Nonsense variants of \textit{STAG2} result in distinct congenital anomalies

Hiromi Aoi\textsuperscript{1,2}, Ming Lei\textsuperscript{1}, Takeshi Mizuguchi\textsuperscript{1}, Nobuko Nishioka\textsuperscript{3}, Tomohide Goto\textsuperscript{4}, Sahoko Miyama\textsuperscript{5}, Toshifumi Suzuki\textsuperscript{2}, Kazuhiro Iwama\textsuperscript{1}, Yuri Uchiyama\textsuperscript{1}, Satomi Mitsuhashi\textsuperscript{1}, Atsuo Itakura\textsuperscript{2}, Satoru Takeda\textsuperscript{2} and Naomichi Matsumoto\textsuperscript{1}

Abstract

Herein, we report two female cases with novel nonsense mutations of \textit{STAG2} at Xq25, encoding stromal antigen 2, a component of the cohesion complex. Exome analysis identified c.3097 C>T, p.(Arg1033*) in Case 1 (a fetus with multiple congenital anomalies) and c.2229 G>A, p.(Trp743*) in Case 2 (a 7-year-old girl with white matter hypoplasia and cleft palate). X inactivation was highly skewed in both cases.

Introduction

Cohesin is a multisubunit protein complex consisting of four core proteins: structural maintenance of chromosome 1 (SMC1), structural maintenance of chromosome 3 (SMC3), RAD21 cohesion complex component (RAD21), and stromal antigen (STAG)\texttextsuperscript{1}. The cohesion subunits STAG1, STAG2, or STAG3 can directly attach to a tripartite ring (comprising SMC1, SMC3, and RAD21) to entrap chromatids\textsuperscript{1}. Other interacting proteins, such as the cohesion loader NIPBL, also regulate the biological functions of cohesion\textsuperscript{1}.

Cohesin is involved in a range of important functions, including functions in sister chromatid cohesion, DNA repair, transcriptional regulation, and architecture\textsuperscript{1,2}. Hence, germline pathogenic variants of genes encoding cohesin subunits and their interacting proteins, such as NIPBL, SMC1A, SMC3, and RAD21, are known to cause developmental disorders referred to as cohesinopathies\textsuperscript{3}, and these are characterized by intellectual disability (ID), growth retardation, and limb abnormalities\textsuperscript{4}.

Recently, \textit{STAG2} was added to the list of genes mutated in cohesinopathies\textsuperscript{5,6}. As \textit{STAG2} is essential for DNA replication fork progression, \textit{STAG2} defects may result in replication fork stalling and collapse with disruption of the interaction between the cohesin ring and the replication machinery as previously described\textsuperscript{7}. To date, 16 pathogenic variants of \textit{STAG2} have been reported, including seven nonsense, four missense, one splicing, and four frameshift variants\textsuperscript{5,8–13}. Notably, seven male patients in three families harbored missense variants. In one family, five affected males showed ID and congenital abnormalities\textsuperscript{11}, and two other sporadic males were reported to have dysmorphic features, short stature, hypotonia, developmental delay (DD) and ID\textsuperscript{9,10}. Female patients had truncated and missense variants\textsuperscript{5,8,10,12,13}. Here, we describe the genetic and clinical features of two female cases with de novo nonsense variants of \textit{STAG2}.

Case 1 was the second conceptus of healthy Japanese nonconsanguineous parents (a 35-year-old mother and 37-year-old father). At 15 gestational weeks, holoprosencephaly, cleft palate, cleft lip, blepharophimosis, nasal bone absence, and hypoplastic left heart were noted by ultrasonography. The fetal karyotype determined by amniocentesis at 18 gestational weeks was normal (46, XX). The pregnancy was terminated at 21 gestational weeks because of multiple fetal abnormalities.
Case 2 was a 7-year-old girl who was born as the second child to healthy nonconsanguineous parents. She was born uneventfully at full term. Her birth weight was 2734 g (~1.3 SD). A cleft palate was noted at birth and surgically repaired at 1 year. She presented with mild dysmorphic features, including a long philtrum. At 8 months, she developed afebrile convulsions for which carbamazepine was effective. Anticonvulsants were discontinued for 4 years with no subsequent attacks. She acquired independent gait at 2 years and spoke only a few words at 7 years. Brain magnetic resonance imaging at 7 years revealed white matter hypoplasia. She currently has mild DD, ID, sensorineural hearing loss, and amblyopia with no neurologic abnormalities. She attends a school for hearing-impaired children.

