Bridging the gap: Short structural variants in the genetics of anorexia nervosa

Natasha Berthold BBMS (Hons)\textsuperscript{1,2,3} | Julia Pytte BBMS (Hons)\textsuperscript{2,3} | Cynthia M. Bulik PhD\textsuperscript{4,5,6} | Monika Tschochner PhD\textsuperscript{1} | Sarah E. Medland PhD\textsuperscript{7} | Patrick Anthony Akkari PhD\textsuperscript{2,8,9,10}

\textsuperscript{1}School of Nursing, Midwifery, Health Sciences & Physiotherapy, University of Notre Dame Australia, Fremantle, Western Australia, Australia
\textsuperscript{2}Perron Institute for Neurological and Translational Science, Nedlands, Western Australia, Australia
\textsuperscript{3}School of Human Sciences, University of Western Australia, Crawley, Western Australia, Australia
\textsuperscript{4}Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
\textsuperscript{5}Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
\textsuperscript{6}Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
\textsuperscript{7}QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia
\textsuperscript{8}Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, Western Australia, Australia
\textsuperscript{9}Centre for Neuromuscular and Neurological Disorders, University of Western Australia, Nedlands, Western Australia, Australia
\textsuperscript{10}Department of Neurology, Duke University, Durham, North Carolina

Correspondence
Patrick Anthony Akkari, Perron Institute for Neurological and Translational Science, Nedlands, Western Australia, Australia.
Email: anthony.akkari@perron.uwa.edu.au

Funding information
Sarah E. Medland was supported by an Australian National Health and...
INTRODUCTION

Anorexia nervosa (AN) is a complex metabo-psychiatric disorder, and novel approaches are required to further elucidate its etiology. AN mortality is six times higher than the general population (Wonderlich et al., 2020). Standard treatment for adult AN patients combines renourishment, psychotherapy, and medications targeting related comorbidities (Kaye & Bulik, 2021). No biologically targeted treatments exist, and treatment efficacy is low (Kaye & Bulik, 2021). Risk factors for AN have been identified; however, mechanisms underlying heterogeneity in clinical presentation (e.g., restricting vs. binge eating/purging) remain to be clarified. Modern genetics uses several methods to estimate the contribution of genetics to a trait (i.e., heritability). Heritability estimates are derived from family (\(h^2_{\text{family}}\)) and twin (\(h^2_{\text{twin}}\)) studies, which provide an estimate of the total contribution of genetics to the trait, and single nucleotide polymorphism (SNP) heritability (\(h^2_{\text{SNP}}\)) estimates, which provide an estimation of the specific contribution of common single nucleotide variations to the trait. Familial and twin studies have yielded \(h^2_{\text{family}}\) and \(h^2_{\text{twin}}\) estimates of ~64% and ~57% for AN, respectively, indicating a notable genetic contribution to the disorder. Two separate genome-wide association studies (GWAS) have identified a total of nine loci significantly associated with AN. The largest GWAS reported a \(h^2_{\text{SNP}}\) of 11%–17% and identified eight loci associated with AN, with cellular adhesion molecule 1 (CADM1) and NCK interacting protein with SH3 domain (NCKIPSD) as the nearest genes to the SNPs in the top two hits (Watson et al., 2019). We address the considerable gap between \(h^2_{\text{family}}/h^2_{\text{twin}}\) estimates and \(h^2_{\text{SNP}}\) estimates, which provide an estimation of the specific contribution of common single nucleotide variations to the trait. Familial and twin studies have yielded \(h^2_{\text{family}}\) and \(h^2_{\text{twin}}\) estimates of ~64% and ~57% for AN, respectively, indicating a notable genetic contribution to the disorder. Two separate genome-wide association studies (GWAS) have identified a total of nine loci significantly associated with AN. The largest GWAS reported a \(h^2_{\text{SNP}}\) of 11%–17% and identified eight loci associated with AN, with cellular adhesion molecule 1 (CADM1) and NCK interacting protein with SH3 domain (NCKIPSD) as the nearest genes to the SNPs in the top two hits (Watson et al., 2019). We address the considerable gap between \(h^2_{\text{family}}/h^2_{\text{twin}}\) estimates and \(h^2_{\text{SNP}}\) (Manolio et al., 2009). The most likely explanation is that the variants that account for the heritability gap may be more informative than SNPs and are unable to be detected by GWAS (Wainschtein et al., 2021). Short structural variants (SSVs) are sequences of DNA 2–50 base pairs in length and are multiallelic, meaning that more than two variations exist within the population (Roses et al., 2016). SSVs have individual and synergistic effects on molecular biological functioning, including altering transcription rates of genes and affecting protein folding (Mis et al., 2017). Fine-mapping SSVs in and around GWAS-identified loci is a potential method of bridging this heritability gap. GWAS-identified loci are viable candidates for identifying SSVs because informative SSVs are often located in regions surrounding the lead SNPs identified in GWAS (Gymrek et al., 2016). Accordingly, using GWAS data to select candidate SSVs is a time- and cost-effective method, particularly for initial investigations; future investigations could utilize whole genome sequencing data to identify an increased number of SSVs in order to uncover greater heritability. The informative power of SSVs has been demonstrated in complex disorders including amyotrophic lateral sclerosis (ALS), late onset Alzheimer’s disease and schizophrenia (Fotinog et al., 2019; Pytte, Anderton, et al., 2020; Pytte, Flynn, et al., 2020; Roses et al., 2016; Theunissen et al., 2021). Exploring SSVs is a viable approach to further explicate genetic contributions to AN.

