Minireview

Mechanisms of disease-associated SINE-VNTR-Alus

Abigail L Pfaff[1,2], Lewis M Singleton[1] and Sulev Kõks[1,2]

[1]Perron Institute for Neurological and Translational Science, Perth, WA 6009, Australia; [2]Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, WA 6150, Australia
Corresponding author: Abigail L Pfaff. Email: Abigail.pfaff@uwa.edu.au

Impact Statement

This review provides a comprehensive overview of disease-associated SVAs identified to date discussing the different mechanisms through which these insertions act. These complex structural variants can be difficult to detect and characterize through routine analysis; however, the number of disease-causing SVAs is increasing. This review highlights the importance of evaluating this type of variation when looking to identify disease-causing variants, particularly in cases where standard pipelines have been unable to determine the causative variant. Understanding the genetic component of disease could potentially lead to novel therapeutic targets.

Abstract

SINE-VNTR-Alus (SVAs) are the youngest retrotransposon family in the human genome. Their ongoing mobilization has generated genetic variation within the human population. At least 24 insertions to date, detailed in this review, have been associated with disease. The predominant mechanisms through which this occurs are alterations to normal splicing patterns, exonic insertions causing loss-of-function mutations, and large genomic deletions. Dissecting the functional impact of these SVAs and the mechanism through which they cause disease provides insight into the consequences of their presence in the genome and how these elements could influence phenotypes. Many of these disease-associated SVAs have been difficult to characterize and would not have been identified through routine analyses. However, the number identified has increased in recent years as DNA and RNA sequencing data became more widely available. Therefore, as the search for complex structural variation in disease continues, it is likely to yield further disease-causing SVA insertions.

Keywords: SINE-VNTR-Alu, retrotransposon, genetic variation, splicing, mobile DNA, disease

***Experimental Biology and Medicine* 2022; 247: 756–764. DOI: 10.1177/15353702221082612**

Introduction

Though previously considered "junk DNA," mobile DNA, or transposable elements (TEs), are a source of both genetic variation[1–4] and gene expression modulation.[5–7] Class I TEs, or retrotransposable elements (RTEs), utilize a copy and paste mechanism when mobilizing, increasing their numbers within the host genome as a result.[8] RTE-specific mobilization encompasses an RNA intermediate that is reverse transcribed into a complementary DNA (cDNA) "copy," which is then inserted into the genome at a locus different to that of the source element.[9] Within the human genome, the only known TEs to be currently active are the non-long terminal repeat (non-LTR) retrotransposons consisting of long interspersed elements (LINEs), short interspersed elements (SINEs) and SINE-VNTR (variable number tandem repeat)-Alus (SVAs).[10] SVAs, the youngest of the RTEs in the human genome and hominid-specific, are so termed after their composite domains.[11] SVAs consist of a hexamer repeat (CCCTCT), *Alu*-like sequence on the antisense strand, GC-rich VNTR, SINE region, a poly A-tail and target site duplications flanking the site of insertion (Figure 1(a)).[12–14]

LINE-1 (L1) elements are the only active autonomous RTE in humans.[10] Encoding the proteins required to retrotranspose, L1s are able to mobilize themselves[15] and other RTEs such as *Alus*[16] and SVAs.[11,17–19] L1s retrotranspose by way of target primed reverse transcription (TPRT), a process that begins when L1 RNA is transcribed by RNA polymerase II before being exported to the cytoplasm of the cell.[20] Within the cytoplasm the two open reading frames, ORF1 and ORF2, encoded by the L1 will be translated into proteins, ORF1p and ORF2p respectively.[21] ORF1p binds to nucleic acids,[22] while ORF2p has both reverse-transcriptase and endonuclease activities.[23,24] Following translation in the cytoplasm, ORF1p and ORF2p join L1 RNA to form a ribonucleoprotein complex (L1 RNP) which is then transported back into the nucleus.[25,26] To ensure a functional L1 is more likely to be inserted into the host genome the L1 encoded proteins demonstrate a *cis* preference for their encoding RNA.[27] Facilitated by its endonuclease activity, ORF2p nicks the bottom strand of the hosts DNA at the 5′TTTTAA3′ consensus sequence at the TA site. Primed by a 3′-hydroxyl group liberated at the TA site, ORFp2 then reverse transcribes the L1 RNA into cDNA before nicking the top strand of the host DNA to allow