This study was approved by the institutional review board of Yokohama City University School of Medicine. WES was performed in the two cases (Cases 1 and 2) and their parents. Blood leukocytes from the patient (Case 2), parents (Cases 1 and 2) and umbilical cord (Case 1) were obtained after obtaining informed consent. Exome data acquisition, processing, annotation, and filtering and variant calling were performed as previously described. Possible pathogenic variants were evaluated based on mutational type (nonsense, missense, frameshift, or splice site) using the SIFT score (http://sift.jcvi.org/), Polyphen-2 (http://genes.genetics.bwh.harvard.edu/pph2/), Mutation Taster (http://www.mutationtaster.org/), and CADD (https://cadd.gs.washington.edu/). Possible pathogenic variants were validated by Sanger sequencing. Parentage was confirmed using 12 microsatellite markers with GeneMapper software v4.1.1 (Life Technologies Inc., Carlsbad, CA).

Total RNA was extracted from lymphoblastoid cell lines (LCLs) with the RNeasy Plus Mini Kit (Qiagen, Germany) and, reverse-transcribed to cDNA with the Super Script First Strand Synthesis System (Takara, Japan), and the cDNA used as templates for RT-PCR. PCR amplicons were subjected to Sanger sequencing.

CNVs were examined using WES data by two algorithms: the eXome Hidden Markov Model, and a program based on the relative depth of coverage ratios developed by Nord et al. X chromosome inactivation was determined using the human androgen receptor gene. X-inactivation ratios (expressed arbitrarily as a ratio of the smaller allele to the larger allele) were calculated twice and judged as published criteria: <20:80 (random), >20:80 (skewed), and >10:90 (highly skewed).

Ten micrograms of sheared DNA was subjected to library preparation using a single-molecule real-time (SMRT)bell Express Template Prep Kit 2.0 (Pacific Biosciences, 100-938-900) and a SMRTbell Enzyme Cleanup Kit (Pacific Biosciences, 101-746-400) in accordance with the manufacturer’s instructions (Procedure & Checklist - Preparing HiFi SMRTbell® Libraries using SMRTbell Express Template Prep Kit 2.0, Pacific Biosciences). One SMRT cell was used for the patient (Case 1). Secondary analysis using base-called data was performed using SMRT analysis v8.0 (Pacific Biosciences). Circular consensus sequencing (CCS) from single molecules was performed, and the generated sequence was mapped to the hg19 human reference genome using the CCS with Mapping application, provided by SMRT analysis, with the default settings. DeepVariant 0.9.0 (https://github.com/google/deepvariant) was used to detect SNVs and indels in CCS reads. The aligned CCS BAM data from the CCS with Mapping application were used as an input. We ran Google DeepVariant with a model trained for PacBio CCS (-model_type=PACBIO) using the prebuilt Docker image from the DeepVariant public repository (https://github.com/google/deepvariant). Small variant calls from DeepVariant were haplotyped and phased using WhatShap 0.18 (https://whatshap.readthedocs.io/en/latest/).

We first performed WES in Case 1. Case 1 had no pathogenic variants in 14 known mutated genes associated with holoprosencephaly, namely, SHH, ZIC2, SIX3, TGFIF1, GLI2, PTCH1, DZIP1, FGF8, FOXH1, NODAL, TGDFI, GAS1, DLI1, and CDON. Moreover, no pathogenic CNVs were identified by exome-based CNV analysis. After analyzing trio-based WES data, three de novo variants were found. (Supplementary Table S1), but two missense variants were likely benign based on computational predictions. The remaining de novo nonsense variant [c.3097 C>T, p.(Arg1033*)] of STAG2 was confirmed by Sanger sequencing (Fig. 1a) and was likely causative. X inactivation was highly skewed (93:7), and the paternal X chromosome was inactivated (Supplementary Fig. S1). Unfortunately, living cells from Case 1 could not be obtained for further mRNA analysis. Using HiFi long-read genome sequencing and haplotype phasing with informative variants, we constructed haplotypes in the vicinity of STAG2 and confirmed that the STAG2 variant occurred de novo on the paternal chromosome in Case 1 (Fig. 2). As the paternal X chromosome is mostly inactivated in blood leukocytes, the X inactivation pattern should be favorable in Case 1. We also identified another STAG2 nonsense mutation [c.2229 G>A, p.(Trp743*)] occurring de novo in Case 2 (Fig. 1a, Table 1). X inactivation was highly skewed (96:4), and the maternal X chromosome was inactivated (Supplementary Fig. S1). RT-PCR indicated that only the wild-type allele was expressed in LCLs of Case 2 (Supplementary Fig. S2). Even after cycloheximide treatment, the mutant allele was completely undetectable, suggesting that it was transcriptionally repressed (through favorably skewed X inactivation) rather than posttranscriptionally diminished (through nonsense-mediated mRNA decay) in cultured
Regardless of the favorable X inactivation pattern in both cases, Case 1 was clinically much more severe than Case 2. Therefore, it is difficult to discuss phenotype severity in relation to the X inactivation pattern.