GWAS AS A STARTING POINT

AN GWAS provide a starting point to initiate investigating SSVs in AN. GWAS examine evidence for association between a trait and common SNPs across the genome. Association with a given SNP indicates an association exists within the surrounding genomic region and does not suggest the SNP is a coding variant (Tam et al., 2019). GWAS are powerful tools for interrogating heritability in complex traits and are followed by downstream molecular interrogations. Two GWAS have identified a total of nine genetic loci associated with AN (Duncan et al., 2017; Watson et al., 2019). The first GWAS (\(N_{\text{cases}} = 3495; N_{\text{controls}} = 10,982\)) revealed a single genome-wide significant locus on chromosome 12, associated with lead SNP rs4622308 (Duncan et al., 2017). The locus had been previously associated with rheumatoid arthritis and type 1 diabetes, with autoimmune-associated loci reported in surrounding regions (Barrett et al., 2009; Okada et al., 2014). Increasing sample sizes yielded a second GWAS (\(N_{\text{cases}} = 16,992; N_{\text{controls}} = 55,525\)) that revealed eight additional significant loci (Table 1); however, the initial locus identified by Duncan et al. (2017) was not replicated (Watson et al., 2019). The lead SNPs for the first five genetic loci were intronic (located in non-coding regions of a gene) and the lead SNPs for the last three genetic loci were intergenic (located in regions of the gene between genes; Table 1). No lead SNPs were located in exonic regions, the coding regions of the gene. Linkage disequilibrium analyses, the analysis of nonrandom co-occurrence, revealed significant positive genetic correlations between AN and obsessive-compulsive disorder, anxiety disorders, schizophrenia, and major depressive disorder (Watson et al., 2019), reinforcing the psychiatric nature of AN (Duncan et al., 2017; Hübel et al., 2021). The study also reaffirmed genetic predisposition to an AN-prone metabolic profile (Duncan et al., 2017; Watson et al., 2019). Consequently, AN is hypothesized to include genetic predispositions to both psychiatric and metabolic traits (Hübel et al., 2021). The GWAS findings have identified likely informative regions within the genome containing causal genetic architecture (Table 1). Characterizing SSVs is a potential method to extend GWAS.
The eight newly identified genetic loci associated with anorexia nervosa