ISSN 1535-3702

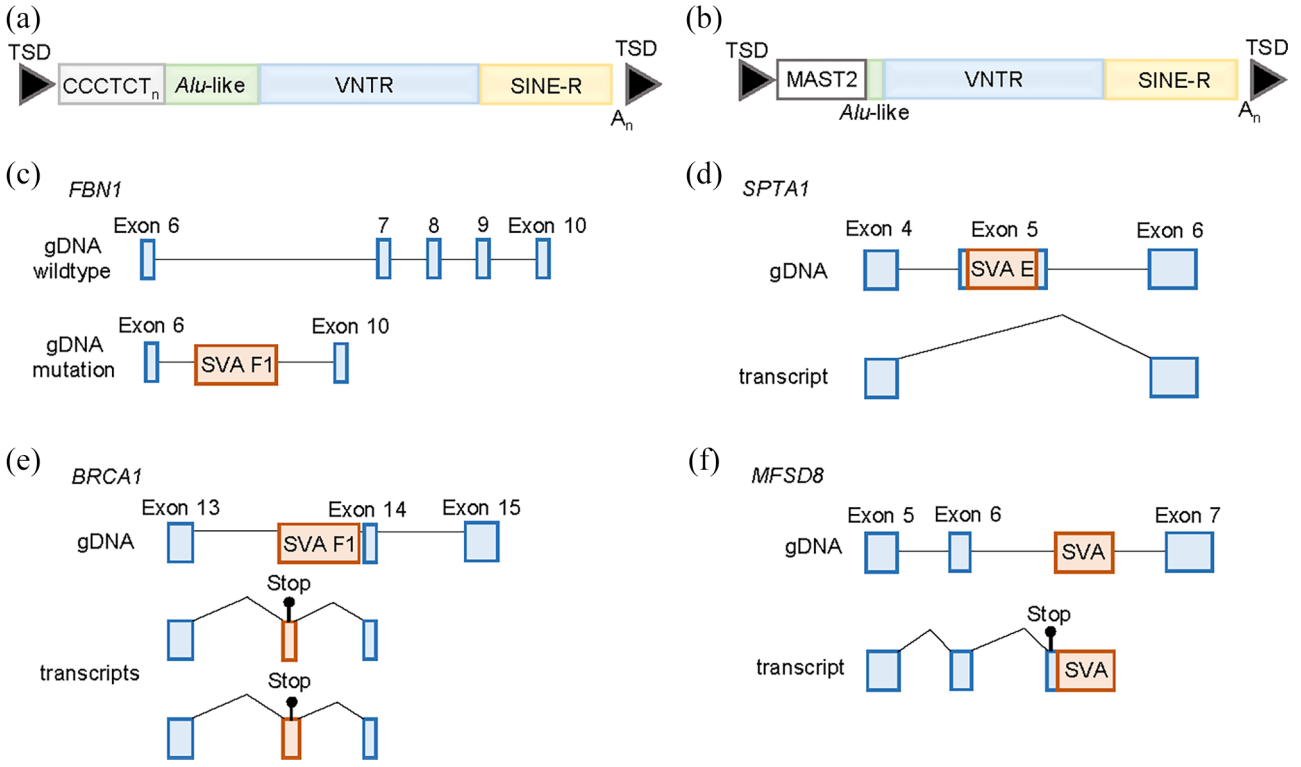

Figure 1. SVA structure and mechanisms through which they are associated with disease. (a) A complete SVA consists of a hexamer repeat (CCCTCT) variable in length, an *Alu*-like sequence on the antisense strand, a variable number tandem repeat (VNTR), a SINE-R domain, a poly A-tail and are usually flanked by target site duplications (TSDs). (b) An SVA F1 consists of the VNTR, SINE-R, and poly A-tail but is lacking the CCCTCT domain and the majority of the *Alu*-like sequence. At the 5′ end is sequence from the MAST2 exon 1 that can vary in size. (c) An example of a genomic deletion upon SVA insertion where a region of ~36.8 kb is deleted that includes exons 7-9 of the *FBN1* gene. (d) Exon skipping induced by the insertion of an SVA into exon 5 of the *SPTA1* gene resulting in an inframe deletion and the production of an abnormal protein. This was associated with hereditary elliptocytosis and hereditary pyropoikilocytosis. (e) Exonization of an intronic SVA insertion that introduced a premature stop codon in the *BRCA1* gene in a family with early onset breast cancer. (f) An SVA insertion into an intron of the *MFSD8* gene that activates existing cryptic splice sites causing missplicing and the introduction of a premature stop codon in the transcript. This was associated with neuronal ceroid lipofuscinosis 7. (A color version of this figure is available in the online journal.)

integration of the cDNA into the genome.[25] Finally, the complementary strand of DNA is synthesized.[20]

SVAs are divided into subtypes (A–F) according to the SINE region, with the oldest elements belonging to A and F being the youngest.[11] A subsequent seventh subtype was identified, SVA F1, containing a 5′ transduction of the sequence from the *MAST2* gene and the incorporated *MAST2* sequence having been shown to act as a positive regulator of transcription (Figure 1(b)).[28–31] Elements from the subtypes D, E, and F1 were found to retrotranspose to varying degrees in multiple cell lines.[17,19] These studies demonstrated that ORF2p was required for retrotransposition, while the need for ORF1p depended on the SVA element co-transfected with the L1 driver construct.[17,19] This may be due to size or sequence differences of the SVAs tested. L1 mobilization requires both proteins; however, *Alu* retrotransposition only requires ORF2p.[16] The rate of retrotransposition in the population for *Alu*, L1, and SVAs had been estimated as 1/21, 1/212, and 1/916 live births, respectively.[32] However, using a pedigree analysis, the rate was reported as 1/40 for *Alus* and 1/63 for both L1s and SVAs.[33] The rate of SVA mobilization was much higher in the pedigree analysis than the previous estimate.

In the literature, 24 SVA insertions were identified as being associated with disease. The mechanisms behind this fell into three broad categories of large genomic deletions upon insertion of an SVA, exonic insertions leading to loss-of-function mutations, and those causing aberrant splicing (Table 1). For the 21 SVAs that the subtype was reported, they all belonged to the three most recently active subtypes (E, F, and F1), which are human specific. The first disease-causing SVA was identified in 1994 and 12 such insertions had been identified up until 2016.[2] In the past 6 years, this number has doubled as technological advances have allowed the analysis of these types of elements more widely. The affordability of high-depth whole-genome sequencing and a range of bioinformatics tools developed to call retrotransposon insertions from these data have been an important advance in characterizing this type of variation genome-wide.[34] It has also enabled population scale analyses to identify polymorphic insertions and provide a reference of variants common to a population for comparison.[3,4] Here we discuss the disease-associated SVAs and the mechanisms through which they act to provide insight into how these elements influence genomic function.

Exonic SVA insertions and loss-of-function mutations

SVAs contain stop codons in their sequence and insertions into exons will often introduce a premature stop codon or cause a frameshift mutation, leading to the transcript being

Table 1. Disease-associated SVAs.

Gene	Disease	SVA subtype	Size (bp)	Location	Mechanism	Reference
A4GNT	Chromothripsis	SVA E	502	Intron	Insertion into intron 2 associated with a 110 kb deletion	Nazaryan-Petersen *et al.*[35]
