

Universiteit
Leiden
The Netherlands

Facts of aggression

Roetman, P.J.

Citation
Roetman, P. J. (2021, June 9). *Facts of aggression*. Retrieved from https://hdl.handle.net/1887/3186436

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3186436

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle https://hdl.handle.net/1887/3186436 holds various files of this Leiden University dissertation.

Author: Roetman, P.J.
Title: Facts of aggression
Issue Date: 2021-06-09

CHAPTER 5

GENOMICS OF HUMAN AGGRESSION: CURRENT STATE OF GENOME-WIDE STUDIES AND AN AUTOMATED SYSTEMATIC REVIEW TOOL

AUTHORS: PETER J. ROETMAN*, VERONIKA V. ODINTSOVA, HILL F. IP, RENÉ POOL, CAMIEL M. VAN DER LAAN, KLODIANA-DAPHNE K. TONA, ROBERT R.J.M. VERMEIREN, DORRET I. BOOMSMA

*SHARED FIRST AUTHORSHIP

PSYCHIATRIC GENETICS (2019)

ABSTRACT

Objectives: There are substantial differences, or variation, between humans in aggression, with its molecular genetic basis mostly unknown. This review summarizes knowledge on the genetic contribution to variation in aggression with three foci: 1) a comprehensive overview of reviews on the genetics of human aggression, 2) a systematic review of genome-wide association studies (GWASs), and 3) an automated tool for the selection of literature based on supervised machine learning.

Methods: The phenotype definition "aggression" (or "aggressive behaviour", or "aggression-related traits") included anger, antisocial behaviour, Conduct Disorder, and Oppositional Defiant Disorder. The literature search was performed in multiple databases, manually and using a novel automated selection tool, resulting in 18 reviews and 17 GWASs of aggression.

Results: Heritability estimates of aggression in children and adults are around 50%, with relatively small fluctuations around this estimate. In 17 GWASs, 817 variants were reported as suggestive ($P \leq 1.0E^{-05}$), including 10 significant associations ($P \leq 5.0E^{-08}$). Nominal associations ($P \leq 1E^{-05}$) were found in gene-based tests for genes involved in immune, endocrine, and nervous systems. Associations were not replicated across GWASs. A complete list of variants and their position in genes and chromosomes is available online. The automated literature search tool produced literature not found by regular search strategies.

Conclusion: Aggression in humans is heritable, but its genetic basis remains to be uncovered. No sufficiently large genome-wide association studies have been carried out yet. With increases in sample size, we expect aggression to behave like other complex human traits for which GWAS has been successful.

INTRODUCTION

Aggression is a common type of human behaviour (Tuvblad & Baker, 2011) and is considered a characteristic that is shared by all humans (Veroude, Zhang-James, et al., 2016). The propensity for aggression, however, varies considerably between individuals. This paper addresses the question to what extent the variation that is seen for aggression has a genetic cause. Broadly, aggression can be defined as a behaviour that intends to cause physical or emotional harm to others (Anderson & Bushman, 2002). High levels of aggression are also seen in individuals with severe mental disorders (e.g., autism, bipolar disorder, schizophrenia) as well as in patients with (rare) Mendelian disorders (Zhang-James et al., 2019). Because of the large impact of aggression on the affected individual, their families, their environment, and society as a whole, there is a substantial interest in studying aggression from a wide range of disciplines. In this context, one goal is to unravel the aetiology of aggression by identifying environmental exposures and biomarkers, including genetic factors, epigenetic marks, and metabolites, that could function as predictors of (excessive) aggression (Boomsma, 2015b).

Research often focuses on the pathological aspects of aggressive behaviour, while aggression does not solely have negative consequences or outcomes. Under certain circumstances, aggressive behaviour is beneficial to individuals, for example when competing for limited resources, like food or mates (Lindenfors & S.Tullberg, 2011), or achieving social dominance (Little et al., 2007). Aggression can further be a powerful deterrent against aggressive behaviour from others. Because both high and low levels of aggression can be detrimental to survival and procreation, it has been postulated that aggression is under stabilizing selection, implying that variation in aggression should show significant heritability. Substantial heritability estimates have indeed been reported in animals (Anholt & Mackay, 2012) and humans, as reviewed below.

Benefits of aggressive acts depend on the type of aggression, its success, environmental circumstances and also vary across cultures (Rubin, Bukowski, & Laursen, 2011). For example, predatory goal-oriented aggression has been associated with social dominance in some instances (Dodge, Lochman, Harnish, Bates, & Pettit, 1997; Hawley & Vaughn, 2003; Voulgaridou & Kokkinos, 2015), but this association seems to vary between groups that are more prosocial and groups that consist predominantly of individuals with disruptive behaviour problems (Wright, Giammarino, & Parad, 1986). A decrease in social status

can also result from aggression, in particular from reactive aggression, which is an uncontrolled type of aggression stemming from internal or external frustration. In reverse, after a conflict, proactive aggression is increased in the victorious party while the losing party is less likely to engage in another aggressive act (Penn, Zito, & Kravitz, 2010; Polman, Orobio de Castro, Koops, van Boxtel, & Merk, 2007). To differentiate between different outcomes of aggression, researchers have distinguished aggression subtypes (e.g., reactive vs. proactive; overt vs. covert), developmental stages (childhood vs. adolescent onset), and comorbidities (e.g., with internalizing problems or with attention deficit hyperactivity disorder (ADHD)). In summary, the outcomes and types of aggressive acts can differ greatly between persons and circumstances, and need not always be dysfunctional.

At the start of the 1990s, research on aggressive behaviour was given a new impulse by a seminal paper of Brunner et al. (1993), in which a Dutch pedigree was described where men exhibited impulsive aggression, arson, violence and borderline mental retardation. The family appeared to have a rare point mutation in the structural gene for monoamine-oxidase-A (MAOA) – which codes for an enzyme that is involved in the oxidative deamination of neurotransmitters like dopamine, serotonin and norepinephrine – resulting in a deficiency of the MAOA enzyme. A study, by Caspi et al. (2002), compared variants of the MAOA gene in children who experienced maltreatment and showed that children with the variant resulting in lower levels of the MAOA enzyme were more likely to develop antisocial behaviour. Efforts to replicate the latter finding have been contradictory, either without replication (Haberstick et al., 2005; Young et al., 2006) or with replication (Foley et al., 2004; Kim-Cohen et al., 2006; Nilsson, Åslund, Comasco, & Oreland, 2018). Nevertheless, the studies of Brunner and Caspi stressed the importance of biological factors in the development of aggression and antisocial behaviour. This instigated extensive efforts to study the genetic basis of aggression.