| CHR | Lead SNP   | Nearest gene   | Functions                                                                                                                                                                                                 | pValue    |
|-----|------------|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 3   | rs9821797  | NCKIPSD        | Growth and cellular signaling in dendrites and sarcomeres; stress fiber formation (Cho et al., 2013).                                                                                                       | 6.99 × 10⁻¹⁵ |
| 11  | rs6589488  | CADM1          | Cellular adhesion; neural network formation; synaptic formation and number (Jin et al., 2019).                                                                                                               | 6.31 × 10⁻¹¹ |
| 2   | rs2287348  | ASB3 and ERLEC1| ASB3: Phosphorylation and ubiquitination (Chung et al., 2005). ERLEC1: N-glycan binding (Cruciati et al., 2006).                                                                                           | 5.62 × 10⁻⁹  |
| 10  | rs2008387  | MGMT           | Alkylation agent removal (Yu et al., 2020).                                                                                                                                                                   | 1.73 × 10⁻⁶  |
| 3   | rs9874207  | FOXP1          | Transcription factor (Siper et al., 2017).                                                                                                                                                                   | 2.05 × 10⁻⁸  |
| 1   | rs10747478 | PTBP2          | RNA splicing in neuronal cell maturation (Romanelli et al., 2013).                                                                                                                                         | 3.13 × 10⁻⁶  |
| 5   | rs370838138| CDH10          | Sodium dependent intercellular adhesion (Kools et al., 1999).                                                                                                                                             | 3.17 × 10⁻⁸  |
| 3   | rs13100344 | NSUN3          | Cytosylation of 5-formylcytidine at position 34 of methionine transfer RNA (Nakano et al., 2016).                                                                                                         | 4.12 × 10⁻³  |

Note: Eight genetic loci were identified in the (2019 ANGI GWA by gene proximity to each lead SNP. Lead SNP was determined as the most strongly associated. p-Value was considered significant (after Bonferroni adjustment) at ≤ 0.05. The major functions for each of the nearest genes to the lead SNP have been described in column 4.

Abbreviations: ASB3, ankyrin repeat and SOCS box containing 3; CADM1, cell adhesion molecule 1; CDH10, cadherin 10; CHR, chromosome; ERLEC1, endoplasmic reticulum lectin 1; FOXP1, forkhead box P1; MGMT, O-6-methylguanine-DNA methyltransferase; NCKIPSD, NCK interacting protein with SH3 domain; NSUN3, NOP2/Sun RNA methyltransferase 3; PTBP2, polyypyrimidine tract binding protein 2; RNA, ribonucleic acid; rs, reference SNP accession number; SNP, single nucleotide polymorphism.

Source: Adapted from “Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa” by Watson et al., 2019; Nature Genetics, 5(8), pp. 1207–1214 (doi: 10.1038/s41588-019-0439-2).

findings to further interrogate genetic factors contributing to AN (Chiba-Falek, 2017; Roses et al., 2016; van Rheenen et al., 2016).

### 3 | EXTENDING GWAS HITS BY STUDYING SSVs

SSVs are multiallelic and include a variety of sizes and sequence arrangements. We focus on short tandem repeats, which are sequence motifs of 1–6 nucleotides repeated numerous times, for example a dinucleotide repeat may contain a repeat of thymine (T) and adenine (A), such as TATATA, while a penta-nucleotide repeat might comprise of a repeat of a sequence of T, A, guanine (G) and cytosine (C), such as TAGGCTAGGCTAGGC (Roses et al., 2016). Short tandem repeats are highly mutable in nature and accumulating evidence suggests that this class is the most variable of SSVs and consequently the most likely to be functionally relevant (Gymrek et al., 2016). SSVs can be located within a gene but are also frequently found within noncoding or regulatory regions (Gharesouran et al., 2021; Theunissen et al., 2020). SSVs can contribute to more biological variability than SNPs as their multiallelic nature engenders them with a greater likelihood of producing diverse results (Chaisson et al., 2019; Gymrek et al., 2016; Roses et al., 2016). Ribonucleic acid (RNA) expression studies suggest that the functional impact of SSVs is significantly greater than SNPs, even though SSVs are less frequent (Jakubosky, D’Antonio, et al., 2020; Jakubosky, Smith, et al., 2020). Growing evidence for informative power of SSVs motivates exploration of their contribution to AN.

