン構築と複雑な構造異常の解明)

<https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-020-00762-1>

## 1. 序論

ロングリードシーケンス (LRS) 解析は、近年技術的進展がめざましく、現在ではショートリードシーケンス (SRS) 解析の短所を補う技術となり、疾患解明に有効な解析手法となっている。

LRS の特徴としては、PCR による増幅を行わないため、GC 含有率の高い領域の解析が可能であることや、10 Kb 以上の長い塩基配列を決定することができることから、リピート配列を含むリードを正確に参照ゲノム配列にアライメントでき、さらにリピート回数の同定が可能である。また、染色体構造異常解析では、切断点の並び方を再構築する必要があり、ロングリードが複数のゲノム切断点を含んでいることは、複雑な構造異常の全体像の正確な解釈を可能にする。しかし、染色体構造変化は健常人にも認められる病原性のない構造多型といわれるものから、疾患の原因となりうる数十箇所の染色体の切断点を認める複雑なクロモソプシス (染色体破碎現象) まで多岐にわたり、患者の塩基配列やその切断点の LRS データを手に入れても、その解釈が困難な場合がある。

今回、患者の LRS データから染色体構造変化を抽出し、コントロール集団と共通の構造変化を差し引いていくことで、患者固有の構造変化を絞り込むという、新しい手法を用いた解析パイプラインを構築した。この解析パイプラインの有用性を検証するために、染色体核型分析やマイクロアレイ解析、SRS 解析等の既存の遺伝学的検査方法では全ての切断点を検出できていなかった、染色体構造異常症例の解析を行った。

また、当研究室でトリオ (患者と両親) ベースの全エクソーム解析 (WES) を行い原因遺伝子の同定ができなかった症例について、実際にこの解析パイプラインを用いた LRS 解析を行い、疾患原因を検索した。

## 2. 実験材料と方法

対象：

- 1) 当研究室で過去に解析された染色体構造異常を持つ症例のうち、G-band 法、Fluorescence in situ hybridization (FISH) 法、サザンブロット法などを用いて、切断点と同定されている 1 例。G-band 法、FISH 法、マイクロアレイ解析、SRS による全ゲノム解析などを用いても、一部の切断点しか同定されていない 3 症例。
- 2) 過去にトリオベースの全エクソーム解析 (WES) を行っているが、疾患原因が未解決である神経発達障害を中心とした先天異常症 10 例。

本研究は横浜市立大学医学部の倫理委員会の承認を得て行われた（承認番号 A19080001）。全例で患者、ご家族に対し研究参加についての説明が行われ書面による同意を得られた。

方法：

ロングリードシーケンサー (PromethION, Oxford Nanopore Technologies 社) から得られた全ゲノム配列データを、配列アライメントのデータ解析ツールである “LAST” を用いて、ヒト参照ゲノム配列 (GRCh38) にアライメントを行った。次に共同開発した “dnarrage” というソフトウェアを用いて、染色体構造変化を有するリードを抽出し、33 人のコントロール集団と共通して認めるリードを差し引いて、患者固有の構造変化であるリードを絞りこんだ。残った構造変化のリードをもとに “lmassemble” でコンセンサス配列を作成し、複雑な構造変化であった場合には “dnarrange-link” でゲノムの切断点を、矛盾なく並べることで、構造変化の全体像を明らかにすることを試みた。

また、“dnarrange” で同定された切断点を、サンガー法を用いて確認した。

発見した構造異常が、遺伝子発現に変化を与える可能性がある場合には、患者のリンパ芽球様細胞株より抽出した RNA を用いて、遺伝子発現量を測定した。

### 3. 結果

- 1) 過去に他手法で解析した 4 症例で検証を行い、2 症例で既存の遺伝学的検査方法では同定されていなかった、微細な欠失、複雑な染色体構造変化／異常が塩基レベルまで同定された。しかし 1 症例で、9 番染色体の切断点が同定できなかった。
- 2) 1 例で *MBD5* 遺伝子のノンコーディングエクソン領域を含む 97 Kb の欠失を同定した。患者の RNA を使用して *MBD5* 遺伝子の発現低下を確認した。