Enormous progress has been made with respect to technology in molecular biology and large-scale genotyping, as well as in the development of statistical methods for genetic association studies and polygenic scores for individual risk assessment, once sufficiently large genetic-association studies are available (Dudbridge, 2016). Costs for genotyping and sequencing of DNA, the epigenome and of RNA, and biomarker assessment, such as metabolomics, have steadily decreased, allowing for large studies, relating aggressive behaviour to genome,

epigenome, transcriptome and other biomarkers (Hagenbeek et al., 2016). Progress also has been made in characterizing the exposome, which reflects the totality of a person's environmental exposures in space and time (Wild, 2005).

Genome-wide association studies (GWAS) provide a conceptual framework to examine whether individual differences in aggression are associated with allelic differences in millions of single nucleotide polymorphisms (SNPs) across the genome (Visscher et al., 2017). Because a GWAS targets the entire human genome, it enables a data-driven approach to identify loci of interest. This hypothesis-free approach could potentially help researchers to overcome limits imposed by multifactorial nature of a trait and incomplete understanding of its physiological basis.

Here we synthesise knowledge deriving from studies on genetics of human aggression and variance in liability to aggression-related traits. Our review has three foci: 1) to give a comprehensive overview of reviews already done on genetics of human aggression, 2) to carry out a systematic review of GWAS studies on human aggression, and 3) to introduce an automated systematic review for the selection of relevant literature based on supervised machine learning. For consistency, in this review we will use the general term "aggression" (or "aggressive behaviour", or "aggression-related traits") to refer to the terminologies used by different authors (see Supplement S1), including anger, hostility dimensions, parent-reported child aggressive behaviour, physical aggression, antisocial behaviour (ASB), violent offending, conduct disorders (CD), oppositional defiant disorder (ODD), and antisocial personality disorder (ASPD).

METHODS

To optimize detection of the relevant literature for our review, we incorporated two strategies:

a) A "traditional" (manual) search strategy where search terms were used to extract the relevant articles from literature databases.
b) An automated screening with Automated Systematic Review Software (ASR) where relevant articles were detected via the utilization of machine learning algorithms and a software development platform.

Traditional approach

Search strategy

Search terms were developed by the authors based on prior literature and discussions with an expert librarian (J.W.S) from the LUMC. A literature search was performed in PubMed, Embase, Web of Science, Cochrane library, PsychInfo and Academic Search Premier with a comprehensive list of general search terms and medical subject headings (Supplement S2). Searches were conducted separately for reviews/meta-analyses and GWA studies. Searches included literature without a specific time limit and were conducted in mid-April 2019.

Selection criteria

A selection was made from all titles and abstracts that were found in the databases using pre-specified inclusion and exclusion criteria (see Table 1). Articles were included if they (1) were written in English and (2) focused on human aggression. Studies were excluded if (1) they focused on animals, or (2) general terms linked to "aggression/violent etc." did not refer to a psychological/ psychiatric perspective but rather to characteristics of disease (e.g., aggressive cancer), or (3) articles discussed only a single gene. Psychiatric disorders which incorporate acts of aggression and are highly correlated to aggression and antisocial lifestyles, like Oppositional Defiant Disorder (ODD), Conduct Disorder (CD), and Antisocial Personality Disorder (ASPD) were included. Papers referring to associations between genetic data and other (neuro)psychiatric disorders as main outcome (e.g., psychosis, borderline personality disorders, schizophrenia, bipolar disorder, anxiety, major depression, intellectual disability, Alzheimer's disease, autism, ADHD, addictions) as main outcome were excluded. This increased the probability that the genetic profile that we examined was not confounded due to high comorbidity of aggression with other psychiatric disorders. Papers referring to aggression from the perspective of victimization and bullying were excluded. The publications were reviewed independently by 2 authors (V.V.O and P.J.R.), and when in doubt other co-authors were consulted until consensus on inclusion was reached.

Selection procedure and analyses

The search on review/meta-analyses resulted in 1,713 records (see Figure 1). Duplicate entries were removed (N=27). Next, 1,660 records were excluded based on screening the titles and abstracts. In total, 26 potentially relevant

reviews were retrieved for a full-text screening. Studies that did not fulfil or only partially fulfilled our criteria were excluded from the analysis (N=12), leading to the inclusion of 14 articles. Four additional reviews were added through the automated selection, leading to a total of 18 articles – 13 targeted and 5 systematic reviews. These were organized into the following categories: review type (targeted or systematic), definition of aggression, type of reviewed studies (heritability, candidate gene, GWAS), population (children, adolescents, adults), quantity and period of the publications included in the reviews (parameters are made on the basis of reference lists with inclusion of publications on the aggression-related traits), described genes and main conclusions.

Table 1 Inclusion and exclusion criteria for the systematic review

Selection Criteria	Inclusion Criteria	Exclusion Criteria
Language	English	Non-English
Population	Human studies (all ages)	Animal studies
Use of term "aggression"	Psychological/psychiatric	Disease characteristics (e.g., aggressive cancer, aggressive form of somatic diseases etc)
		Victimization, victims of bullying
Psychiatric disorders	ODD, CD, ASPD	Other neuropsychiatric and psychiatric disorders (e.g., psychosis, anxiety etc)
Discussion of genes	At least 2 genes associated with aggression*	No genetic methods and information on genes associated with aggression

*This was done to exclude reviews focussing on a single candidate gene.

The search for GWASs on aggression resulted in 356 records. A total of 331 were excluded based on screening of the titles and abstracts. This led to the retrieval of 25 potentially relevant studies for full-text screening. Studies that did not fulfil or only partially fulfilled our criteria were excluded (N=8), leading to the inclusion of 17 GWAS articles. Three additional studies were selected from the automated selection, including 1 SNP-heritability and 2 linkage studies. The studies were analysed by phenotype, sample characteristics, SNPs or genetic variants associated with aggression-related traits at $p < 1E^{-05}$, genetic variants position in genes and chromosomes.