SSV functional mechanisms include influencing gene expression, regulation of gene expression, and RNA splicing. Such influences may modify disorder presentation, indicate risk, and influence therapeutic response among patients (Gharesouran et al., 2021; Mahmoud et al., 2019; Pytte, Anderton, et al., 2020; Pytte, Flynn, et al., 2020; Roses et al., 2016; Theunissen et al., 2020). Technically, SSVs located within promoter regions can regulate gene expression by modifying histones, which alters accessibility of the region to transcriptional machinery, and influences promoter binding specificity (Fotsing et al., 2019; Gharesouran et al., 2021). The variation of SSV length within intronic regions can alter secondary RNA structure, affecting the availability and accessibility of the region to splicing factor binding, altering transcription efficiency. Additionally, SSV polymorphisms can alter the binding sites of intronic and exonic splicing enhancers or silencers to favor one or the other, thus modifying splicing to alter the final messenger RNA (mRNA) transcript (Gymrek et al., 2016). Downstream effects of this have been observed by SSVs such as the G₃C₂ repeat expansion in chromosome 9 open reading frame 72-Smith–Magenis chromosome region 8 complex subunit (C9orf72-SMCR8 complex subunit:C9orf72) and CAG trinucleotide repeat expansion in Ataxin 2 (ATXN2) for ALS (Mis et al., 2017; Van Damme et al., 2011). Both SSVs alter the native structure of the respective protein product leading to aberrant binding, producing truncated protein products, or toxic protein aggregates, which has downstream effects on neurobiological functions that contribute to disease pathogenesis (Mis et al., 2017; Van Damme et al., 2011). A second example of the informative power of SSVs in ALS is in the CA dinucleotide Stathmin-2 (STMN2) intronic repeat (Theunissen et al., 2021). Here, the presence of two long alleles in a cohort of sporadic ALS (N₉ₙₐ₃ₑₙₛ = 321) was associated with increased disease risk, earlier age of onset, and decreased survival duration for
cases of bulbar onset, and disease severity compared with controls \((N_{\text{controls}} = 332; \text{Theunissen et al., 2021})\). In another sporadic ALS cohort \((N_{\text{cases}} = 67)\), the presence of two long alleles was associated with lower ALS functional rating scale scores and revealed variation in expression levels of Statmin-2 mRNA between sporadic ALS cases and control laser-captured spinal motor neurons based on the CA genotype \((\text{Theunissen et al., 2021})\). With such effects on gene expression and regulation, and the ability to act as genetic markers, uncharacterized SSVs potentially possess considerable power in reducing the heritability gap in AN and may further our understanding of AN etiology, mechanisms, and heterogeneity.

### 4 | SSVs ELUCIDATING UNDERLYING GENETIC MECHANISMS OF AN

Fine-mapping poorly characterized regions in and around AN GWAS-associated loci are likely to elucidate AN heritability, risk factors, and novel pathogenic mechanisms. We utilized an SSV bioinformatics algorithm, “SSV evaluation system” \((\text{Saul et al., 2016})\) to prioritize candidate SSVs in AN-associated genetic loci NCKIPSD and CADM1 for an initial genetic investigation (Table 2). These loci were prioritized based on association strength reported in the latest AN GWAS (Table 1) and were considered valid for investigation. Both genes have functions that could be biologically relevant to the pathological mechanisms of AN. The NCKIPSD protein possesses several functions across numerous tissues, centered on its signal transduction abilities. Within the context of its role in the nervous system, its function is uncharacterized SSVs potentially possess considerable power in reducing the heritability gap in AN and may further our understanding of AN etiology, mechanisms, and heterogeneity.