### 4. 考察

- 1) 9 番染色体の切断点が同定できなかった理由として、切断点がセントロメア領域に存在し、ヒト参照ゲノム配列 (GRCh38) で配列の情報が得られない場所であったため、リードのアライメントができなかったことが考えられた。GRCh38 に正確に存在しないゲノム領域が関わる構造変化の同定は困難な場合があり、ヒト参照ゲノム配列の正確性や網羅性が向上することで、今後克服できる可能性があると考えられる。
- 2) WES 解析で疾患原因が同定できなかった症例で病因遺伝子の同定ができた。さらに診断率を向上させるため、LRS を適用する症例の選択や他の遺伝学的検査との組み合わせ等の検討も必要と考えられた。

今回検証した解析手法が、ヒト参照ゲノム配列に存在する領域の染色体構造変化を正確に検知していることを確認できたことから、本解析パイプラインは、疾患原因となる染色体構造異常の検出に有用であることが示された。

キーワード：ロングリードシーケンス (LRS) 解析，染色体構造変化／異常，ヒト参照ゲノム配列

## 引用文献

Aneichyk, T., Hendriks, W. T., Yadav, R., Shin, D., Gao, D., Vaine, C. A., Collins, R. L., Domingo, A., Currall, B., Stortchevoi, A., Mulhaupt-Buell, T., Penney, E. B., Cruz, L., Dhakal, J., Brand, H., Hanscom, C., Antolik, C., Dy, M., Ragavendran, A., Underwood, J., Talkowski, M. E. (2018). Dissecting the Causal Mechanism of X-Linked Dystonia-Parkinsonism by Integrating Genome and Transcriptome Assembly. *Cell*, 172(5), 897–909.e21. <https://doi.org/10.1016/j.cell.2018.02.011>

Baslan, T., Kovaka, S., Sedlazeck, F. J., Zhang, Y., Wappel, R., Tian, S., Lowe, S. W., Goodwin, S., and Schatz, M. C. (2021). High resolution copy number inference in cancer using short-molecule nanopore sequencing. *Nucleic acids research*, gkab812. Advance online publication. <https://doi.org/10.1093/nar/gkab812>

Bertelsen, B., Nazaryan-Petersen, L., Sun, W., Mehrjouy, M. M., Xie, G., Chen, W., Hjermind, L. E., Taschner, P. E., and Tümer, Z. (2016). A germline chromothripsis event stably segregating in 11 individuals through three generations. *Genetics in medicine : official journal of the American College of Medical Genetics*, 18(5), 494–500. <https://doi.org/10.1038/gim.2015.112>

Beyter, D., Ingimundardottir, H., Oddsson, A., Eggertsson, H. P., Bjornsson, E., Jonsson, H., Atlason, B. A., Kristmundsdottir, S., Mehringer, S., Hardarson, M. T., Gudjonsson, S. A., Magnusdottir, D. N., Jonasdottir, A., Jonasdottir, A., Kristjansson, R. P., Sverrisson, S. T., Holley, G., Palsson, G., Stefansson, O. A., Eyjolfsson, G., ... Stefansson, K. (2021). Long-read sequencing of 3,622 Icelanders provides insight into the role of structural variants in human diseases and other traits. *Nature genetics*, 53(6), 779–786. <https://doi.org/10.1038/s41588-021-00865-4>

Bharti, R., and Grimm, D. G. (2021). Current challenges and best-practice protocols for microbiome analysis. *Briefings in bioinformatics*, 22(1), 178–193.