In addition to the two variants in our cases, a total of 16 pathogenic variants of STAG2 have been reported in unrelated families (Table 1), including 12 truncated variants [p.(Arg69*), p.(Gln140*), p.(Arg259*), p.(Cys535*), p.(Val547Cysfs*29), p.(Lys553Ilefs*6), p.(Arg614*), p.(Arg614*), p.(Arg1012*), p.(Glu968Serfs*15), p.(Trp743*)], and four missense variants [p.(Tyr159Cys), p.(Ser327Asn), p.(Arg604Gln), and p.(Lys1009Asn)]5,8–13.

For a female patient with p.(Ala638Valfs*10), no detailed phenotype was provided in the DECIPHER database, therefore, this patient was omitted for further comparison of clinical features8. Twelve cases with STAG2 truncation variants reported in the literature were all females, and one missense variant was reported in a female patient. The 12 female cases with STAG2 truncation shared microcephaly (10/12), abnormal brain MRI findings (10/12, including holoprosencephaly 7/12), thoracic vertebral anomalies (6/12), DD (9/12), and ID (4/12). Case 1 showed severe clinical features, such as holoprosencephaly and hypoplastic left heart, similar to previous literature, while Case 2’s clinical features were relatively mild. p.(Arg69*) was recurrent in two unrelated patients, both showing middle brain anomalies8,12.

Our two cases showed highly skewed X inactivation (93:7 in Case 1 and 96:4 in Case 2) (Supplementary Fig. S1). To date, X inactivation analysis has been reported in only two cases, one with skewed X inactivation and
another with random X inactivation, but the exact ratios were not shown in the literature\textsuperscript{10,12}. Interestingly, one splicing variant (c.2533+1 G) was inherited from the mother, but unfortunately, an X inactivation study was not conducted\textsuperscript{12}.

In contrast, null \textit{STAG2} variants in males have never been reported. We speculate that males with a hemizygous truncating \textit{STAG2} aberration are lethal or show severe fetal clinical ends. Interestingly, one missense variant [p.(Ser327-Asn)] was transmitted in an X-linked recessive manner in a

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
#CHROM & POS & ID & Allele1 & Allele2 & Fa & Mo \\
\hline
X & 123195035 & rs2297651 & A & C & C/C & A/A \\
\hline
X & 12320440 & \textbf{NM_001042751.1:c.3097C>T:p.R1033*} & C & T & C/C & C/C \\
\hline
X & 123307694 & rs2279726 & T & C & C/C & T/T \\
\hline
X & 123308295 & rs7355854 & C & T & T/T & C/T \\
\hline
X & 12346811 & rs12393728 & A & G & G/G & A/A \\
\hline
X & 123467158 & \textbf{rs11381271} & G & T & T/T & G/G \\
\hline
X & 123556628 & & A & G & G/G & A/G \\
\hline
\end{tabular}
\caption{Confirmation of the pathogenic \textit{STAG2} variant in the paternal chromosome in Case 1. a The c.3097 C>T variant could be successfully mapped within the 450-kb phased haplotype block in Case 1 using HiFi sequence and haplotype phasing. b SNP typing confirmed that the mutation occurred in the paternal chromosome (Allele 2, the brown haplotype block in Fig. 2a). POS: position of sequence, Pt: patient, Fa: father, Mo: mother.}
\end{table}
| Characteristics | Case 1 | Case 2 | Mullegama et al. | Lan Yu et al. | Yuan et al. |
|-----------------|--------|--------|-----------------|---------------|-------------|
| **Sex**         | Female | Female | Female          | Female        | Female      |
| **STAG2 variants** | c.3097 T>C | c.2229 G>A | c.1913_1922 del | c.1840 C>T | c.1605 T>A |
| **Protein change** | p.(Arg1033*) | p.(Trp743*) | p.(Ala638Valfs*10) | p.(Arg614*) | p.(Gln140*) |
| **Inheritance**  | de novo | de novo | de novo | de novo | de novo |
| **Microcephaly** | NA | – | + | NA | + |
| **Brain MRI/US findings** | HPE | white matter hypoplasia | dysgenesis of the splenium of the corpus callosum, subarachnoid cyst, subgaleal hematoma | NA | HPE (microform) |
| **Dysmorphic features** | NA | + | + | NA | + |
| **Craniofacial anomalies** | cleft lip/palate | cleft palate | NA | NA | NA |
| **Congenital heart defects** | left heart hypoplasia | – | VSD | NA | – |
| **Thoracic vertebra anomalies** | NA | hemi vertebra | scoliosis, hemivertebra, butterfly vertebra | NA | scoliosis, rib fusion, vertebral clefts |
| **Developmental delay** | NA | + | + | NA | + |
| **Intellectual disability** | NA | + | + | NA | + |
| **Other features** | amblyopia, seizures | left facial palsy, mild left pelvectomy | sacral dimple, CDH | gastro-esophageal reflux, CDH, seizures | hyptonia |