**TABLE 2** Initial short structural variants prioritized by the short structural variant evaluation algorithm for future characterization and investigation in anorexia nervosa case/control studies

| Gene | rs Number | Symbol | Gene feature |
|------|-----------|--------|--------------|
| CADM1 | rs11358670 | 28T | Intronic variant |
| CADM1 | rs58589028 | 29T | Intronic variant |
| CADM1 | rs61694033 | 25A | Intronic variant |
| CADM1 | rs72085573 | 20T | 5' Intergenic region |
| CADM1 | rs140815983 | 15A | 3' Intergenic region |
| CADM1 | rs147798460 | 32T | 3' Intergenic region |
| CADM1 | rs148209064 | 33A | Intronic variant |
| CADM1 | rs386374979 | 29T | 3' UTR downstream contiguous variant |
| CADM1 | rs747352768 | 11TGG | Exonic variant (Coding exon 8) |
| CADM1 | rs991408884 | 33T | 3' UTR variant |
| NCKIPSD | rs71074264 | 24T | Intronic variant (NCKIPSD and LINC02585) |
| NCKIPSD | rs71627345 | 21A | Intronic variant (NCKIPSD and LINC02585) |
| NCKIPSD | rs37547983 | 5ACAA | Intronic variant |
| NCKIPSD | rs377051084 | 9AGGG | Intronic variant |
| NCKIPSD | rs545029045 | 20AC | Intronic variant |
| NCKIPSD | rs757842104 | 31T | Downstream variant IP6K2 |
| NCKIPSD | rs34837885 | 8AAAT | Intronic variant IP6K2; upstream variant NCKIPSD |
| NCKIPSD | rs35746542 | 24T | Intronic variant IP6K2; upstream variant NCKIPSD |
| NCKIPSD | rs67509214 | 28A | Intronic variant IP6K2; upstream variant NCKIPSD |
| NCKIPSD | rs71074266 | 22A | Intronic variant |

Note: The 20 SSVs prioritized by the SSV evaluation algorithm (designed by Saul et al., 2016) as candidates for further investigation to elucidate potential roles in AN risk. Ten SSVs have been prioritized for each genetic candidate loci, NCKIPSD, and CADM1. The column titled “Gene” refers to which candidate loci the SSV was reported for. The column labeled “rs Number” refers to the unique identifier supplied by the current human reference genome for that variant. The column titled “Symbol” refers to the most frequently occurring variation of that SSV according to the Allele Frequency Aggregator Project. The final column, titled “Gene Feature,” refers to the functional property of the region of the genome in which SSV is situated. The symbol, rs number and gene feature listed here are as reported in the current human reference genome GRCh38.p13.

Abbreviations: A, adenine; C, cytosine; CADM1, cell adhesion molecule 1; G, guanine; IP6K2, inositol hexakisphosphate kinase 2; LINC02585, long intergenic non-protein coding RNA 02585; NCKIPSD, NCK interacting protein with SH3 domain; rs, reference SNP accession number; T, thymine; UTR, untranslated region.
for causal SSVs, enabling efficient prioritization of potential trait modifying SSVs in complex disorders (Saul et al., 2016). Implicated SSVs are prioritized and further interrogated via molecular biology techniques in case/control association studies (Saul et al., 2016). SSV polymorphisms are initially identified via Sanger sequencing and quantified via fragment analysis with end-labeled fluorescent primers in small control cohorts. Sanger sequencing allows sequence visualization at a single base pair resolution—an effective method for initially determining polymorphisms of an SSV. Several characteristics make Sanger sequencing less suited to genotyping large cohorts as it is particularly prone to errors when sequencing repetitive stretches of identical nucleic acids, such as SSVs, and can be time consuming. Fragment analysis provides an empirical measure of the total size of the genomic region of interest in base pairs, thus catering for repeats of any length with high accuracy and can be performed in a time efficient and high-throughput manner, rendering it suitable for genotyping large cohorts for confirmed SSV polymorphisms. This approach has been employed effectively in multiple studies, such as Theunissen et al., (2021), which identified and characterized polymorphisms for the CA dinucleotide repeat in the STMN2 gene prior to performing the association studies between the SSV and ALS, presenting it as an exciting future avenue in AN genetics research (Pytte, Anderton, et al., 2020; Pytte, Flynn, et al., 2020; Theunissen et al., 2021). The SSVs we select for the NCKIPSD and CADM1 genetic loci are predominantly noncoding intronic or intergenic variants, as functionally relevant SSVs occur more frequently in noncoding regions of the genome (Table 2; Fotsing et al., 2019). The variant rs747352768 is the only SSV, between both target loci, which falls in a coding exon region (Table 2). The outlined potential of these SSVs to have impact on the CADM1 and NCKIPSD genetic loci that are potentially functionally relevant to AN pathiology makes this a valid direction of investigation.