<https://doi.org/10.1093/bib/bbz155>

Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alföldi, J., Francioli, L. C., Khera, A. V., Lowther, C., Gauthier, L. D., Wang, H., Watts, N. A., Solomonson, M., O'Donnell-Luria, A., Baumann, A., Munshi, R., Walker, M., Whelan, C. W., Huang, Y., Brookings, T., Sharpe, T., Talkowski, M. E. (2020). A structural variation reference for medical and population genetics. *Nature*, 581(7809), 444–451. <https://doi.org/10.1038/s41586-020-2287-8>

Corbett, M. A., Kroes, T., Veneziano, L., Bennett, M. F., Florian, R., Schneider, A. L., Coppola, A., Licchetta, L., Franceschetti, S., Suppa, A., Wenger, A., Mei, D., Pendziwiat, M., Kaya, S., Delledonne, M., Straussberg, R., Xumerle, L., Regan, B., Crompton, D., van Rootselaar, A. F., Gecz, J. (2019). Intronic ATTTTC repeat expansions in *STARD7* in familial adult myoclonic epilepsy linked to chromosome 2. *Nature communications*, 10(1), 4920. <https://doi.org/10.1038/s41467-019-12671-y>

Cretu Stancu, M., van Roosmalen, M. J., Renkens, I., Nieboer, M. M., Middelkamp, S., de Ligt, J., Pregno, G., Giachino, D., Mandrile, G., Espejo Valle-Inclan, J., Korzelius, J., de Bruijn, E., Cuppen, E., Talkowski, M. E., Marschall, T., de Ridder, J., and Kloosterman, W. P. (2017). Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nature communications*, 8(1), 1326. <https://doi.org/10.1038/s41467-017-01343-4>

de Pagter, M. S., van Roosmalen, M. J., Baas, A. F., Renkens, I., Duran, K. J., van Binsbergen, E., Tavakoli-Yaraki, M., Hochstenbach, R., van der Veken, L. T., Cuppen, E., and Kloosterman, W. P. (2015). Chromothripsis in healthy individuals affects multiple protein-coding genes and can result in severe congenital abnormalities in offspring. *American journal of human genetics*, 96(4), 651–656. <https://doi.org/10.1016/j.ajhg.2015.02.005>

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D., and Cohen, D. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database : the journal of biological databases and curation, 2017, bax028.  
<https://doi.org/10.1093/database/bax028>

Florian, R. T., Kraft, F., Leitão, E., Kaya, S., Klebe, S., Magnin, E., van Rootselaar, A. F., Buratti, J., Kühnel, T., Schröder, C., Giesselmann, S., Tschernoster, N., Altmueller, J., Lamiral, A., Keren, B., Nava, C., Bouteiller, D., Forlani, S., Jornea, L., Kubica, R., Depienne, C. (2019). Unstable TTTTA/TTTCA expansions in MARCH6 are associated with Familial Adult Myoclonic Epilepsy type 3. Nature communications, 10(1), 4919.  
<https://doi.org/10.1038/s41467-019-12763-9>

Frith, M. C., and Kawaguchi, R. (2015). Split-alignment of genomes finds orthologies more accurately. Genome biology, 16(1), 106. <https://doi.org/10.1186/s13059-015-0670-9>

Gonçalves, A., Oliveira, J., Coelho, T., Taipa, R., Melo-Pires, M., Sousa, M., and Santos, R. (2017). Exonization of an Intronic LINE-1 Element Causing Becker Muscular Dystrophy as a Novel Mutational Mechanism in Dystrophin Gene. Genes, 8(10), 253.  
<https://doi.org/10.3390/genes8100253>

Hamada, M., Ono, Y., Asai, K., and Frith, M. C. (2017). Training alignment parameters for arbitrary sequencers with LAST-TRAIN. Bioinformatics (Oxford, England), 33(6), 926–928. <https://doi.org/10.1093/bioinformatics/btw742>