Several GWAS papers report findings on multiple (stratified) GWASs. Tielbeek et al. (2017) adjusted for the fact that they performed three genome-wide association meta-analyses (GWAMA) by setting the genome-wide significance

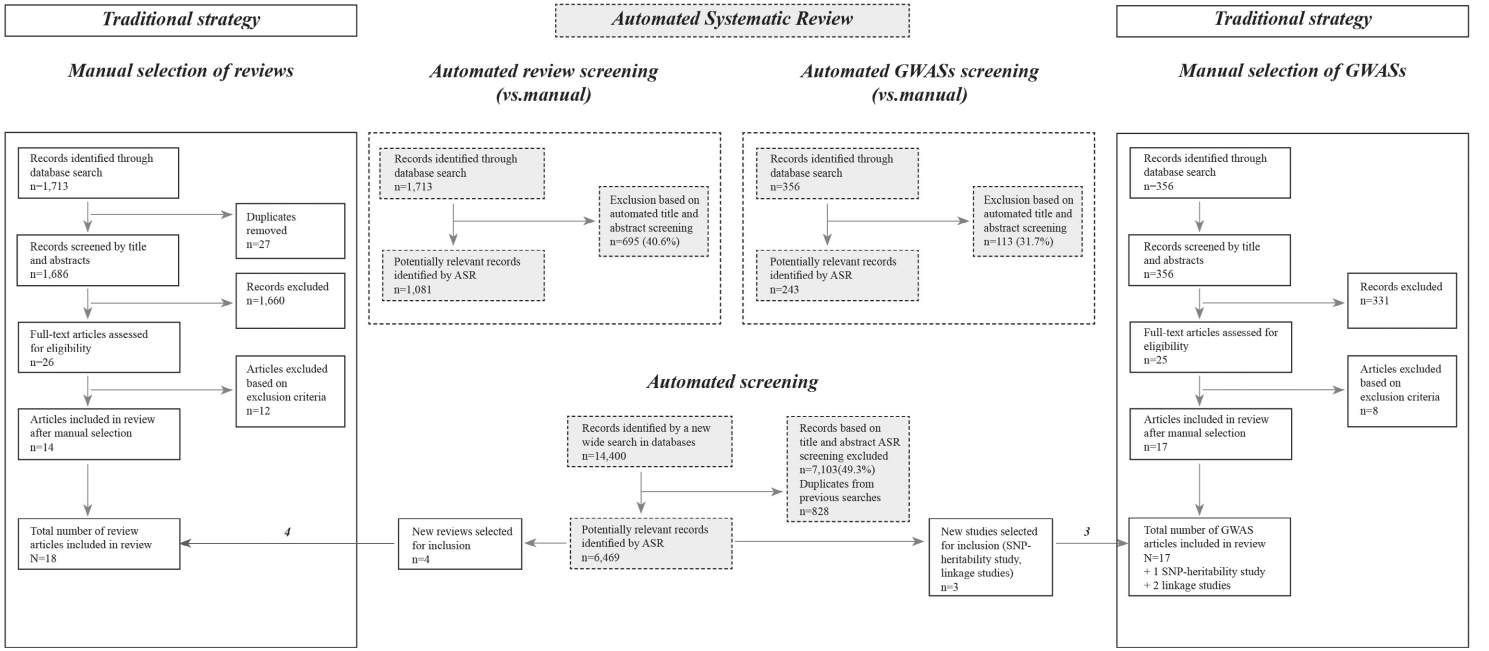
threshold at $p = 1.67E^{-08}$, whereas others did not apply such a correction. This threshold might be overly conservative as the GWAMAs are stratified, which makes the p-values non-independent across GWAMA. Therefore, we maintained a significance threshold of $p = 5.0E^{-08}$ for all studies, and denote any SNP with a p-value below this threshold as genome-wide significant. While the traditional threshold might be too lenient in this context, we note that, when discussing GWASs, the p-value of a SNP in any given study is of less relevance than replication across GWASs.

Automated titles and abstracts screening

In parallel with the manual selection of titles and abstracts, another selection was made with the use of an automated selection tool "Automated Systematic Review" (ASR) – software hosted at https://github.com (*Automated systematic reviews by using Deep Learning and Active Learning*, 2019). This software allows for automated in- and exclusion of articles for systematic reviews based on the titles and abstracts of articles. This enabled a comparison between "traditional" manual selection and the automated screening on performance characteristics (e.g., time spent on selection, false negative results). Furthermore, an additional selection was performed with the ASR on a large dataset of references to retrieve any new additional papers to our review, which would have been missed in the traditional search strategy (see Supplement S3).

We trained a model using ASR. To do so, the model requires a training set based on expert knowledge, consisting of papers that are either labelled relevant or non-relevant (labels 1 = included , 0 = not) (see Supplement S3: Figure S3.1). To study the operating characteristics of the ASR, we used a dataset (N=2,955) consisting of relevant and non-relevant papers on the genetics of human aggression, as labelled by researchers. From this labelled dataset of N=2,955, 500 records were repeatedly drawn at random as training sets. The number of relevant records in the training sets varied between 10 and 80 (e.g., 10 relevant records versus 490 non-relevant records), in increments of 10. These sets were used to train models to include relevant records and exclude non-relevant records. For each model we computed Receiver Operating Characteristic (ROC) parameters that were then used to select the optimal model (see Supplement S1: Table S3.1, Figure S3.2).

Figure 1 Flow diagram of literature selection



We selected the model that returned the lowest false positive rate (FPR) while allowing for a maximum false negative rate of FNR = 0.03 at most. Note that FNR = 0.03 corresponds with a true positive rate of TPR = 0.97.

We applied the optimal model to predict classification in different searches: (1) reviews of genetics of human aggression (1,713 records); (2) GWASs on human aggression (356 records); (3) searches 1 and 2 combined (2,069 records) to analyse parameters of automated selection in comparison to manual selection.

Training sets were provided to the ASR for the reviews on aggression (26 relevant records out of 1,713 [1.5%]) and the GWASs on aggression (25 relevant records out of 356 [7.0%]) (see Supplement S3: Table S3.2). The automated selection predicted 1,018 records out of 1,713 (59.4%) as relevant for reviews (including all pre-labelled positives: TPR = 1.0; FPR = 0.59) and 243 records out of 356 (68.3%) for GWAS (including 24 pre-labelled positives: TPR = 0.96; FPR = 0.66). Automated selection predicted 1,261 records out of 2,069 (60.9%) as important (including 50 pre-labelled positives: TPR = 0.98; FPR = 0.60). The workload for manual selection was ~60 hours. This means that for the applied model and these set(s), the reduction in workload is expected to be ~23.5 hours. By allowing for a higher FNR in model selection, the workload could be reduced even further, although at the expense of missing more true positives.

Our automated selection repeated the traditional manual search with inclusion rates (100% for reviews [58.8% false positives], 96.0% for GWASs [66.2% false positives], 98.4% for reviews and GWASs combined [60.0% false positives]), 0 cases were false negatives for reviews, 1 case for GWASs, and 1 case for reviews and GWASs combined.

A new search on "human aggression genes" was performed in the same databases without additional search terms and time limitation (14,400 records) to detect new contributions to the systematic review, resulting in 55.8% included records. Exclusion of duplicate records resulted in 6,469 records. From these, four reviews were added to the overview of reviews on aggression, and one SNP-heritability and two linkage studies were added to the GWASs review as additional information for the interpretation of GWAS findings. These seven studies were detected only by the ASR approach and did not appear in the traditional approach.

RESULTS

We included 18 reviews on the genetics of human aggression in our analyses, each covering different periods and including varying numbers of studies (see Table 2). The reviews cover more than 2,000 studies on aggression.