5 | CONCLUSION AND FUTURE DIRECTIONS

AN is severe and potentially fatal with an underexplored heterogeneous etiology. No medications exist that target the underlying biology of AN (Kaye & Bulik, 2021). Outcomes could be improved with increasing understanding of pathogenic mechanisms responsible for the development of AN (Kaye & Bulik, 2021). GWAS have been reported and larger studies are underway to expand upon the GWAS findings and deepen understandings of the heritability of eating disorders more broadly, including bulimia nervosa and binge-eating disorder (Bulik et al., 2021). Accumulating evidence indicates that SSVs have diverse functional effects on genotypic variability (Fotsing et al., 2019; Pytte, Anderton, et al., 2020; Pytte, Flynn, et al., 2020; Theunissen et al., 2021). Successful SSV mapping has the potential to extend genomic discovery in AN, unveil undetected heritability, elucidate novel pathogenic mechanisms, and identify targets for new therapies, with the long-term objective of reducing mortality and improving the quality of life for individuals with AN.

ACKNOWLEDGMENT

Open access publishing facilitated by The University of Notre Dame Australia, as part of the Wiley - The University of Notre Dame Australia agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST

Cynthia M. Bulik reports: Shire (grant recipient, Scientific Advisory Board member); Idorsia (consultant); Lundbeckfonden (grant recipient); Pearson (author, royalty recipient); Equip Health Inc. (Clinical Advisory Board). The other authors report no conflicts.

AUTHOR CONTRIBUTIONS

Natasha Berthold: Conceptualization; investigation; methodology; writing — original draft; writing — review and editing. Julia Pytte: Methodology; supervision; writing — original draft; writing — review and editing. Cynthia M. Bulik: Supervision; writing — review and editing. Monika Tschochner: Supervision; writing — review and editing. Sarah E. Medland: Supervision; writing — review and editing. P. Anthony Akkari: Methodology; supervision; writing — original draft; writing — review and editing.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Natasha Berthold https://orcid.org/0000-0001-7432-3353

REFERENCES

Barrett, J. C., Clayton, D. G., Concannon, P., Akolkar, B., Cooper, J. D., Erlich, H. A., Julier, C., Morahan, G., Nerup, J., Nierras, C., Plagnol, V., Pociot, F., Schuijlenburg, H., Smyth, D. J., Stevens, H., Todd, J. A., Walker, N. M., Rich, S. S., & Type 1 Diabetes Genetics Consortium. (2009). Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nature Genetics, 41(6), 703–707. https://doi.org/10.1038/ng.381

Bulik, C. M., Thornton, L. M., Parker, R., Kennedy, H., Baker, J. H., MacDermod, C., Guintivano, J., Cleland, L., Miller, A. L., Harper, L., Larsen, J. T., Yilmaz, Z., Grove, J., Sullivan, P. F., Petersen, L. V., Jordan, J., Kennedy, M. A., & Martin, N. G. (2021). The Eating Disorders Genetics Initiative (EDGI): Study protocol. BMC Psychiatry, 21(1), 234. https://doi.org/10.1186/s12888-021-03212-3