Jain, M., Koren, S., Miga, K. H., Quick, J., Rand, A. C., Sasani, T. A., Tyson, J. R., Beggs, A. D., Dilthey, A. T., Fiddes, I. T., Malla, S., Marriott, H., Nieto, T., O'Grady, J., Olsen, H. E., Pedersen, B. S., Rhie, A., Richardson, H., Quinlan, A. R., Snutch, T. P., Loose, M. (2018). Nanopore sequencing and assembly of a human genome with ultra-long reads. Nature biotechnology, 36(4), 338–345. <https://doi.org/10.1038/nbt.4060>

Lee, H., Deignan, J. L., Dorrani, N., Strom, S. P., Kantarci, S., Quintero-Rivera, F., Das, K., Toy, T., Harry, B., Yourshaw, M., Fox, M., Fogel, B. L., Martinez-Agosto, J. A., Wong, D. A., Chang, V. Y., Shieh, P. B., Palmer, C. G., Dipple, K. M., Grody, W. W., Vilain, E., ... Nelson, S. F. (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA*, 312(18), 1880–1887.

<https://doi.org/10.1001/jama.2014.14604>

Lei, M., Liang, D., Yang, Y., Mitsuhashi, S., Katoh, K., Miyake, N., Frith, M. C., Wu, L., and Matsumoto, N. (2020). Long-read DNA sequencing fully characterized chromothripsis in a patient with Langer-Giedion syndrome and Cornelia de Lange syndrome-4. *Journal of human genetics*, 65(8), 667–674. <https://doi.org/10.1038/s10038-020-0754-6>

Logsdon, G. A., Vollger, M. R., and Eichler, E. E. (2020). Long-read human genome sequencing and its applications. *Nature reviews. Genetics*, 21(10), 597–614.

<https://doi.org/10.1038/s41576-020-0236-x>

MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., Scherer, S. W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic acids research*, 42(Database issue), D986–D992.

<https://doi.org/10.1093/nar/gkt958>

Magi, A., Bolognini, D., Bartalucci, N., Mingrino, A., Semeraro, R., Giovannini, L., Bonifacio, S., Parrini, D., Pelo, E., Mannelli, F., Guglielmelli, P., and Maria Vannucchi, A. (2019). Nano-GLADIATOR: real-time detection of copy number alterations from nanopore sequencing data. *Bioinformatics (Oxford, England)*, 35(21), 4213–4221.

<https://doi.org/10.1093/bioinformatics/btz241>



Miga, K. H., Koren, S., Rhie, A., Vollger, M. R., Gershman, A., Bzikadze, A., Brooks, S., Howe, E., Porubsky, D., Logsdon, G. A., Schneider, V. A., Potapova, T., Wood, J., Chow, W., Armstrong, J., Fredrickson, J., Pak, E., Tigyi, K., Kremitzki, M., Markovic, C., Phillippy, A. M. (2020). Telomere-to-telomere assembly of a complete human X chromosome. *Nature*, 585(7823), 79–84. <https://doi.org/10.1038/s41586-020-2547-7>

Miller, D. E., Sulovari, A., Wang, T., Loucks, H., Hoekzema, K., Munson, K. M., Lewis, A. P., Fuerte, E., Paschal, C. R., Walsh, T., Thies, J., Bennett, J. T., Glass, I., Dipple, K. M., Patterson, K., Bonkowski, E. S., Nelson, Z., Squire, A., Sikes, M., Beckman, E., ... Eichler, E. E. (2021). Targeted long-read sequencing identifies missing disease-causing variation. *American journal of human genetics*, 108(8), 1436–1449. <https://doi.org/10.1016/j.ajhg.2021.06.006>

Mizuguchi, T., Toyota, T., Adachi, H., Miyake, N., Matsumoto, N., and Miyatake, S. (2019a). Detecting a long insertion variant in SAMD12 by SMRT sequencing: implications of long-read whole-genome sequencing for repeat expansion diseases. *Journal of human genetics*, 64(3), 191–197. <https://doi.org/10.1038/s10038-018-0551-7>