BBS1	Bardet–Biedl syndrome	SVA F	2435	Exon	Insertion into exon 13 introducing a premature stop codon	Delvallee *et al.*,[36] Tavares *et al.*[37]
BRCA1	Breast cancer	SVA F1	2856	Intron	Exonization of SVA sequence introducing a premature stop codon	Walsh *et al.*[38]
BRCA2	Breast cancer	SVA E	~2000	Exon	Disruption of the BRCA2 DNA-binding domain	Deuitch *et al.*[39]
BTK	X-linked agammaglobulinemia	–	253	Intron	Disruption of the 5′ splice site of intron 9 resulting in the skipping of exon 9	Conley *et al.*,[40] Rohrer *et al.*[41]
CHM	Choroideremia	NR	NR	Exon	Insertion into exon 2 results in exon 2 skipping and loss of the REP-1 protein encoded by the *CHM* gene	Jones *et al.*[42]
EBP	X-linked dominant chrondrodysplasia punctata	SVA F1	~4500	Exon	Insertion into exon 2 introducing a premature stop codon	Hiraide *et al.*[43]
FBN1	Dilated aortic sinus	SVA F1	~1500	Intron	Deletion of approximately 36.8 kb, which included three exons of *FBN1*	Brett *et al.*[44]
FIX	Hemophilia B	SVA F	2524	Exon	Insertion into exon 6 at the intron 5–exon 6 boundary inducing aberrant splicing and may result in reduced transcript levels due to degradation by nonsense-mediated decay	Nakamura *et al.*[45]
FKTN	Fukuyama-type congenital muscular dystrophy	SVA E	3023	3′-UTR	Insertion into 3′-UTR induces aberrant splicing by the presence of an acceptor site in the SVA and the activation of an alternative donor site in exon 10	Kobayashi *et al.*,[46] Watanabe *et al.*,[47] Taniguchi-Ikeda *et al.*[48]
GAA	Pompe disease	SVA E with 3′ transduction of L1ME3	3394	Intron	Homozygous insertion leads to the near complete loss of the full-length GAA isoform due to exonization of the insertion and early transcript termination at inserted polyA tail	Bychkov *et al.*[49]
HLA-A	Leukemia	SVA F1	2000	–	A 14 kb deletion that included the entire *HLA-A* gene	Hancks *et al.*,[30] Takasu *et al.*[50]
LDLRAP1	Autosomal recessive hypercholesterolemia	SVA E	2600	Intron	SVA insertion into intron 1 leads to exonization of the SVA sequence and degradation of the transcript by nonsense-mediated decay	Taniguchi-Ikeda *et al.*,[48] Wilund *et al.*[51]
MFSD8	Neuronal ceroid lipofuscinosis 7 (Batten's disease)	NR	~2000	Intron	Insertion led to the activation of a cryptic splice acceptor site in intron 6 199 bp upstream of the SVA introducing a premature stop codon	Kim *et al.*[52]
MSH2	Lynch syndrome	SVA F1	~3000	Exon	Insertion into exon 12 is predicted to introduce either a premature stop codon or cause a frameshift mutation	Yang *et al.*[53]
MSH2	Lynch syndrome	SVA F	~2400	Exon	Insertion into exon 3 induced aberrant splicing and created an intron in the middle of exon 3 leading to a frameshift mutation	Yamamoto *et al.*[54]
MSH6	Lynch syndrome	SVA E	~2400	Exon	Insertion into *exon 5* gene causing aberrant splicing and the use of a cryptic 3′ splice leading to an inframe deletion of the MSH6 protein	Yamamoto *et al.*[54]
PMS2	Lynch syndrome	SVA F	2200	Intron	The insertion into intron 7 activates a cryptic splice acceptor site leading to Exonization of part of the SVA sequence causing a frameshift and a premature stop codon	van der Klift *et al.*[55]
PNPLA2	Neutral lipid storage disease with subclinical myopathy	SVA E	1800	Exon	Insertion into exon 3 and degradation of the transcript by nonsense-mediated decay	Akman *et al.*[56]
TAF1	X-linked dystonia parkinsonism	SVA F	2627	Intron	Insertion into intron 32 that leads to partial intron retention proximal to the SVA and a reduction in overall TAF1 expression	Makino *et al.*,[57] Aneichyk *et al.*[58]
SMARCB1	Atypical teratoid rhabdoid tumor	SVA E	2763	Intron	The insertion into intron causes aberrant splicing by the presence of a splice acceptor site in the SVA sequence leading to loss of the functional transcript	Sabatella *et al.*[59]
SPTA1	Hereditary elliptocytosis and hereditary pyropoikilocytosis	SVA E	632	Exon	Insertion into exon 5 leads to skipping of the exon, which causes an inframe deletion and abnormalities in the protein structure and function	Ostertag *et al.*,[18] Hassoun *et al.*[60]
SUZ12P	Neurofibromatosis type 1	SVA F1	1700	Intron	Insertion into intron 8 of SUZ12P associated with a 1 Mb deletion that included the *NF1* gene	Vogt *et al.*[61]
SUZ12P	Neurofibromatosis type 1	SVA F	1300	Intron	Insertion into intron 8 of SUZ12P associated with a 867 kb deletion that included the *NF1* gene	Vogt *et al.*[61]