Aoi et al. Human Genome Variation (2020) 7:26 Page 5 of 7

Official Journal of the Japan Society of Human Genetics
| Characteristics             | Kruszka et al.\(^\text{11}\) | Berkovic et al.\(^\text{12}\) | Mullegama et al.\(^\text{8}\) | Soardi et al.\(^\text{10}\) |
|-----------------------------|-------------------------------|---------------------------------|-------------------------------|-----------------------------|
| **Sex**                     | Female                        | Female                          | Female                        | Female                      |
| **STAG2 variants**          | c.205 C>T                     | c.436 C>T                       | c.775 C>T                     | c.2898_2899del              |
| **Protein change**          | p.(Arg69*)                    | p.(Arg146*)                     | p.(Arg259*)                   | p.(Glu968Serfs*15)          |
| **Inheritance**             | de novo                       | de novo                         | de novo                       | de novo                     |
| **Microcephaly**            | +                             | +                               | +                             | NA                          |
| **Brain MRI/US findings**   | HPE (semi-lobar)              | HPE (alobar)                    | HPE (septo-optic dysplasia)   | HPE (microform)             |
| **Dysmorphic features**     | +                             | +                               | +                             | +                           |
| **Craniofacial anomalies**  | cleft palate, micrognathia    | cyclopia, absent nose, hypognathia | cleft lip/palate              | NA                          |
| **Congenital heart defects**| PFO, PDA                      | VSD                             | NA                            | left heart hypoplasia, DORV |
| **Thoracic vertebra anomalies** | --                           | hemi vertebra                   | spina bifida                  | NA                          |
| **Developmental delay**     | +                             | NA                              | NA                            | +                           |
| **Intellectual disability** | NA                            | NA                              | NA                            | +                           |
| **Other features**          | duodenal atresia              | left hip dysplasia, bilateral optic nerve dysplasia | gastro-esophageal reflux | seizures, hearing loss |

\(c.980 \text{G} > \text{A}\) was reported in a family with five affected males.  
NA not available, MRI magnetic resonance imaging, US ultrasonography, HPE holoprosencephaly, VSD ventricular septal defect, CDH congenital diaphragmatic hernia, CA coarctation of the aorta, PFO patent foramen ovale, PDA patent ductus arteriosus, DORV double-outlet right ventricle.
family with five affected males and two healthy carrier females 11. These five males showed ID (5/5), several facial dysmorphisms [large nose (5/5), prominent ears (5/5), frontal baldness (4/5)], hearing loss (3/5), short stature (5/5), and cleft palate (1/5). An additional two hemizygous missense variants [p.(Tyr159Cys) and p.(Lys1009Asn)] were recently reported in two unrelated males 9,10. They showed facial dysmorphisms (2/2), cleft lip and palate (1/2), pituitary gland abnormality (1/2), patent foramen ovale (1/2), hypotonia (2/2), DD (2/2), and ID (2/2), as seen in the above family. Male patients with missense variants exhibited milder mutant effects than those with truncated variants, as expected.

In conclusion, of the two female patients with STAG2 variants, one showed a severe prenatal phenotype, while the other showed a mild pediatric phenotype. X inactivation was highly skewed in both cases. This phenotypic difference might depend on another factor, such as a modifier, that is yet to be found.