Mizuguchi, T., Suzuki, T., Abe, C., Umemura, A., Tokunaga, K., Kawai, Y., Nakamura, M., Nagasaki, M., Kinoshita, K., Okamura, Y., Miyatake, S., Miyake, N., and Matsumoto, N. (2019b). A 12-kb structural variation in progressive myoclonic epilepsy was newly identified by long-read whole-genome sequencing. *Journal of human genetics*, 64(5), 359–368. <https://doi.org/10.1038/s10038-019-0569-5>

Mizuguchi, T., Okamoto, N., Yanagihara, K., Miyatake, S., Uchiyama, Y., Tsuchida, N., Hamanaka, K., Fujita, A., Miyake, N., and Matsumoto, N. (2021). Pathogenic 12-kb copy-neutral inversion in syndromic intellectual disability identified by high-fidelity long-read sequencing. *Genomics*, 113(1 Pt 2), 1044–1053. <https://doi.org/10.1016/j.ygeno.2020.10.038>

Nazaryan-Petersen, L., Bjerregaard, V. A., Nielsen, F. C., Tommerup, N., and Tümer, Z. (2020). Chromothripsis and DNA Repair Disorders. *Journal of clinical medicine*, 9(3), 613. <https://doi.org/10.3390/jcm9030613>

Nishimura-Tadaki, A., Wada, T., Bano, G., Gough, K., Warner, J., Kosho, T., Ando, N., Hamanoue, H., Sakakibara, H., Nishimura, G., Tsurusaki, Y., Doi, H., Miyake, N., Wakui, K., Saitsu, H., Fukushima, Y., Hirahara, F., and Matsumoto, N. (2011). Breakpoint determination of X:autosome balanced translocations in four patients with premature ovarian failure. *Journal of human genetics*, 56(2), 156–160. <https://doi.org/10.1038/jhg.2010.155>

Pellestor, F., Anahory, T., Lefort, G., Puechberty, J., Liehr, T., Hédon, B., and Sarda, P. (2011). Complex chromosomal rearrangements: origin and meiotic behavior. *Human reproduction update*, 17(4), 476–494. <https://doi.org/10.1093/humupd/dmr010>

Prabakar, R. K., Xu, L., Hicks, J., and Smith, A. D. (2019). SMURF-seq: efficient copy number profiling on long-read sequencers. *Genome biology*, 20(1), 134. <https://doi.org/10.1186/s13059-019-1732-1>

Redin, C., Brand, H., Collins, R. L., Kammin, T., Mitchell, E., Hodge, J. C., Hanscom, C., Pillalamarri, V., Seabra, C. M., Abbott, M. A., Abdul-Rahman, O. A., Aberg, E., Adley, R., Alcaraz-Estrada, S. L., Alkuraya, F. S., An, Y., Anderson, M. A., Antolik, C., Anyane-Yeboah, K., Atkin, J. F., Talkowski, M. E. (2017). The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nature genetics*, 49(1), 36–45. <https://doi.org/10.1038/ng.3720>

Riggs, E. R., Andersen, E. F., Cherry, A. M., Kantarci, S., Kearney, H., Patel, A., Raca, G., Ritter, D. I., South, S. T., Thorland, E. C., Pineda-Alvarez, D., Aradhya, S., and Martin, C. L. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genetics in medicine : official journal of the American College of Medical Genetics*, 22(2), 245–257. <https://doi.org/10.1038/s41436-019-0686-8>

Saitsu, H., Kurosawa, K., Kawara, H., Eguchi, M., Mizuguchi, T., Harada, N., Kaname, T., Kano, H., Miyake, N., Toda, T., and Matsumoto, N. (2009). Characterization of the complex 7q21.3 rearrangement in a patient with bilateral split-foot malformation and hearing loss. *American journal of medical genetics. Part A*, 149A(6), 1224–1230. <https://doi.org/10.1002/ajmg.a.32877>