What is considered to be aggression?

Reviews indicate that the phenotypic definitions of aggression vary considerably, and heterogeneity of the phenotypic definition is mentioned as a major hurdle in aggression research by multiple papers. Definitions of aggression, as well as the focal points of reviews, range from broadly-defined externalizing and antisocial behaviours (see Supplement S1), which also include potentially non-aggressive behaviours like rule-breaking behaviour (Fernàndez-Castillo & Cormand, 2016), to a narrow focus on chronic physical aggression (Tremblay, Vitaro, & Côté, 2018). Other reviews and studies focus more explicitly on psychiatric classifications like Oppositional Defiant Disorder (ODD), Conduct Disorder (CD), and Antisocial Personality Disorder (ASPD), which encompass aggressive acts and are correlated to antisocial behaviour (ASB; Raine, 2019; Veroude, Zhang-James, et al., 2016)One review incorporated the analysis of genetics of aggression in suicidal behaviour (Baud, 2005). Classifications which are useful in clinical practice (e.g., American Psychiatric Association, 2013), tend to consist of constellations of heterogeneous antisocial behaviours (e.g., "often initiates physical fights" vs. "is often truant from school") and personality characteristics (e.g., "having difficulty sustaining long-term relationships" vs. "lacks concern, regret or remorse about other people's distress").

Table 2 Reviews on genetics of human aggression

Review	Type of studies included	N papers with trait-related studies	Taxonomy of aggressive behaviour (phenotype)	Samples
Baud (2005)	heritability studies, CGS	91	limited discussion of genetics studies of aggression, impulsivity and anger-related traits in suicidal behaviour	humans
Moffitt (2005)	Heritability studies (twins, adoption, family)	117	antisocial behaviour	children, adolescents, adults
Craig and Halton (2009)	heritability studies, CGS, GWAS	117	human aggressive behaviour; instrumental (proactive) and reactive	humans
Tuvblad and Baker (2011)	heritability studies (twin and adoption studies), CGS	138	human aggressive behaviour	children, adolescents, adults
Anholt and Mackay (2012)	CGS, GWAS	127	aggression as quantitative trait, pathological aggression (in substance abuse, psychiatric disorders, Alzheimer), externalizing behaviour	humans and animals
Vassos et al. (2014)	CGS	185	aggression and violence (categorical and continuous outcomes)	general population and specific subgroups

Discussed genes and polymorphisms in association with aggressive behaviour	Main conclusions
TPH, MAOA, COMT polymorphism	Aggression and unprovoked anger could be associated with intronic polymorphism in the *TPH* gene, VNTR regulatory polymorphism in the promoter region of the gene for *MAO-A*. The *COMT* genotype could differentially affect outwardly and inwardly directed aggressive behaviour.
MAOA, 5-HTTLPR polymorphisms	Environmental and genetic causes are equally important for antisocial outcomes. Heritability estimates form a curve with its peak at 50%, and small tails to the left (0% h^2) and right (80% h^2). Candidate genes should be chosen for GxE research based on a biologically plausible hypothesis that gene moderates responses to an environmental risk.
DGKA (DAGK1), GRIA3, CAG repeats, *MAOA, MAOB, SLC6A4, TPH1, TPH2, 5HT2A*, G861C, T102C, C-1021T polymorphisms, *COMT, ADRB1, NET1, SLC6A2, SLC2A1, NOS1, AVPR1A*	Genetic factors and common environment are equally important in childhood, heritability became more prominent in adulthood. Male heritability is slightly higher than that for females that implies specific genes on the X and/or Y chromosome. Genes do not operate independently, but function against a background in which other genetic and environmental factors are crucial.
MAOA, SLC6A4 (5HTTLPR), DRD2, DRD3, DRD4, DAT1, COMT, VNTR alleles of 5HTTLPR, SNPs of epinephrine and norepinephrine	About half (50%) of the variance in aggressive behaviour is explained by genetic influences in both males and females, 50% is explained by nonshared environmental factors. Form of aggression, method of assessment, and age of the subjects seem to be significant moderators. Study design and sex seem to be not significant moderators. Identification of genetic risks at the level of specific genes will reflect only an increased (probabilistic) risk and not a biological determinism.
apolipoprotein E e4 allele, tryptophan hydroxylase, serotonin 5HT-2A and 5HT-2C receptors and serotonin transporter, *COMT, MAOA, SLC6A4, DRD4, NOS-I, NOS-III*	Aggression is a quantitative trait, the manifestation of which is attributable to multiple segregating genes that are sensitive to the environment. Aggression is under stabilizing selection. It is difficult to discriminate correlations with disease status from causality in the aggressive phenotype. Polymorphisms in genes encoding the serotonin transporter and MAOA have been definitevly implicated in predisposition to aggression.
HTR1B, SLC6A4 (5HTTLPR), 5HTT-VNTR, *BDNF, COMT, SLC6A3, DRD4, MAOA-F, MAOA-M, TPH1, AR* (CAG), *DRD2*	No strong associations between selected polymorphisms and aggression outcomes are found. The candidate gene approach has not succeeded in identifying genes associated with aggression.

Table 2 Continued.

Review	Type of studies included	N papers with trait-related studies	Taxonomy of aggressive behaviour (phenotype)	Samples
Provencal et al. (2015)	heritability studies, CGS, GWAS, EWAS	176	chronic physical aggression	humans and animals
Zhang-James and Faraone (2016)	CGS	524 OMIM records	aggressive and antisocial behaviour, conduct disorder	humans
Fernandez-Castillo and Cormand (2016)	CGS, GWAS, pathways and functions	198	aggressive behaviours including aggression traits (aggressiveness, impulsive aggression, anger, externalizing behaviour, violence, delinquency or criminality) or diagnostic categories (OD, CD, ASPD, CU, and psychopathy)	humans
Veroude et al. (2016)	heritability studies, animal models, CGS, GWAS	378	RDoC nomenclature: frustrative non-reward, defensive and offensive (or proactive) aggression. ODD, CD, APD	humans (children, adolescents, adults) and animals