Source: Data also taken from Hancks *et al.*[2]
SVA: SINE-VNTR-Alus (short interspersed elements–variable number tandem repeat–*Alu*-like sequence); DNA: deoxyribonucleic acid; NR: not reported.

degraded by nonsense-mediated decay. Several examples of SVAs causing loss-of-function mutations via exonic insertion have been reported (Table 1). These include an SVA F in the *PNPLA2* gene associated with neutral lipid storage disease with subclinical myopathy, an SVA E in the *BRCA2* gene associated with breast cancer and an SVA F1 in the *MSH2* gene associated with the cancer predisposition disease Lynch syndrome.[39,53,56] An SVA F insertion in exon 13 of the *BBS1* gene was identified as the second most common pathogenic variant associated with the rare recessive disease Bardet–Beidl syndrome, which occurred in a common ancestor at an estimated 74 generations ago.[36,37] In addition, a recent report of SVA F1 insertion with accompanying 5′ and 3′ transductions of *Alu* sequences was associated with X-linked dominant chondrodysplasia punctata (CDPX2). CDPX2 occurs almost exclusively in females and is associated with skin, bone, and eye abnormalities caused by pathogenic variants in the *EBP* gene.[43] Structural variant analysis of whole-genome sequencing data identified the SVA F1 in exon 2 of the *EBP* gene in the proband and her affected mother and inspection of the SVA F1 sequence provides evidence for the introduction of a premature stop codon.[43]

SVA insertions associated with large genomic deletions

A deletion of three *FBN1* exons (7–9) from intron 6 to intron 9, spanning approximately 38.6 kb at 15q21.1 with a concomitant SVA F1 insertion, was identified in a study of a child with mildly dilated aortic sinus (Figure 1(c)).[44] Mutations in the *FBN1* gene typically result in an autosomal dominant connective tissue disorder termed Marfan syndrome (MFS), which affects the skeletal, cardiovascular, and ocular systems.[62] In this case, the child displayed no other features indicative of an MFS diagnosis. Loss of exons 7 and 8 likely influence correct folding of the fibrillin protein,[63,64] while additional disruption of the coding sequence of the messenger RNA (mRNA) may result from the acceptor and donor splice sites contained within the inserted SVA. Large intragenic deletions and chromosomal imbalances are only responsible for an estimated 1–2% of MFS cases,[65,66] and no previously reported cases involved the deletion of exons 7–9.

Chromothripsis (CTH) is the occurrence of hundreds of DNA double-stranded breaks (DSBs) typically within small genomic regions, though regions as large as entire chromosomes have been identified.[67,68] Initially, CTH was exclusively associated with cancer pathogenesis; however, when later identified in some congenital and developmental disorders, limited cases of germline-CTH (GCTH) were described.[69–71] The DSB are typically repaired to generate genomic fragment rearrangements involving translocations, inversions, and insertions. With few resultant deletions being observed, CTH is relatively balanced.[70] In a case of familial G-CTH, numerous break points were identified in chromosome 3 accompanied by a 502 bp 5′-truncated SVA E element insertion into intron 2 of the *A4GNT* gene.[35] Associated with this insertion was a 110 kb deletion at the 5′-end of the insertion. The insertion location within a sequence reminiscent of the L1 endonuclease cleavage site and truncation of the insertion itself at its 5′-end suggest L1-mediated retrotransposition as the mechanism by which this insertion occurred.[18,23] In addition, the location of the deleted sequence at the 5′-end of the insertion and lack of target-site duplications, are features reminiscent of previously described retrotransposon insertions with concomitant deletion events.[72,73] Finally, the 100% sequence match between the inserted SVA E element and its source element, an SVA E located on chromosome 7, further indicates retrotransposition as the mechanism of this insertion event.

The insertion and deletion events most likely occurred concomitantly due to several factors. The authors suggest that ORF2p endonuclease activity precipitated multiple DNA breaks leading to G-CTH and mediated SVA E retrotransposition associated with a large-scale deletion. Specifically, the authors postulate a model whereby two mispaired AluSx elements flanking the deleted segment, both in the same orientation, may have predisposed the segment for deletion, mediating chromatin looping and arranging the segment proximally to ORF2p. This facilitated cleavage and SVA retrotransposition.

When investigating genetic samples provided for bone marrow testing individuals from three apparently unrelated Japanese families, another study found SVA retrotransposition accompanied by a large 14 kb deletion comprising the entire *HLA-A* gene.[50] The 2 kb insertion SVA was later identified as belonging to the SVA F1 subfamily.[30] Interestingly, though the three families were apparently unrelated, they shared identical *HLA-A* haplotypes and each family originated from the same area of Japan, North Kanto. This suggested a possible common ancestor from which the original deletion and insertion was inherited as a founder mutation.[50]

Accounting for approximately 8–10% of large *NF1* deletions, "atypical" deletions have non-recurrent breakpoints and vary by size and the number of genes within the deleted region.[74,75] Prior analysis of the few highly characterized atypical *NF1* deletions suggested they were due to non-homologous end joining (NHEJ). Two instances of large atypical *NF1* deletions within *SUZ12P* intron 8 in unrelated patients were found to be accompanied by SVA retrotransposition at the deletion breakpoints.[61] One patient possessed a 1 Mb deletion with a 1.7 kb SVA insertion likely sourced from one of the most active SVA elements and belonging to the SVA F1 subfamily, H10_1 on chromosome 10q24.2.[29,30] The other patient displayed a 867 kb deletion accompanied by an SVA insertion highly homologous to SVA F element H6_1084 on chromosome 6q22.31. The sites of both insertions were separated by 3067 bp, highlighting the concept of retrotransposition hotspots within the human genome as postulated by earlier studies.[40,76–78] The study concluded that the origin deletion and insertion events occurred concomitantly during early post-zygotic development for each case (indicated by the somatic mosaicism present in one patient and grandmother of the other). Endonuclease cleavage sites located within *SUZ12P* intron 8 and long polyT tracts at integration sites suggested the insertion events were mediated by L1-associated TPRT. Again, it was postulated that SVA insertion resulted in loop-like conformational changes to chromatin. This conformational change brought the *SUZ12P* sequence ending with a newly inserted SVA element into

proximity of the telomeric *NF1* gene region, facilitating their ligation most likely by NHEJ.[79]