Saitsu, H., Osaka, H., Sugiyama, S., Kurosawa, K., Mizuguchi, T., Nishiyama, K., Nishimura, A., Tsurusaki, Y., Doi, H., Miyake, N., Harada, N., Kato, M., and Matsumoto, N. (2012). Early infantile epileptic encephalopathy associated with the disrupted gene encoding Slit-Robo Rho GTPase activating protein 2 (SRGAP2). *American journal of medical genetics. Part A*, 158A(1), 199–205. <https://doi.org/10.1002/ajmg.a.34363>

Schneider, V. A., Graves-Lindsay, T., Howe, K., Bouk, N., Chen, H. C., Kitts, P. A., Murphy, T. D., Pruitt, K. D., Thibaud-Nissen, F., Albracht, D., Fulton, R. S., Kremitzki, M., Magrini, V., Markovic, C., McGrath, S., Steinberg, K. M., Auger, K., Chow, W., Collins, J., Harden, G., Church, D. M. (2017). Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome research*, 27(5), 849–864. <https://doi.org/10.1101/gr.213611.116>

Sergey Nurk, Sergey Koren, Arang Rhie, Mikko Rautiainen, Andrey V. Bzikadze, Alla Mikheenko, Mitchell R. Vollger, Nicolas Altemose, Lev Uralsky, Ariel Gershman, Sergey Aganezov, Savannah J. Hoyt, Mark Diekhans, Glennis A. Logsdon, Michael Alonge, Stylianos E. Antonarakis, Matthew Borchers, Gerard G. Bouffard, Shelise Y. Brooks, Gina V. Caldas, Haoyu Cheng, Chen-Shan Chin, William Chow, Leonardo G. de Lima, Philip C. Dishuck, Richard Durbin, Tatiana Dvorkina, Ian T. Fiddes, Giulio Formenti, Robert S. Fulton, Arkarachai Functammasan, Erik Garrison, Patrick G.S. Grady, Tina A. Graves-Lindsay, Ira M. Hall, Nancy F. Hansen, Gabrielle A. Hartley, Marina Haukness, Kerstin Howe, Michael W. Hunkapiller, Chirag Jain, Miten Jain, Erich D. Jarvis, Peter Kerpedjiev, Melanie Kirsche, Mikhail Kolmogorov, Jonas Korlach, Milinn Kremitzki, Heng Li, Valerie V. Maduro, Tobias Marschall, Ann M. McCartney, Jennifer McDaniel, Danny E. Miller, James C. Mullikin, Eugene W. Myers, Nathan D. Olson, Benedict Paten, Paul Peluso, Pavel A. Pevzner, David Porubsky, Tamara Potapova, Evgeny I. Rogae, Jeffrey A. Rosenfeld, Steven L. Salzberg, Valerie A. Schneider, Fritz J. Sedlazeck, Kishwar Shafin, Colin J. Shew, Alaina Shumate, Yumi Sims, Arian F. A. Smit, Daniela C. Soto, Ivan Sović, Jessica M. Storer, Aaron Streets, Beth A. Sullivan, Françoise Thibaud-Nissen, James Torrance, Justin Wagner, Brian P. Walenz, Aaron Wenger, Jonathan M. D. Wood, Chunlin Xiao, Stephanie M. Yan, Alice C. Young, Samantha Zarate, Urvashi Surti, Rajiv C. McCoy, Megan Y. Dennis, Ivan A. Alexandrov, Jennifer L. Gerton, Rachel J. O'Neill, Winston Timp, Justin M. Zook, Michael C. Schatz, Evan E. Eichler, Karen H. Miga, Adam M. Phillippy. (2021). The complete sequence of a human genome. bioRxiv 445798; doi: <https://doi.org/10.1101/2021.05.26.445798>