Discussed genes and polymorphisms in association with aggressive behaviour	Main conclusions
5-HT, MAOA, DRD2, SLC6A4, methylation patterns of *NR3C1*, *PCDH*, *SLC6A4*, *GR* and *CRH* genes, *AVPR1A*, *HTR1D*, HPA-regulating genes (*NR3C1*, *CRHBP*) and others	The response to early-life social adversity and aggression has an immune component. The immune system and the brain are interconnected through the hypothalamic-pituitary-adrenal (HPA) axis and the 5-HT system, and might play a role in the response to social adversity and in the development of chronic physical aggression through epigenetic mechanisms. T-cells could be useful to investigate.
Genes associated with aggressive behaviours in human (n=86)	A list of human disorder (n=95) have documented aggressive symptoms in at least one individual with a well-defined genetic variant; 86 causal genes were retreived.
Genes of dopamine and serotonin neurotransmission, hormone regulation and others in CGS. *BDNF, CAMK2A, DYRK1AFYN, ILVBL(FLJ39061), KIRREL3, LOC729257, LRRC7, MYRFL(c12orf28), NTRK2, PAWR, RBFOX1(A2BP1), RGL1, SHISA6* and others in GWASs.	Most CGS have identified associations with genes involved in dopaminergic and serotonergic neurotransmission and in hormone regulation. GWAS have not yet identified genome-wide significant associations, but top nominal findings are related to several signalling pathways, such as axon guidance or estrogen receptor signalling, and to neurodevelopmental processes and synaptic plasticity.
5HTT, 5HTTLPR, A2BP1, ABCG1, ADH1C, AKAP5, androgen receptor haplotype, *ANK3, AVP, AVPR1A, AVPR1B, BDNF, CAMK2A, COMT, DRD2, DRD4, DYRK1A, ESR1, FYN, HTR1B, ILVBL (FLJ39061), KIRREL3, MAOA, MFHAS1, MYRFL, NTRK2, OXTR, PAWR, PURG, RBFOX1, RIT1, ROBO2, SHISA6, SLC6A1* and others	Both CGS and GWAS approaches have identified potential susceptibility genes for aggressive behaviour. CGS have focused mainly in dopaminergic and serotonergic genes. GWAS, although not reaching genome-wide significance, have highlighted genes involved in neurodevelopmental processes and synaptic plasticity.

Table 2 Continued.

Review	Type of studies included	N papers with trait-related studies	Taxonomy of aggressive behaviour (phenotype)	Samples
Waltes et al. (2016)	heritability studies, animal models, CGS, GWAS, EWAS	248	human aggressive behaviour, reactive (impulsive) and proactive (pre-mediated) aggression	humans
Manchia and Fanos (2017)	CGS, GWAS, epigenetic, metabolomic, microbiomic association studies	87	aggression in mental illness	humans
Zhang-James et al. (2018)	GWAS	9	aggression	children, adults
Beaver et al. (2018)	heritability studies, CGS, GWAS	40	antisocial behaviour, aggression, violence	humans
Tremblay et al., (2018)	heritability studies (twin studies, adoption studies), CGS, epigenetic studies	123	physical aggression	human (children, adolescents) and animals

Discussed genes and polymorphisms in association with aggressive behaviour	Main conclusions
ABCG1, APOE, AR, AVPR1A, AVPR1B, BDNF, COMT, CRHR1, DRD1, DRD2, DRD3, DRD4, ESR1, HTR1A, HTR1B, HTR2A, MAOA, NOS1, NOS3, NR3C2, OXTR, SLC6A3, SLC6A4, TPH1, TPH2.	Heritability estimates from twin studies are highly variable. Several CTG are related to the monoaminergic neurotransmitter systems, genes regulating the HPA axis, and hormone pathways. Targeted analysis of genes known to be associated with aggressive behaviour suggests the epigenetic modulations.
ADNP2, BDNF, HTR2A, ITGB3, MTHFR, NRGN, PARD6G-AS1, TPH1, TRPS1	Specific genetic signatures of aggressive behaviour are present, which might result in substantial neurobiological alteration predisposing to behavioural dysregulation, particularly in individuals with severe mental illnesses. Environmental moderators act on the predisposing liability threshold set by genetic factors altering the expression of specific genes through, but not exclusively, changes in DNA methylation.
ACHE, ALDH5A1, ALK, AVPR1A, CACNB3, CADM1, CHMP2B, CRHR1, DNAJB5, EN2, ERBB4, FGF14, GRIA3, HDAC4, KCNJ18, LAMA2, LRRC7, MAOA, MECP2, NFKB1, OSMR, PRNP, RBFOX1, SERPINI1, WDR62	Among the top enriched pathways, several were previously well-known pathways for aggression (the dopamine, serotonin, glutamate, and GABA signalling pathways). The adult and child GWAS sets had six genes in common: *ALK, LAMA2, NFKB1, OSMR, RBFOX1*, and *WDR62*. Ranked gene list highlights 40 top genes, involved in neurotransmission, axon guidance, synaptic plasticity, learning and memory, neuronal development, or hormone signalling.
COMT, DAT1, DRD2/ ANKK1, DRD3, DRD4, DRD5, MAOA; 5HTTLPR, 5HTR2A, 5HTR1B, 5HTR2C polymorphisms, SNPs located in *C1QTNF7, DYRK1A, CDH13*	The heritability of antisocial behaviour is approximately 50%. Nonshared environmental influences account for the overwhelming majority of all environmental variance. Genetic polymorphisms involved in neurotransmission have most frequently been connected to antisocial phenotypes. Genetic and environmental influences frequently interact to predict variation in antisocial outcomes.
MAOA, DRD2, 5-HTT, SLC6A4. Methylation of glucocorticoid receptor gene and serotonergic system genes.	The development of chronic physical aggression is generally influenced by genetic and environmental factors through numerous interrelated bio-psycho-social channels from conception onwards. Involved genes vary with age and interact with the environment.

Table 2 Continued.

Review	Type of studies included	N papers with trait-related studies	Taxonomy of aggressive behaviour (phenotype)	Samples
Davydova et al. (2018)	Heritability studies, CGS, GWAS	78	aggressive behaviour	children, adults
Salvatore and Dick (2018)	Heritability studies, CGS, linkage, GWAS, GxE studies, rGE studies, epigenetics	96	conduct disorder	humans
Gard et al (2018)	heritability studies, CTG, GWAS (metaanlyses)	56	antisocial behaviour, including aggression, violence and rule-breaking	humans

CGS=candidate gene studies, GWAS = genome-wide association study, EWAS=epigenome-wide association study, GxE=genome-environment interaction, rGE=genome-environment correlation, VNTR=variable number tandem repeat, Chr=chromosome
Genes are sorted in alphabetic order. When gene name has a new name in HUGO, the old name used in the article is given in brackets.