SVA-induced aberrant splicing

More than half of the disease associated SVA insertions outlined in Table 1 cause aberrant splicing patterns, that include exon skipping, activation of cryptic splice sites up or downstream of the SVA, and the inclusion of the SVA sequence in gene transcripts. This can result in smaller proteins that can impair their normal function or frameshift mutations and nonsense-mediated decay leading to disease. Although, not in all cases can the precise effects be determined. For example, an SVA insertion into exon 6 of the *FIX* gene that causes Hemophilia B alters normal splicing of exons 5 and 6; however, whether this was due to exon skipping or exonization was unknown.[45]

Exon skipping

The skipping of exons due to mutation can lead to disease either through altering the reading frame or protein structure. SVA insertions in exons or at the exon–intron boundary have led to three cases of disease (Table 1). An intronic insertion, close to the exon–intron boundary, in the *BTK* gene caused skipping of exon 9 by disrupting the 5′ splice site of intron 9 resulting in an inframe deletion, an unstable protein, and the immune system affecting disorder X-linked agammaglobulinemia.[40,41] The insertion of an SVA into exon 5 of the *SPTA1* gene causes exon 5 skipping that leads to an inframe deletion producing abnormalities in the protein structure and function and is associated with the heterogeneous red blood cell disorder hereditary elliptocytosis and hereditary pyropoikilocytosis (Figure 1(e)).[18,60] The third example of SVA-induced skipping is an insertion in exon 2 of the *CHM* gene and is associated with choroideremia, a condition characterized by progressive vision loss.[42] This insertion causes exon 2 skipping and an absence of the REP-1 protein encoded by the *CHM* gene.

Exonization of SVA sequence

An SVA in the sense orientation contains multiple cryptic splice sites throughout its sequence and several instances of these SVA splice sites being used have been reported.[30] In addition, the most recent subfamily of SVAs (SVA F1) was created when exon 1 of the *MAST2* gene was spliced into the *Alu*-like region of an SVA F and was subsequently retrotransposed (Figure 1(b)).[28–30] This subfamily rapidly expanded with more than 80 SVA F1s in the human genome[29] and six disease-associated insertions (Table 1). The SVA sequence also contains multiple stop codons within its sequence; therefore, exonization of the SVA is likely to introduce premature stop codons causing nonsense-mediated decay or produce truncated proteins. Six of the disease-associated SVAs reported in Table 1 act through this mechanism, five of which are located within introns and one in a 3′-UTR. The SVA in the 3′ UTR of the *FKTN* gene is a founder insertion and is associated with Fukuyama-type congenital muscular dystrophy (FCMD).[46,47] FCMD is a type of muscular dystrophy accompanied by abnormalities in the brain and eyes and is predominantly found in Japan. The insertion causes an abnormal splicing event with a rare alternative donor site in exon 10 (last exon) and an acceptor site in the SVA being used. This leads to a truncation of the normal fukutin protein, and the final 129 amino acids being coded for by the SVA causing mislocalization of the protein.[48]

An SVA in intron 1 of the *LDLRAP1* gene is associated with autosomal recessive hypercholesterolemia leading to no detectable *LDLRAP1* mRNA in patient cells due to abnormal splicing and degradation of the SVA-containing transcript.[48,51] Three of the intronic SVA insertions that undergo exonization are in genes in which a germline mutation predisposes an individual to certain cancers. A case of Lynch syndrome was associated with an SVA in intron 7 of the *PMS2* gene that led to the exonization of 71 bp sequence consisting of nucleotides from the target site duplication and the SVA itself.[55] The remainder of the SVA sequence was spliced out using a cryptic donor site in the SVA sequence and the canonical splice acceptor of intron 7. The inclusion of this sequence caused a frameshift, which introduced a stop codon and the transcript being degraded. Using a long-read sequencing approach to analyze the DNA of a family affected by early-onset breast cancer, an SVA insertion was identified in intron 13 of the *BRCA1* gene in the proband, which segregated with breast cancer in the family.[38] Previous panel testing and exome sequencing had been unable to identify the genetic variant involved. The SVA resulted in two additional transcripts, depending on the size of the SVA sequence included, to be expressed both of which contained premature stop codons (Figure 1(d)). Atypical teratoid rhabdoid tumor is a rare and aggressive pediatric tumor caused by the biallelic inactivation of *SMARCB1* with nearly one-third of cases carrying a predisposing germline variant. One such germline variant was identified as an SVA in a pair of siblings diagnosed with the disease.[59] The SVA was located in intron 2 of the gene, which caused splicing from exon 2 into the SVA using a splice acceptor in the *Alu*-like region. This is likely a highly penetrant variant due to both siblings being affected, and the insertion only present in the mother in a mosaic state.[59] The final example of an exonized SVA in disease is located in the *GAA* gene and is associated with Pompe disease, an autosomal recessive lysosomal storage disorder.[49] The patient was homozygous for an insertion in intron 15 of *GAA* gene, which caused exon 15 to be spliced into the SVA and termination of the transcript at the insertion's poly A-tail and the almost complete absence of the full-length isoform.

Activation of cryptic splice sites

SVAs have been shown to alter splicing by activating existing cryptic splice sites at the locus in which they insert, in both introns and exons. An SVA F insertion in exon 3 of the *MSH2* gene in a patient with Lynch Syndrome led to the use of a new donor and acceptor splice site up and downstream of the SVA insertion. This divided exon 3 into two, creating an intron in the middle of the existing exon.[54] This resulted in shorter mRNA and a frameshift mutation. A second case of Lynch Syndrome was associated with an SVA E insertion into exon 5 of the *MSH6* gene, also through changes to normal

splicing patterns.[54] This insertion altered the splice acceptor site that was used to downstream of the SVA causing an inframe deletion in exon 5.

An SVA in intron 32 of the *TAF1* gene was associated with X-linked dystonia-parkinsonism (XDP).[57] The presence of the SVA was associated with reduced *TAF1* expression and the inclusion of a cryptic exon in intron 32 5′ of the SVA. Removal of the SVA using CRISPR/Cas9 rescued this aberrant transcription normalizing *TAF1* expression.[58] In addition to the presence of the SVA influencing XDP, the length of the CCCTCT domain at the 5′-end of the SVA was associated with age of disease onset.[80] The median repeat number was found to be higher in the basal ganglia and cerebellum compared to the blood of the same individual suggesting somatic repeat instability is a feature of specific brain regions and could play a role in disease manifestation.[81] Therefore, size variation of the SVA should also be an important factor when evaluating their functional impact.

Whole-genome sequencing was performed on a patient diagnosed with neuronal ceroid lipofuscinosis 7 (CLN7), a form of Batten's disease, who was heterozygous for a known pathogenic mutation in the *MFSD8* gene to identify the second mutation involved. An SVA in intron 6 of the gene was identified, which was absent from over 800 whole genomes and present in the mother of the proband. Analysis of RNA from the patient demonstrated missplicing of exon 6 into a cryptic splice acceptor site 199 bp upstream of the SVA introducing a stop codon and was predicted to lead to premature termination of translation (Figure 1(g)).