HGV Database
The relevant data from this Data Report are hosted at the Human Genome Variation Database at https://doi.org/10.6084/m9.figshare.hgv.2891, https://doi.org/10.6084/m9.figshare.hgv.2894.

Acknowledgements
We would like to thank the patients and their families for participating in this study. We also thank N. Watanabe, T. Miyama, M. Sato, and K. Takabe for their technical assistance. This work was supported by AMED under Grant numbers JP18ek0109280, JP18dk0107090, JP18ek0109301, JP18ek0109348, and JP18dk0205001; by JSPS KAKENHI under Grant numbers JP17H01539, JP16H05357, JP16H06254, JP17K16132, JP17K10080, JP17K15630, and JP17H06994; by the Ministry of Health, Labour, and Welfare; and by the Takeda Science Foundation. We thank Sarah Williams, PhD, from Edanz Group (www.edanzediting.com) for editing a draft of this manuscript.

Author details
1Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan. 2Department of Obstetrics and Gynecology, Faculty of Medicine, Juntendo University, Tokyo, Japan. 3Department of Obstetrics and Gynecology, Koshigaya Municipal Hospital, Saitama, Japan. 4Department of Neurology, Kanagawa Children’s Medical Center, Kanagawa, Japan. 5Department of Neurology, Tokyo Metropolitan Children’s Medical Center, Fuchu, Tokyo, Japan

Conflict of interest
The authors declare that they have no conflict of interest.

References
1. Kline, A. D. et al. Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. Nat. Rev. Genet. 19, 649–666 (2018).
2. Leroy, C. et al. Xq25 duplication: the crucial role of the STAG2 gene in this novel human cohesinopathy. Clin. Genet. 89, 68–73 (2016).
3. Deardorff, M. A. et al. RAD21 mutations cause a human cohesinopathy. Am. J. Hum. Genet. 90, 1014–1027 (2012).
4. Baquerizo-Montoya, C. et al. Could a patient with SMC1A duplication be classified as a human cohesinopathy? Clin. Genet. 85, 446–451 (2014).
5. Yu, L. et al. Increased burden of de novo predicted deleterious variants in complex congenital diaphragmatic hernia. Hum. Mol. Genet. 24, 4764–4773 (2015).
6. Yingjun, X. et al. Microduplication of chromosome Xq25 encompassing STAG2 gene in a boy with intellectual disability. Eur. J. Med. Genet. 58, 116–121 (2015).
7. Mondal, G., Stevens, M., Goode, B., Ashworth, A. & Solomon, D. A. A requirement for STAG2 in replication fork progression creates a targetable synthetic lethality in cohesin-mutant cancers. Nat. Commun. 10, 1866 (2019).
8. Mullegama, S. V. et al. De novo loss-of-function variants in STAG2 are associated with developmental delay, microcephaly, and congenital anomalies. Am. J. Med. Genet. A 173, 1319–1327 (2017).
9. Mullegama, S. V., Klein, S. D., Signer, R. H., Vilain, E. & Martinez-Agosto, J. A. Mutations in STAG2 cause an X-linked cohesinopathy associated with undergrowth, developmental delay, and dysmorphism: expanding the phenotype in males. Mol. Genet. Genom. Med. 7, e00501 (2019).
10. Yuan, B. et al. Clinical exome sequencing reveals locus heterogeneity and phenotypic variability of cohesinopathies. Genet. Med. 21, 663–675 (2019).
11. Soardi, F. C. et al. Familial STAG2 germline mutation defines a new human cohesinopathy. NPJ Genom. Med. 2, 7 (2017).
12. Kruzhika, P. et al. Cohesin complex-associated holoprosencephaly. Brain 142, 2631–2643 (2019).
13. Berkovic, S. F. et al. The epilepsy genetics initiative: systematic reanalysis of diagnostic exomes increases yield. Epilepsia 60, 797–806 (2019).
14. Iwama, K. et al. A novel SLC9A1 mutation causes cerebellar ataxia. J. Hum. Genet. 63, 1049–1054 (2018).
15. Former, M. et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. Am. J. Hum. Genet. 91, 597–607 (2012).
16. Nord, A. S., Lee, M., King, M. C. & Walsh, T. Accurate and exact CNV identification from targeted high-throughput sequence data. BMC Genomics 12, 184 (2011).
17. Amos-Landgraf, J. M. et al. X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. Am. J. Hum. Genet. 79, 493–499 (2006).

Publisher’s note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information is available for this paper at https://doi.org/10.1038/s41439-020-00114-w.