Discussed genes and polymorphisms in association with aggressive behaviour	Main conclusions
AR, AVPR1A, AVPR1B, BAXBDNF, CASP3, COMT, DRD4, ESR1 (ER1), HTR2A, MAOA, OXT, OXTR, SLC6A, SLC6A4, TNR2A, TPH1, TPH2	Genes involved in cell adhesion, synaptic plasticity, and neurogenesis as key processes in development of aggressive phenotype may be considered as potential genetic markers for further research of aggressive behaviour
A2BP1, AVPR1A, ILVBL (FLJ39061), GABRA2, KIRREL3, LOC729257, LRRTM4/SNAR-H, MAOA, MYRFL (c12orf28), PAWR, PKD1L2, PKD1L3, RGL1, SLC6A4	Linkage studies identified regions of interest in different chromosomes, but few regions reach conventional thresholds. There is little consistency among regions identified across samples, with the exception of the region on chromosome 2. Suggestive evidence was found for SNP rs11126630 and between conduct disorder related phenotypes and *GABRA2, MAOA, SLC6A4*, and *AVPR1A* across independent samples.
Dopamine genes *DRD4, D4, DAT1, DRD2, DRD5, D5.* Serotonin genes 5-HTTLPR in *SLC6A4.* Catecholamine catabolism genes *MAOA, COMT.* Chr 1, 4, 7, 11, 13, and X. *ABCB1, C1QTNF7, LRRTM4/SNAR-H.*	The current body of work is limited by single candidate gene and GxE interaction studies that often utilize small sample sizes and imprecise measures of ASB. GWAS has not been able to identify any single gene(s) linked to ASB, emphasizing the need to look for biological substrates through which genes may indirectly impact ASB. Novel approaches, including neurogenetics and GxE studies, represent exciting potential avenues to better understanding the mechanisms of ASB.

Several reviews proposed a focus on more homogeneous or dimensional constructs of aggression (Fernàndez-Castillo & Cormand, 2016; Tremblay et al., 2018). A dimensional construct is in line with the conceptualization that pathological aggression is situated on the extreme ends of a normal distribution (Veroude, Zhang-James, et al., 2016). Some authors see a risk in the dimensional approach and note that findings might become predominantly driven by variations within normal, adaptive levels of aggression (Ferguson, 2010). However, if pathological levels of aggression are indeed the extreme end of a continuous phenotype, the same genetic and environmental factors should apply to both the normal range and extremes of the distribution.

In the end, concerns regarding heterogeneity and the impact of different phenotype definitions are empirical questions, which are currently also being asked in other GWASs of psychiatric disorders such as depression (Cai et al., 2018). Such questions can be resolved, once well-powered GWASs are available, by estimation of genetic correlations among different phenotype definitions of aggression and can also be addressed through genetic modelling of twin and family data. For example, Hendriks et al. (2020) analysed twin data collected by multiple instruments, commonly employed to measure aggression in children. While phenotypic correlations between different aggression scales could be low, a genetic multivariate analysis of these data showed high genetic correlations among different instruments. Such observations mean that different instrument tap into the same genetic liability and could be analysed simultaneously in GWAS.

Reviews that propose some sort of differentiation among aggressive behaviours, often return to a distinction between reactive and proactive aggression. Reactive aggression is commonly described as impulsive and defensive, while proactive aggression is considered predatory and premeditated. Both types of aggression may involve similar biological systems. The aminergic systems (e.g., serotonergic, dopaminergic) have been proposed as likely to regulate both forms of aggression (Waltes, Chiocchetti, & Freitag, 2016). Interestingly, Runions and colleagues (2019) argue that researchers studying reactive and proactive forms of aggression have conflated motivation (aversive vs. appetitive) and implementation (impulsive vs. premeditated) and propose that predatory aggression can also be impulsive in nature, defined as recreation instead of rage, while reactive aggression could also be delivered after a longer period of time, referring to reward instead of revenge.

The developmental aspect of aggression is a major theme in reviews (Davydova, Litvinov, Enikveeva, Malykh, & Khusnutdinova, 2018; Moffit, 2005; Provencal, Booij, & Tremblay, 2015; Tuvblad & Baker, 2011; Veroude, Zhang-James, et al., 2016; Waltes et al., 2016). Age of onset is often mentioned as an important differentiating factor for subtypes of antisocial behaviour, with aggression usually already present in early childhood, while rule-breaking behaviour and delinquency usually develop during adolescence. Tremblay (2010) proposes a developmental framework of aggression among a covert/overt axis and a second destructive/non-destructive axis as the most viable constructs to subtype disruptive behaviour (aggression, opposition-defiance, rule breaking, and stealing-vandalism). Children who display destructive and overt disruptive behaviours, especially those exhibiting chronic physical aggression, experience more risk factors early in life, engage in aggression from a young age, and have a more persistent developmental course of aggression and antisocial behaviour. A differentiation on age of onset is considered especially relevant in reviews which include epigenetics. Epigenetic changes may be triggered by early life adversity (Curry, 2019; Manchia & Fanos, 2017; Provencal et al., 2015; Tremblay et al., 2018), although variation in epigenetic marks can also reflect influences of DNA polymorphisms (van Dongen et al., 2016).

In research, aggressive behaviour often is measured by questionnaires, such as the Achenbach System of Empirically Based Assessment scales (ASEBA; Achenbach, Ivanova, & Rescorla, 2017), the Strengths and Difficulties Questionnaire (SDQ; Goodman, Lamping, & Ploubidis, 2010), or the Buss Durkee Hostility Inventory (BDHI; Buss & Durkee, 1957). Aggression scales in such instruments may include items which reflect behaviour that is related to aggression, but would not be considered aggression based on item content. For example, the ASEBA Aggressive Behaviour scale of for children contains items like "Argues a lot" or "Gets in many fights", but also "Unusually loud" or "Suspicious". Measures can also derive from observational studies, especially in younger children, and some experimental paradigms are available to measure aggression in across wider age ranges. Such experiments can, however, not cover the full spectrum of aggressive behaviour and, perhaps even more critically, cannot be applied in epidemiological samples.

There is a divergence between measurement of aggression in research projects compared to how (pathological) aggression is defined in clinical practice. Questionnaires are used as tools by clinicians, but the presence of these