Shorokhova, M., Nikolsky, N., and Grinchuk, T. (2021). Chromothripsis-Explosion in Genetic Science. *Cells*, 10(5), 1102. <https://doi.org/10.3390/cells10051102>

Sone, J., Mitsuhashi, S., Fujita, A., Mizuguchi, T., Hamanaka, K., Mori, K., Koike, H., Hashiguchi, A., Takashima, H., Sugiyama, H., Kohno, Y., Takiyama, Y., Maeda, K., Doi, H., Koyano, S., Takeuchi, H., Kawamoto, M., Kohara, N., Ando, T., Ieda, T., Sobue, G. (2019). Long-read sequencing identifies GGC repeat expansions in NOTCH2NLC associated with neuronal intranuclear inclusion disease. *Nature genetics*, 51(8), 1215–1221. <https://doi.org/10.1038/s41588-019-0459-y>

Stephens, P. J., Greenman, C. D., Fu, B., Yang, F., Bignell, G. R., Mudie, L. J., Pleasance, E. D., Lau, K. W., Beare, D., Stebbings, L. A., McLaren, S., Lin, M. L., McBride, D. J., Varela, I., Nik-Zainal, S., Leroy, C., Jia, M., Menzies, A., Butler, A. P., Teague, J. W., Campbell, P. J. (2011). Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*, 144(1), 27–40. <https://doi.org/10.1016/j.cell.2010.11.055>

Suzuki, T., Tsurusaki, Y., Nakashima, M., Miyake, N., Saitsu, H., Takeda, S., and Matsumoto, N. (2014). Precise detection of chromosomal translocation or inversion breakpoints by whole-genome sequencing. *Journal of human genetics*, 59(12), 649–654. <https://doi.org/10.1038/jhg.2014.88>

Takayama, J., Tadaka, S., Yano, K., Katsuoka, F., Gocho, C., Funayama, T., Makino, S., Okamura, Y., Kikuchi, A., Sugimoto, S., Kawashima, J., Otsuki, A., Sakurai-Yageta, M., Yasuda, J., Kure, S., Kinoshita, K., Yamamoto, M., and Tamiya, G. (2021). Construction and integration of three de novo Japanese human genome assemblies toward a population-specific reference. *Nature communications*, 12(1), 226. <https://doi.org/10.1038/s41467-020-20146-8>

Talkowski, M. E., Mullegama, S. V., Rosenfeld, J. A., van Bon, B. W., Shen, Y., Repnikova, E. A., Gastier-Foster, J., Thrush, D. L., Kathiresan, S., Ruderfer, D. M., Chiang, C., Hanscom, C., Ernst, C., Lindgren, A. M., Morton, C. C., An, Y., Astbury, C., Brueton, L. A., Lichtenbelt, K. D., Ades, L. C., Elsea, S. H. (2011). Assessment of 2q23.1 microdeletion syndrome implicates MBD5 as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder. *American journal of human genetics*, 89(4), 551–563. <https://doi.org/10.1016/j.ajhg.2011.09.011>

Bonnet, C., Ali Khan, A., Bresso, E., Vigouroux, C., Béri, M., Lejczak, S., Deemer, B., Andrieux, J., Philippe, C., Moncla, A., Giurgea, I., Devignes, M. D., Leheup, B., and Jonveaux, P. (2013). Extended spectrum of MBD5 mutations in neurodevelopmental disorders. *European journal of human genetics : EJHG*, 21(12), 1457–1461.