Modulation of SVA-induced aberrant splicing using antisense oligonucleotides

Antisense oligonucleotides (AOs) are single-stranded synthetic nucleic acid analogues that can be designed to target specific sequences to modulate gene expression and several have been licensed to treat diseases such as Duchenne muscular dystrophy and spinal muscular atrophy.[82] AOs have been designed to target missplicing caused by two SVA insertions, one of which (milasen) was approved by the Food and Drug Administration and expedited institutional review board to be administered by an intrathecal bolus injection to the patient it was designed for in a "n-of-1" situation.[48,52,83]

AOs were designed to restore translation of the full-length fukutin protein to correct the missplicing caused by the SVA insertion in the 3′ UTR of the *FTKN* gene that causes FCMD. The AOs targeted acceptor, donor, and exonic splicing enhancer sites and a cocktail of three AOs achieved the greatest recovery of *FKTN* mRNA in lymphoblast and myotube patient cell lines.[48] The AO cocktail was also tested in mice rescuing the full-length *FKTN* mRNA and restoring the normal protein.[48] This provides a potential treatment for FCMD patients through the modulation of splicing.

One of the pathogenic mutations in a single patient with the fatal neurodegenerative disorder neuronal ceroid lipofuscinosis 7 (CLN7) was identified to be an SVA insertion that altered normal splicing of the *MFSD8* gene and splice-modulating AOs were designed to correct this. The AOs targeted the cryptic splice acceptor site activated by the SVA insertion and nearby splicing enhancers. The lead candidate was chosen and was named milasen.[52] Milasen was tested in patient fibroblasts more than tripling the amount of normal splicing and alleviated cellular phenotypes associated with lysosomal dysfunction. Due to deterioration of the patient, expedited approval was given for use of the AO and treatment with milasen resulted in the reduction and frequency of seizures experienced by the patient.[52] This demonstrates the rapid development of a targeted personalized therapeutic, and that SVAs, which alter normal splicing, may be amenable to modulation using AOs.

Common SVA insertions and disease risk

The insertions discussed so far are rare and cause a robust phenotype leading to disease. However, common insertions could influence disease risk with more subtle functional effects. For example, an SVA in intron 8 of the *CASP8* gene is a common polymorphic element that was associated with the retention of intron 8, an increased risk of breast cancer and protective for prostate cancer.[84] This polymorphic SVA is in strong linkage disequilibrium (LD) with single nucleotide polymorphisms (SNPs) that were identified as risk variants in genome-wide association studies (GWAS).[84] This approach of LD analysis of SVA polymorphisms and known GWAS risk variants would be useful to identify those SVAs that could be the causative variant at known disease-associated loci. A focused analysis of reference SVA polymorphisms and Parkinson's disease (PD)-associated SNPs identified an SVA on chromosome 17 in strong LD with a known PD risk variant.[85] A genome-wide approach across multiple diseases, similar to analysis performed for *Alu* variation,[86] could provide insight into common SVA variation and disease risk.

Conclusions

The list of SVA insertions associated with disease is increasing and as the number of genomes analyzed for complex structural variants expands it is likely that the number of disease-causing SVAs will as well. Understanding the mechanisms through which these elements cause disease not only contributes to the overall knowledge regarding their function but also to potential novel therapeutics that modify their effects.

AUTHORS' CONTRIBUTIONS

ALP and LMS reviewed the literature, ALP generated the table and figure and all authors contributed to the writing and editing of the manuscript.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: ALP, LMS, and SK are funded by Multiple Sclerosis Society of Western Australia.

ORCID IDS

Abigail L Pfaff https://orcid.org/0000-0002-2231-9800
Lewis M Singleton https://orcid.org/0000-0002-3317-2701
Sulev Kõks https://orcid.org/0000-0001-6087-6643

REFERENCES

1. Chuang NT, Gardner EJ, Terry DM, Crabtree J, Mahurkar AA, Rivell GL, Hong CC, Perry JA, Devine SE. Mutagenesis of human genomes by endogenous mobile elements on a population scale. *Genome Res* 2021;**31**:2225–35
2. Hancks DC, Kazazian HH Jr. Roles for retrotransposon insertions in human disease. *Mob DNA* 2016;**7**:9
3. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, Fritz MH, Konkel MK, Malhotra A, Stutz AM, Shi X, Casale FP, Chen J, Hormozdiari F, Dayama G, Chen K, Malig M, Chaisson MJP, Walter K, Meiers S, Kashin S, Garrison E, Auton A, Lam HYK, Mu XJ, Alkan C, Antaki D, Bae T, Cerveira E, Chines P, Chong Z, Clarke L, Dal E, Ding L, Emery S, Fan X, Gujral M, Kahveci F, Kidd JM, Kong Y, Lameijer EW, McCarthy S, Flicek P, Gibbs RA, Marth G, Mason CE, Menelaou A, Muzny DM, Nelson BJ, Noor A, Parrish NF, Pendleton M, Quitadamo A, Raeder B, Schadt EE, Romanovitch M, Schlattl A, Sebra R, Shabalin AA, Untergasser A, Walker JA, Wang M, Yu F, Zhang C, Zhang J, Zheng-Bradley X, Zhou W, Zichner T, Sebat J, Batzer MA, McCarroll SA, Genomes Project C, Mills RE, Gerstein MB, Bashir A, Stegle O, Devine SE, Lee C, Eichler EE, Korbel JO. An integrated map of structural variation in 2,504 human genomes. *Nature* 2015;**526**:75–81
4. Collins RL, Brand H, Karczewski KJ, Zhao X, Alfoldi J, Francioli LC, Khera AV, Lowther C, Gauthier LD, Wang H, Watts NA, Solomonson M, O'Donnell-Luria A, Baumann A, Munshi R, Walker M, Whelan CW, Huang Y, Brookings T, Sharpe T, Stone MR, Valkanas E, Fu J, Tiao G, Laricchia KM, Ruano-Rubio V, Stevens C, Gupta N, Cusick C, Margolin L, Genome Aggregation Database Production T, Genome Aggregation Database C, Taylor KD, Lin HJ, Rich SS, Post WS, Chen YI, Rotter JI, Nusbaum C, Philippakis A, Lander E, Gabriel S, Neale BM, Kathiresan S, Daly MJ, Banks E, MacArthur DG, Talkowski ME. A structural variation reference for medical and population genetics. *Nature* 2020;**581**:444–51
5. Koks S, Pfaff AL, Bubb VJ, Quinn JP. Expression quantitative trait loci (eQTLs) associated with retrotransposons demonstrate their modulatory effect on the transcriptome. *Int J Mol Sci* 2021;**22**:6319
6. Wang L, Norris ET, Jordan IK. Human retrotransposon insertion polymorphisms are associated with health and disease via gene regulatory phenotypes. *Front Microbiol* 2017;**8**:1418
7. Wang L, Rishishwar L, Marino-Ramirez L, Jordan IK. Human population-specific gene expression and transcriptional network modification with polymorphic transposable elements. *Nucleic Acids Res* 2017;**45**:2318–28
8. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 2009;**10**:691–703
9. Kazazian HH Jr. Mobile elements: drivers of genome evolution. *Science* 2004;**303**:1626–32
10. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet* 2011;**12**:187–215
11. Wang H, Xing J, Grover D, Hedges DJ, Han K, Walker JA, Batzer MA. SVA elements: a hominid-specific retroposon family. *J Mol Biol* 2005;**354**:994–1007
12. Ono M, Kawakami M, Takezawa T. A novel human nonviral retroposon derived from an endogenous retrovirus. *Nucleic Acids Res* 1987;**15**:8725–37
13. Shen L, Wu LC, Sanlioglu S, Chen R, Mendoza AR, Dangel AW, Carroll MC, Zipf WB, Yu CY. Structure and genetics of the partially duplicated gene RP located immediately upstream of the complement C4A and the C4B genes in the HLA class III region. Molecular cloning, exon-intron structure, composite retroposon, and breakpoint of gene duplication. *J Biol Chem* 1994;**269**:8466–76
14. Zhu ZB, Hsieh SL, Bentley DR, Campbell RD, Volanakis JE. A variable number of tandem repeats locus within the human complement C2 gene is associated with a retroposon derived from a human endogenous retrovirus. *J Exp Med* 1992;**175**:1783–7
15. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. High frequency retrotransposition in cultured mammalian cells. *Cell* 1996;**87**:917–27
16. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* 2003;**35**:41–8
17. Hancks DC, Goodier JL, Mandal PK, Cheung LE, Kazazian HH Jr. Retrotransposition of marked SVA elements by human L1s in cultured cells. *Hum Mol Genet* 2011;**20**:3386–400
18. Ostertag EM, Goodier JL, Zhang Y, Kazazian HH Jr. SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am J Hum Genet* 2003;**73**:1444–51
19. Raiz J, Damert A, Chira S, Held U, Klawitter S, Hamdorf M, Lower J, Stratling WH, Lower R, Schumann GG. The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. *Nucleic Acids Res* 2012;**40**:1666–83
20. Babushok DV, Kazazian HH Jr. Progress in understanding the biology of the human mutagen LINE-1. *Hum Mutat* 2007;**28**:527–39
21. Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, Cooley T, Heath P, Smith KD, Margolet L. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. *Genomics* 1987;**1**:113–25
22. Hohjoh H, Singer MF. Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. *EMBO J* 1997;**16**:6034–43
23. Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 1996;**87**:905–16
24. Mathias SL, Scott AF, Kazazian HH Jr, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science* 1991;**254**:1808–10
25. Cost GJ, Feng Q, Jacquier A, Boeke JD. Human L1 element target-primed reverse transcription in vitro. *EMBO J* 2002;**21**:5899–910
26. Hohjoh H, Singer MF. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J* 1996;**15**:630–9
27. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JD, Moran JV. Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol* 2001;**21**:1429–39
28. Bantysh OB, Buzdin AA. Novel family of human transposable elements formed due to fusion of the first exon of gene MAST2 with retrotransposon SVA. *Biochemistry* 2009;**74**:1393–9
29. Damert A, Raiz J, Horn AV, Lower J, Wang H, Xing J, Batzer MA, Lower R, Schumann GG. 5′-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. *Genome Res* 2009;**19**:1992–2008
30. Hancks DC, Ewing AD, Chen JE, Tokunaga K, Kazazian HH Jr. Exon-trapping mediated by the human retrotransposon SVA. *Genome Res* 2009;**19**:1983–91
31. Zabolotneva AA, Bantysh O, Suntsova MV, Efimova N, Malakhova GV, Schumann GG, Gayfullin NM, Buzdin AA. Transcriptional regulation of human-specific SVAF(1) retrotransposons by cis-regulatory MAST2 sequences. *Gene* 2012;**505**:128–36
32. Xing J, Zhang Y, Han K, Salem AH, Sen SK, Huff CD, Zhou Q, Kirkness EF, Levy S, Batzer MA, Jorde LB. Mobile elements create structural variation: analysis of a complete human genome. *Genome Res* 2009;**19**:1516–26
33. Feusier J, Watkins WS, Thomas J, Farrell A, Witherspoon DJ, Baird L, Ha H, Xing J, Jorde LB. Pedigree-based estimation of human mobile element retrotransposition rates. *Genome Res* 2019;**29**:1567–77
34. Ewing AD. Transposable element detection from whole genome sequence data. *Mob DNA* 2015;**6**:24
35. Nazaryan-Petersen L, Bertelsen B, Bak M, Jonson L, Tommerup N, Hancks DC, Tumer Z. Germline chromothripsis driven by L1-mediated retrotransposition and Alu/Alu homologous recombination. *Hum Mutat* 2016;**37**:385–95