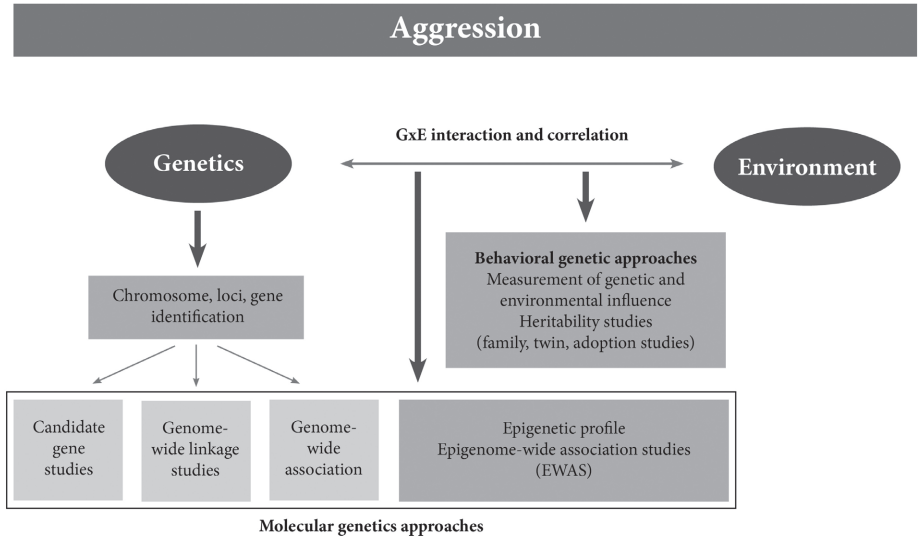
behaviours is mostly determined by interviews with the patient, and others who know the person (e.g., parents, teachers), by observation, and by the patient's (criminal) records. Psychiatric disorders that include aggressive behaviours or disorders which are correlated to aggressive and antisocial lifestyles, are dependent on classification systems like the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Diseases (ICD). In these classifications a dichotomy is applied in which a disorder is either present or absent, largely ignoring the dimensional nature of human behaviour. In genetic studies, a focus on the dichotomy rather than on continuous variation, may lead to a loss of statistical power (Van der Sluis, Posthuma, Nivard, Verhage, & Dolan, 2013).

Another important question, especially in clinical settings, is when aggression becomes pathological. Some aggressive behaviours are clearly defined as pathological, like aggressive behaviours that define Conduct Disorder (e.g., "Has used a weapon that can cause serious physical harm to others), or Antisocial Personality Disorder (e.g., "Irritability and aggressiveness, as indicated by repeated physical fights or assaults"). In contrast, other aggressive behaviours are less clearly considered pathological, because they occur to some extent in all individuals, like anger or hostility. This even is the case for some aggressive behaviours which are part of disruptive behaviour disorders (e.g., ODD: often argues with authority figures). For aggression to be pathological, it is essential that aggressive behaviours cause clinically significant impairment in social, academic, or occupational functioning.

Approaches in genetics of aggression studies and the current status quo

There are several designs to study the genetic aetiology of aggression, with the two major ones being genetic epidemiological / behavioural genetic approaches on the one hand and molecular genetic approaches on the other (see Figure 2). Behavioural genetic studies have a long and successful history (Loehlin, 2009). More recently, molecular genetic studies have seen enormous breakthroughs with the development of techniques like GWASs (Visscher et al., 2017).

Figure 2 Interplay of genetic, epigenetic, and environmental factors in behaviour and genetic studies of aggression



Behavioural genetic approaches

Numerous studies focused on explaining the aetiology of aggression and antisocial behaviour through family, twin, and adoption studies, which can disentangle genetic and environmental influences. Twin models enable researchers to divide the variance for a trait, or the liability to a disorder, into genetic and non-genetic components. The genetic variance component often is defined as the additive (A) effects of many genes. Environmental variance components consist of environmental influences common to siblings from the same family (C), creating resemblance of family members through environment rather than through genetics, and a unique or non-shared environmental component (E). Unique environmental influences affect family members in different ways (Boomsma, Busjahn and Peltonen, 2002). Unsystematic influences such as measurement error also are included in the E component, unless explicitly modelled. In general, reviews indicate that additive genetic factors explain around 50% of the variability of aggressive behaviour (Craig & Halton, 2009; Fernàndez-Castillo & Cormand, 2016; Rhee & Waldman, 2002; Tuvblad & Baker, 2011). The estimate varies around 50% across studies, with some reviews reporting somewhat higher heritability estimates (65%) and others giving estimates for aggression and

antisocial behaviour that vary more (e.g., 38%-88%; Veroude, 2016; 28%-78%; Tuvblad & Baker, 2011). Physical aggression seems to show larger heritability estimates (65%) than reactive (20-43%) and proactive aggression (32-48%), while rule-breaking behaviour, which is often aggregated with aggression indices, also shows a heritability around 50% (Gard, Dotterer, & Hyde, 2019; Waltes, Chiocchetti, & Freitag, 2015). Heritability estimates of aggressive behaviour were higher in children with stable callous unemotional traits (81%) compared to children low in callous unemotional traits (30%; Gard et al., 2019). This suggests a larger influence of genes on children with more severe aggressive tendencies (Gard et al., 2019). Contributions of shared environment are relatively small and decrease with age, with the vast majority of adult studies not reporting any shared environmental influences (Tuvblad & Baker, 2011; Veroude, Zhang-James, et al., 2016; Waltes et al., 2015). Thus, research in behaviour genetics clearly indicates that there is a substantial genetic component to aggressive behaviour in humans. In longitudinal studies, heritability estimates of aggression and antisocial behaviour increase somewhat from childhood through adulthood (Tuvblad & Baker, 2011; Veroude, Zhang-James, et al., 2016; Waltes et al., 2015). Genetic factors also contribute to the stability of aggressive behaviour during preschool and school age, and puberty (Porsch et al., 2016; Waltes et al., 2015). Measurement instrument, and also rater seem to influence heritability estimates, with heritability based on parent-report and teacher-report estimated as higher than those based on self-report and observational studies. Studies based on self-report tend not to find any shared environmental influences (Tuvblad & Baker, 2011), but such studies are not available for younger children. Unlike parent or teacher reports, observational studies more often give an assessment of aggression at one particular moment in time only. Parent- and teacher-reports tend to provide phenotype information that is more averaged over longer periods of time and are similar in terms of heritability estimates. Parent-report leads to higher estimates of shared environmental influences than teacher-report, when parental characteristics that influence ratings of multiple children (e.g., twins or siblings) are not taken into account. When twins have different teachers, similarities between them tend to decrease. This may reflect actual differences in aggressive behaviour with different teachers and/or different settings, but may also reflect teacher characteristics that influence assessments of multiple children.

In summary, heritability is estimated consistently around 50%, with some variation that may be due to different conceptualization of aggressive and

antisocial behaviours, with more severe types of aggression showing higher heritability.