Chaissen, M. J. P., Sanders, A. D., Zhao, X., Malhotra, A., Porubsky, D., Rausch, T., Gardner, E. J., Rodriguez, O. L., Guo, L., Collins, R. L., Fan, X., Wen, J., Handsaker, R. E., Fairley, S., Kronenberg, Z. N., Kong, X., Hormozdari, F., Lee, D., Wenger, A. M., ... Lee, C. (2019). Multi-platform discovery of haplotype-resolved structural variation in human genomes. Nature Communications, 10(1), 1784. https://doi.org/10.1038/s41467-018-08148-z

Chiba-Falek, O. (2017). Structural variants in SNCA gene and the implication to synucleinopathies. Current Opinion in Genetics & Development, 44, 110–116. https://doi.org/10.1016/j.gde.2017.01.014

Cho, I. H., Kim, D. H., Lee, M. J., Bae, J., Lee, K. H., & Song, W. K. (2013). SPIN90 phosphorylation modulates spine structure and synaptic function. PLoS One, 8(1), e54276. https://doi.org/10.1371/journal.pone.0054276

Chung, A. S., Guan, Y.-J., Yuan, Z.-L., Albina, J. E., & Chin, Y. E. (2005). Ankyrin repeat and SOCS box 3 (ASB3) mediates ubiquitination and...
heritability in sALS. *Frontiers in Neuroscience*, 14, 47. https://doi.org/10.3389/fnins.2020.00047

Van Damme, P., Veldink, J. H., van Blitterswijk, M., Corveleyn, A., van Vught, P. W., Thijs, V., Dubois, B., Matthijs, G., van den Berg, L., & Robberecht, W. (2011). Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2. *Neurology*, 76(24), 2066–2072.

van Rheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Diekstra, F. P., Pult, S. L., van der Spek, R., Vösa, U., de Jong, S., Robinson, M. R., Yang, J., Fogh, I., van Doormaal, P., Tazelaar, G. H., Koppers, M., Blokhuis, A. M., Sproviero, W., Jones, A. R., Kenna, K. P., ... Veldink, J. H. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature Genetics*, 48(9), 1043–1048. https://doi.org/10.1038/ng.3622

Wainschtein, P., Jain, D., Zheng, Z., TOPMed Anthropometry Working Group, Adrienne Cupples, L., Shadyab, A. H., McKnight, B., Shoemaker, B. M., Mitchell, B. D., Psaty, B. M., Kooperberg, C., Liu, C.-T., Albert, C. M., Roden, D., Chasman, D. I., Darbar, D., Lloyd-Jones, D. M., Arnett, D. K., & Visscher, P. M. (2021). Recovery of trait heritability from whole genome sequence data. bioRxiv, 588020. https://doi.org/10.1101/588020

Wonderlich, S. A., Bulik, C. M., Schmidt, U., Steiger, H., & Hoek, H. W. (2020). Severe and enduring anorexia nervosa: Update and observations about the current clinical reality. *International Journal of Eating Disorders*, 53(8), 1303–1312. https://doi.org/10.1002/eat.23283

Yu, W., Zhang, L., Wei, Q., & Shao, A. (2020). O6-Methylguanine-DNA methyltransferase (MGMT): Challenges and new opportunities in glioma chemotherapy. *Frontiers in Oncology*, 9, 1547. https://doi.org/10.3389/fonc.2019.01547

How to cite this article: Berthold, N., Pytte, J., Bulik, C. M., Tschochner, M., Medland, S. E., & Akkari, P. A. (2022). Bridging the gap: Short structural variants in the genetics of anorexia nervosa. *International Journal of Eating Disorders*, 55(6), 747–753. https://doi.org/10.1002/eat.23716

Yu, W., Zhang, L., Wei, Q., & Shao, A. (2020). O6-Methylguanine-DNA methyltransferase (MGMT): Challenges and new opportunities in glioma chemotherapy. *Frontiers in Oncology*, 9, 1547. https://doi.org/10.3389/fonc.2019.01547