36. Delvallee C, Nicaise S, Antin M, Leuvrey AS, Nourisson E, Leitch CC, Kellaris G, Stoetzel C, Geoffroy V, Scheidecker S, Keren B, Depienne C, Klar J, Dahl N, Deleuze JF, Genin E, Redon R, Demurger F, Devriendt K, Mathieu-Dramard M, Poitou-Bernert C, Odent S, Katsanis N, Mandel JL, Davis EE, Dollfus H, Muller J. A BBS1 SVA F retrotransposon insertion is a frequent cause of Bardet-Biedl syndrome. *Clin Genet* 2021;**99**:318–24
37. Tavares E, Tang CY, Vig A, Li S, Billingsley G, Sung W, Vincent A, Thiruvahindrapuram B, Heon E. Retrotransposon insertion as a novel mutational event in Bardet-Biedl syndrome. *Mol Genet Genomic Med* 2019;**7**:e00521
38. Walsh T, Casadei S, Munson KM, Eng M, Mandell JB, Gulsuner S, King MC. CRISPR-Cas9/long-read sequencing approach to identify cryptic mutations in BRCA1 and other tumour suppressor genes. *J Med Genet* 2021;**58**:850–2
39. Deuitch N, Li ST, Courtney E, Shaw T, Dent R, Tan V, Yackowski L, Torene R, Berkofsky-Fessler W, Ngeow J. Early-onset breast cancer in a woman with a germline mobile element insertion resulting in BRCA2 disruption: a case report. *Hum Genome Var* 2020;**7**:24
40. Conley ME, Partain JD, Norland SM, Shurtleff SA, Kazazian HH Jr. Two independent retrotransposon insertions at the same site within the coding region of BTK. *Hum Mutat* 2005;**25**:324–5
41. Rohrer J, Minegishi Y, Richter D, Eguiguren J, Conley ME. Unusual mutations in Btk: an insertion, a duplication, an inversion, and four large deletions. *Clin Immunol* 1999;**90**:28–37
42. Jones KD, Radziwon A, Birch DG, MacDonald IM. A novel SVA retrotransposon insertion in the CHM gene results in loss of REP-1 causing choroideremia. *Ophthalmic Genet* 2020;**41**:341–4
43. Hiraide T, Masunaga Y, Honda A, Kato F, Fukuda T, Fukami M, Nakashima M, Saitsu H, Ogata T. Retrotransposition disrupting EBP in a girl and her mother with X-linked dominant chondrodysplasia punctata. *J Hum Genet*. Epub ahead of print 9 January 2022. DOI: 10.1038/s10038-021-01000-1
44. Brett M, Korovesis G, Lai AHM, Lim ECP, Tan EC. Intragenic multiexon deletion in the FBN1 gene in a child with mildly dilated aortic sinus: a retrotransposal event. *J Hum Genet* 2017;**62**:711–5
45. Nakamura Y, Murata M, Takagi Y, Kozuka T, Nakata Y, Hasebe R, Takagi A, Kitazawa J, Shima M, Kojima T. SVA retrotransposition in exon 6 of the coagulation factor IX gene causing severe hemophilia B. *Int J Hematol* 2015;**102**:134–9
46. Kobayashi K, Nakahori Y, Miyake M, Matsumura K, Kondo-Iida E, Nomura Y, Segawa M, Yoshioka M, Saito K, Osawa M, Hamano K, Sakakihara Y, Nonaka I, Nakagome Y, Kanazawa I, Nakamura Y, Tokunaga K, Toda T. An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature* 1998;**394**:388–92
47. Watanabe M, Kobayashi K, Jin F, Park KS, Yamada T, Tokunaga K, Toda T. Founder SVA retrotransposal insertion in Fukuyama-type congenital muscular dystrophy and its origin in Japanese and Northeast Asian populations. *Am J Med Genet A* 2005;**138**:344–8
48. Taniguchi-Ikeda M, Kobayashi K, Kanagawa M, Yu CC, Mori K, Oda T, Kuga A, Kurahashi H, Akman HO, DiMauro S, Kaji R, Yokota T, Takeda S, Toda T. Pathogenic exon-trapping by SVA retrotransposon and rescue in Fukuyama muscular dystrophy. *Nature* 2011;**478**:127–31
49. Bychkov I, Baydakova G, Filatova A, Migiaev O, Marakhonov A, Pechatnikova N, Pomerantseva E, Konovalov F, Ampleeva M, Kaimonov V, Skoblov M, Zakharova E. Complex transposon insertion as a novel cause of Pompe disease. *Int J Mol Sci* 2021;**22**:10887
50. Takasu M, Hayashi R, Maruya E, Ota M, Imura K, Kougo K, Kobayashi C, Saji H, Ishikawa Y, Asai T, Tokunaga K. Deletion of entire HLA-A gene accompanied by an insertion of a retrotransposon. *Tissue Antigens* 2007;**70**:144–50
51. Wilund KR, Yi M, Campagna F, Arca M, Zuliani G, Fellin R, Ho YK, Garcia JV, Hobbs HH, Cohen JC. Molecular mechanisms of autosomal recessive hypercholesterolemia. *Hum Mol Genet* 2002;**11**:3019–30
52. Kim J, Hu C, Moufawad El Achkar C, Black LE, Douville J, Larson A, Pendergast MK, Goldkind SF, Lee EA, Kuniholm A, Soucy A, Vaze J, Belur NR, Fredriksen K, Stojkovska I, Tsytsykova A, Armant M, DiDonato RL, Choi J, Cornelissen L, Pereira LM, Augustine EF, Genetti CA, Dies K, Barton B, Williams L, Goodlett BD, Riley BL, Pasternak A, Berry ER, Pflock KA, Chu S, Reed C, Tyndall K, Agrawal PB, Beggs AH, Grant PE, Urion DK, Snyder RO, Waisbren SE, Poduri A, Park PJ, Patterson A, Biffi A, Mazzulli JR, Bodamer O, Berde CB, Yu TW. Patient-customized oligonucleotide therapy for a rare genetic disease. *N Engl J Med* 2019;**381**:1644–52
53. Yang C, Li Y, Trottier M, Farrell MP, Rai VK, Salo-Mullen EE, Gallagher DJ, Stadler ZK, van der Klift HM, Zhang L. Insertion of an SVA element in MSH2 as a novel cause of Lynch syndrome. *Genes Chromosomes Cancer* 2021;**60**:571–6
54. Yamamoto G, Miyabe I, Tanaka K, Kakuta M, Watanabe M, Kawakami S, Ishida H, Akagi K. SVA retrotransposon insertion in exon of MMR genes results in aberrant RNA splicing and causes Lynch syndrome. *Eur J Hum Genet* 2021;**29**:680–6
55. van der Klift HM, Tops CM, Hes FJ, Devilee P, Wijnen JT. Insertion of an SVA element, a nonautonomous retrotransposon, in PMS2 intron 7 as a novel cause of Lynch syndrome. *Hum Mutat* 2012;**33**:1051–5
56. Akman HO, Davidzon G, Tanji K, Macdermott EJ, Larsen L, Davidson MM, Haller RG, Szczepaniak LS, Lehman TJ, Hirano M, DiMauro S. Neutral lipid storage disease with subclinical myopathy due to a retrotransposal insertion in the PNPLA2 gene. *Neuromuscul Disord* 2010;**20**:397–402
57. Makino S, Kaji R, Ando S, Tomizawa M, Yasuno K, Goto S, Matsumoto S, Tabuena MD, Maranon E, Dantes M, Lee LV, Ogasawara K, Tooyama I, Akatsu H, Nishimura M, Tamiya G. Reduced neuron-specific expression of the TAF1 gene is associated with X-linked dystonia-parkinsonism. *Am J Hum Genet* 2007;**80**:393–406
58. Aneichyk T, Hendriks WT, Yadav R, Shin D, Gao D, Vaine CA, Collins RL, Domingo A, Currall B, Stortchevoi A, Multhaupt-Buell T, Penney EB, Cruz L, Dhakal J, Brand H, Hanscom C, Antolik C, Dy M, Ragavendran A, Underwood J, Cantsilieris S, Munson KM, Eichler EE, Acuna P, Go C, Jamora RDG, Rosales RL, Church DM, Williams SR, Garcia S, Klein C, Muller U, Wilhelmsen KC, Timmers HTM, Sapir Y, Wainger BJ, Henderson D, Ito N, Weisenfeld N, Jaffe D, Sharma N, Breakefield XO, Ozelius LJ, Bragg DC, Talkowski ME. Dissecting the causal mechanism of X-linked dystonia-parkinsonism by integrating genome and transcriptome assembly. *Cell* 2018;**172**:897–909.e21
59. Sabatella M, Mantere T, Waanders E, Neveling K, Mensenkamp AR, van Dijk F, Hehir-Kwa JY, Derks R, Kwint M, O'Gorman L, Tropa Martins M, Gidding CE, Lequin MH, Kusters B, Wesseling P, Nelen M, Biegel JA, Hoischen A, Jongmans MC, Kuiper RP. Optical genome mapping identifies a germline retrotransposon insertion in SMARCB1 in two siblings with atypical teratoid rhabdoid tumors. *J Pathol* 2021;**255**:202–11
60. Hassoun H, Coetzer TL, Vassiliadis JN, Sahr KE, Maalouf GJ, Saad ST, Catanzariti L, Palek J. A novel mobile element inserted in the alpha spectrin gene: spectrin dayton. *J Clin Invest* 1994;**94**:643–8
61. Vogt J, Bengesser K, Claes KB, Wimmer K, Mautner VF, van Minkelen R, Legius E, Brems H, Upadhyaya M, Hogel J, Lazaro C, Rosenbaum T, Bammert S, Messiaen L, Cooper DN, Kehrer-Sawatzki H. SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints. *Genome Biol* 2014;**15**:R80
62. De Paepe A, Devereux RB, Dietz HC, Hennekam RC, Pyeritz RE. Revised diagnostic criteria for the Marfan syndrome. *Am J Med Genet* 1996;**62**:417–26
63. Aoyama T, Tynan K, Dietz HC, Francke U, Furthmayr H. Missense mutations impair intracellular processing of fibrillin and microfibril assembly in Marfan syndrome. *Hum Mol Genet* 1993;**2**:2135–40
64. Saharinen J, Keski-Oja J. Specific sequence motif of 8-Cys repeats of TGF-beta binding proteins, LTBPs, creates a hydrophobic interaction surface for binding of small latent TGF-beta. *Mol Biol Cell* 2000;**11**:2691–704
65. Furtado LV, Wooderchak-Donahue W, Rope AF, Yetman AT, Lewis T, Plant P, Bayrak-Toydemir P. Characterization of large genomic deletions in the FBN1 gene using multiplex ligation-dependent probe amplification. *BMC Med Genet* 2011;**12**:119
66. Singh KK, Elligsen D, Liersch R, Schubert S, Pabst B, Arslan-Kirchner M, Schmidtke J. Multi-exon out of frame deletion of the FBN1 gene leading to a severe juvenile onset cardiovascular phenotype in Marfan syndrome. *J Mol Cell Cardiol* 2007;**42**:352–6