<https://doi.org/10.1038/ejhg.2013.22>

Tang, H. L., Tang, H. M., Mak, K. H., Hu, S., Wang, S. S., Wong, K. M., Wong, C. S., Wu, H. Y., Law, H. T., Liu, K., Talbot, C. C., Jr, Lau, W. K., Montell, D. J., and Fung, M. C. (2012). Cell survival, DNA damage, and oncogenic transformation after a transient and reversible apoptotic response. *Molecular biology of the cell*, 23(12), 2240–2252. <https://doi.org/10.1091/mbc.E11-11-0926>

Tsuchida, N., Nakashima, M., Kato, M., Heyman, E., Inui, T., Haginoya, K., Watanabe, S., Chiyonobu, T., Morimoto, M., Ohta, M., Kumakura, A., Kubota, M., Kumagai, Y., Hamano, S. I., Lourenco, C. M., Yahaya, N. A., Ch'ng, G. S., Ngu, L. H., Fattal-Valevski, A., Weisz Hubshman, M., Matsumoto, N. (2018). Detection of copy number variations in epilepsy using exome data. *Clinical genetics*, 93(3), 577–587.

<https://doi.org/10.1111/cge.13144>

Trujillano, D., Bertoli-Avella, A. M., Kumar Kandaswamy, K., Weiss, M. E., Köster, J., Marais, A., Paknia, O., Schröder, R., Garcia-Aznar, J. M., Werber, M., Brandau, O., Calvo Del Castillo, M., Baldi, C., Wessel, K., Kishore, S., Nahavandi, N., Eyaid, W., Al Rifai, M. T., Al-Rumayyan, A., Al-Twaijri, W., Abou Jamra, R. (2017). Clinical exome sequencing: results from 2819 samples reflecting 1000 families. *European journal of human genetics : EJHG*, 25(2), 176–182. <https://doi.org/10.1038/ejhg.2016.146>

Yeetong, P., Pongpanich, M., Srichomthong, C., Assawapitaksakul, A., Shotelersuk, V., Tantirukdham, N., Chunharas, C., Suphapeetiporn, K., and Shotelersuk, V. (2019). TTTCA repeat insertions in an intron of YEATS2 in benign adult familial myoclonic epilepsy type 4. *Brain : a journal of neurology*, 142(11), 3360–3366.

<https://doi.org/10.1093/brain/awz267>

Yuan, B., Wang, L., Liu, P., Shaw, C., Dai, H., Cooper, L., Zhu, W., Anderson, S. A., Meng, L., Wang, X., Wang, Y., Xia, F., Xiao, R., Braxton, A., Peacock, S., Schmitt, E., Ward, P. A., Vetrini, F., He, W., Chiang, T., Bi, W. (2020). CNVs cause autosomal recessive genetic diseases with or without involvement of SNV/indels. *Genetics in medicine : official journal of the American College of Medical Genetics*, 22(10), 1633–1641. <https://doi.org/10.1038/s41436-020-0864-8>

## 論文目録

### I. 主論文

・ A pipeline for complete characterization of complex germline rearrangements from long DNA reads.

Mitsuhashi, S.#, Ohori, S.# (#: co-first authors), Katoh, K., Frith, M. C., and Matsumoto, N.

Genome Medicine, Vol 12, Issue1, Page 67. 2020.

### II. 副論文

・ Long-read whole-genome sequencing identified a partial MBD5 deletion in an exome-negative patient with neurodevelopmental disorder.

Ohori, S., Tsuburaya, R. S., Kinoshita, M., Miyagi, E., Mizuguchi, T., Mitsuhashi, S., Frith, M. C., and Matsumoto, N.

Journal of Human Genetics, Vol 66, Issue 7, Page 697–705. 2021.

### ・ (総説)

長鎖シーケンサーを用いたヒト疾患解析

尾堀佐知子, 三橋里美, 松本直通

遺伝子医学, 33号, Vol 10. No.3. 23頁～28頁. 2020

### III. 参考論文

・ A novel PAK1 variant causative of neurodevelopmental disorder with postnatal macrocephaly.

Ohori, S., Mitsuhashi, S., Ben-Haim, R., Heyman, E., Sengoku, T., Ogata, K., Matsumoto, N.

Journal of Human Genetics. Vol 65, Issue5, Page 481-485.2020.