67. Kloosterman WP, Hoogstraat M, Paling O, Tavakoli-Yaraki M, Renkens I, Vermaat JS, van Roosmalen MJ, van Lieshout S, Nijman IJ, Roessingh W, van't Slot R, van de Belt J, Guryev V, Koudijs M, Voest E, Cuppen E. Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. *Genome Biol* 2011;**12**:R103
68. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, McLaren S, Lin ML, McBride DJ, Varela I, Nik-Zainal S, Leroy C, Jia M, Menzies A, Butler AP, Teague JW, Quail MA, Burton J, Swerdlow H, Carter NP, Morsberger LA, Iacobuzio-Donahue C, Follows GA, Green AR, Flanagan AM, Stratton MR, Futreal PA, Campbell PJ. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 2011;**144**:27–40
69. Chiang C, Jacobsen JC, Ernst C, Hanscom C, Heilbut A, Blumenthal I, Mills RE, Kirby A, Lindgren AM, Rudiger SR, McLaughlan CJ, Bawden CS, Reid SJ, Faull RL, Snell RG, Hall IM, Shen Y, Ohsumi TK, Borowsky ML, Daly MJ, Lee C, Morton CC, MacDonald ME, Gusella JF, Talkowski ME. Complex reorganization and predominant nonhomologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. *Nat Genet* 2012;**44**:390–7, S1
70. Kloosterman WP, Tavakoli-Yaraki M, van Roosmalen MJ, van Binsbergen E, Renkens I, Duran K, Ballarati L, Vergult S, Giardino D, Hansson K, Ruivenkamp CA, Jager M, van Haeringen A, Ippel EF, Haaf T, Passarge E, Hochstenbach R, Menten B, Larizza L, Guryev V, Poot M, Cuppen E. Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms. *Cell Rep* 2012;**1**:648–55
71. Nazaryan L, Stefanou EG, Hansen C, Kosyakova N, Bak M, Sharkey FH, Mantziou T, Papanastasiou AD, Velissariou V, Liehr T, Syrrou M, Tommerup N. The strength of combined cytogenetic and mate-pair sequencing techniques illustrated by a germline chromothripsis rearrangement involving FOXP2. *Eur J Hum Genet* 2014;**22**:338–43
72. Okubo M, Horinishi A, Saito M, Ebara T, Endo Y, Kaku K, Murase T, Eto M. A novel complex deletion-insertion mutation mediated by Alu repetitive elements leads to lipoprotein lipase deficiency. *Mol Genet Metab* 2007;**92**:229–33
73. Su LK, Steinbach G, Sawyer JC, Hindi M, Ward PA, Lynch PM. Genomic rearrangements of the APC tumor-suppressor gene in familial adenomatous polyposis. *Hum Genet* 2000;**106**:101–7
74. Messiaen L, Vogt J, Bengesser K, Fu C, Mikhail F, Serra E, Garcia-Linares C, Cooper DN, Lazaro C, Kehrer-Sawatzki H. Mosaic type-1 NF1 microdeletions as a cause of both generalized and segmental neurofibromatosis type-1 (NF1). *Hum Mutat* 2011;**32**:213–9
75. Pasmant E, Sabbagh A, Spurlock G, Laurendeau I, Grillo E, Hamel MJ, Martin L, Barbarot S, Leheup B, Rodriguez D, Lacombe D, Dollfus H, Pasquier L, Isidor B, Ferkal S, Soulier J, Sanson M, Dieux-Coeslier A, Bieche I, Parfait B, Vidaud M, Wolkenstein P, Upadhyaya M, Vidaud D, Members of the NF France Network. NF1 microdeletions in neurofibromatosis type 1: from genotype to phenotype. *Hum Mutat* 2010;**31**:E1506–18
76. Chen JM, Stenson PD, Cooper DN, Ferec C. A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. *Hum Genet* 2005;**117**:411–27
77. Stewart C, Kural D, Stromberg MP, Walker JA, Konkel MK, Stutz AM, Urban AE, Grubert F, Lam HY, Lee WP, Busby M, Indap AR, Garrison E, Huff C, Xing J, Snyder MP, Jorde LB, Batzer MA, Korbel JO, Marth GT, 1000 Genomes Project. A comprehensive map of mobile element insertion polymorphisms in humans. *PLoS Genet* 2011;**7**:e1002236
78. Wimmer K, Callens T, Wernstedt A, Messiaen L. The NF1 gene contains hotspots for L1 endonuclease-dependent de novo insertion. *PLoS Genet* 2011;**7**:e1002371
79. Suzuki J, Yamaguchi K, Kajikawa M, Ichiyanagi K, Adachi N, Koyama H, Takeda S, Okada N. Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. *PLoS Genet* 2009;**5**:e1000461
80. Bragg DC, Mangkalaphiban K, Vaine CA, Kulkarni NJ, Shin D, Yadav R, Dhakal J, Ton ML, Cheng A, Russo CT, Ang M, Acuna P, Go C, Franceour TN, Multhaupt-Buell T, Ito N, Muller U, Hendriks WT, Breakefield XO, Sharma N, Ozelius LJ. Disease onset in X-linked dystonia-parkinsonism correlates with expansion of a hexameric repeat within an SVA retrotransposon in TAF1. *Proc Natl Acad Sci USA* 2017;**114**:E11020–28
81. Reyes CJ, Laabs BH, Schaake S, Luth T, Ardicoglu R, Rakovic A, Grutz K, Alvarez-Fischer D, Jamora RD, Rosales RL, Weyers I, Konig IR, Bruggemann N, Klein C, Dobricic V, Westenberger A, Trinh J. Brain regional differences in hexanucleotide repeat length in X-linked dystonia-parkinsonism using nanopore sequencing. *Neurol Genet* 2021;**7**:e608
82. Li D, McIntosh CS, Mastaglia FL, Wilton SD, Aung-Htut MT. Neurodegenerative diseases: a hotbed for splicing defects and the potential therapies. *Transl Neurodegener* 2021;**10**:16
83. Woodcock J, Marks P. Drug regulation in the era of individualized therapies. *N Engl J Med* 2019;**381**:1678–80
84. Stacey SN, Kehr B, Gudmundsson J, Zink F, Jonasdottir A, Gudjonsson SA, Sigurdsson A, Halldorsson BV, Agnarsson BA, Benediktsdottir KR, Aben KK, Vermeulen SH, Cremers RG, Panadero A, Helfand BT, Cooper PR, Donovan JL, Hamdy FC, Jinga V, Okamoto I, Jonasson JG, Tryggvadottir L, Johannsdottir H, Kristinsdottir AM, Masson G, Magnusson OT, Iordache PD, Helgason A, Helgason H, Sulem P, Gudbjartsson DF, Kong A, Jonsson E, Barkardottir RB, Einarsson GV, Rafnar T, Thorsteinsdottir U, Mates IN, Neal DE, Catalona WJ, Mayordomo JI, Kiemeney LA, Thorleifsson G, Stefansson K. Insertion of an SVA-E retrotransposon into the CASP8 gene is associated with protection against prostate cancer. *Hum Mol Genet* 2016;**25**:1008–18
85. Pfaff AL, Bubb VJ, Quinn JP, Koks S. Reference SVA insertion polymorphisms are associated with Parkinson's disease progression and differential gene expression. *NPJ Parkinsons Dis* 2021;**7**:44
86. Payer LM, Steranka JP, Yang WR, Kryatova M, Medabalimi S, Ardeljan D, Liu C, Boeke JD, Avramopoulos D, Burns KH. Structural variants caused by Alu insertions are associated with risks for many human diseases. *Proc Natl Acad Sci USA* 2017;**114**:E3984–92