Heritability estimates of aggression and antisocial behaviour may differ between environments suggesting an interaction between genes and environment (GxE). Proposed putative environmental moderators are familial adversity (e.g., maltreatment, parental delinquency), social disadvantage (e.g., poverty, bad neighbourhoods), violent media exposure, and alcohol use. Tuvblad and Baker (2011) argue that, compared to genetic factors, environmental influences are relatively more pronounced for antisocial behaviours in the presence of high environmental risk and disadvantaged environments. Conversely, genetic influences will be more pronounced when environmental risk factors are absent or less prominent. In one study, the moderating effects of neighbourhood seemed to be specific to the heritability of nonaggressive antisocial behaviour, while heritability estimates of aggressive antisocial behaviour were not influenced by neighbourhood disadvantage (Burt et al., 2016). Such findings underscore the differential influence of environmental adversity on certain types of antisocial behaviour, with aggressive behaviour showing less sensitivity to environmental influences than other types of antisocial behaviour. Later reviews, however, indicate mixed findings. Some reported an increase in genetic variance in the presence of environmental risk. To illustrate, when young children were subjected to high levels of maternal disengagement, genetic factors explained more variance in later conduct problems (Boutwell, Beaver, Barnes, & Vaske, 2012; Waltes et al., 2015). An increase in heritability of externalizing disorders was also found when young adults were exposed to a combination of risk factors (e.g., antisocial or lack of prosocial peers, relationship problems with parents (e.g., antisocial or lack of prosocial peers, relationship problems with parents; Hicks, South, DiRago, Iacono, & McGue, 2009; Veroude, Zhang-James, et al., 2016).

Depending on the type of aggression, mean levels of aggression often are higher in males than in females. Differences in heritability estimates, however, between males and females are modest or absent. According to Tuvblad (2011) heritability did not differ significantly between genders across different twin studies, either quantitatively or qualitatively (see also: Vink et al., 2012). These studies mainly included mother-reports of childhood aggression and heritability estimates were higher in males than in females when self-report data were analysed (Waltes et al., 2015). It has been suggested that gender differences in heritability become more pronounced from adolescence, which could be

indicative of the "Young Male Syndrome", in which the onset of puberty and increasing levels of testosterone are related to increases in aggression in 12- to 25-year-old males (Craig & Halton, 2009). This would also suggest a possible role of genes related to androgen synthesis and function in the development of aggression from puberty onwards.

In summary, twin studies highlight the importance of genetic influences, with estimates of the heritability of aggression and antisocial behaviour often reported to be around 50% (Moffitt, 2005), without much evidence for sex differences in heritability estimates. Such significant heritability is a first requirement for initiating studies that aim to find molecular signatures in the DNA sequence that are associated or causally related to the phenotype.

Integrating data on genetics of aggression from molecular genetic studies

Genetic linkage and candidate gene studies

Molecular genetic studies include genetic linkage and association studies, either genome-wide or with a focus on a limited number of candidate genes or candidate regions. In *linkage studies*, DNA markers are assessed in related individuals to investigate the inheritance of markers with known chromosomal locations together with aggression in pedigrees. Sometimes candidate regions to be investigated are suggested from studies in other species. With the arrival of large scale association studies, linkage studies, which require family-based designs, have become less common, but early studies have suggested regions on three chromosomes that could be associated with aggression. Dick et al. (2004) analysed retrospectively reported childhood conduct disorder in an adult sample from COGA (Collaborative Study on the Genetics of Alcoholism). Regions on chromosomes 19 and 2 may contain genes associated with risk of CD. The same region on chromosome 2 has been linked do alcohol dependence in this sample. Criado et al. (2012) in a linkage study of cortical even-related oscillations associated with ASPD and CD suggested that chromosome 1 may contain a genetic locus for ASPD/CD.

Genetic association studies initially were *candidate gene studies*. These require *a priori* knowledge of or hypotheses about which genes are implicated in the aetiology of the trait of interest.

For aggression, associations were considered for genes from the serotoninergic [*5-HTTLPR* (5-hydroxytryptamine (serotonin) receptors), *SLC6A4* (solute carrier family 6 member 4)], dopaminergic [dopamine receptors genes *DRD4, DRD2, DRD5*, and *SLC6A3* (solute carrier family 6 member 3)] and GABAergic systems [e.g., genes that code GABA (gamma-aminobutyric acid) receptors, like *GABRA2* (gamma-aminobutyric acid type A receptor alpha2 subunit)], as well as genes related to catecholamine catabolism [*MAOA* (monoamine oxidase A), *COMT* (catechol-O-methyltransferase)] (Davydova et al., 2018; Fernàndez-Castillo & Cormand, 2016; Gard et al., 2019; Provencal et al., 2015; Veroude, Zhang-James, et al., 2016). Other studies focused on associations with the genes involved in stress response pathways (Craig & Halton, 2009; Waltes et al., 2015); hormone regulation (e.g., *AVPR1A* (argenine vasopressin receptor 1A)) (Fernàndez-Castillo & Cormand, 2016; Salvatore & Dick, 2018; Veroude, Zhang-James, et al., 2016; Waltes et al., 2015); hypoglycaemia and insulin secretion (Craig & Halton, 2009); and neuronal transcripts and brain expression patterns (Anholt & Mackay, 2012; Craig & Halton, 2009; Gard et al., 2019; Waltes et al., 2015). Candidate gene studies have been criticised (e.g., Duncan and Keller, 2011), since it became clear that findings for candidate genes are often not replicated in well-powered genome-wide association studies (e.g., Bosker et al., 2011; Luo et al., 2016). It is likely that this also extends to studies of aggression, but the status of the candidate genes for aggression must await well-powered GWASs.

Many reviews agree that aggression is a polygenic trait influenced by many genes, that each explains a small proportion of the phenotypic differences. There may however be an overlap between genes of large effect underlying monogenic disorders and those affecting continuous variability of related quantitative traits. Extending the idea of a shared genetic basis between Mendelian disorders and polygenic traits, one alternative approach based on the search for genes for aggression in studies of rare, functional genetic variants associated with aggression phenotypes catalogued in Online Mendelian Inheritance in Man (OMIM; Zhang-James & Faraone, 2016). Most of these genes had not been implicated in human aggression before, but the most significantly enriched pathways (e.g., serotonin and dopamine signalling) had been previously implicated in aggression. Among these genes, only two were previously related to aggression (*MAOA*, *GRIA3* (glutamate ionotropic receptor AMPS type subunit 3). New associations were found with genes [e.g., *CAMTA1* (calmodulin binding transcription activator 1), *APBB2* (amyloid beta precursor protein binding family

B member 2), *DISC1* (DISC1 scaffold protein) and others], which are implicated in cell-to-cell signalling and interaction, nervous system development and function, and behaviour. The novel genes and pathways identified in this study suggested additional mechanisms underlying aggression.

Genome-wide association studies

Genome-wide association studies (GWASs*)* investigate millions of Single Nucleotide Polymorphisms (SNPs), under a continuous or dichotomous, case/control, model. The result is a list that, for every variant, indicates the expected increase in a trait (continuous) or genetic liability (dichotomous) for every copy of an effect allele. Due to the large number of tests, the genome-wide significance level is set at $p = 5.0E^{-08}$ (Sham & Purcell, 2014), to properly control for the type I error rate. This adjusted threshold already considers the fact that neighbouring SNPs are not inherited independently from one another. However, the non-independent inheritance of SNPs indicates that association tests between non-causal SNPs and the trait of interest contain a part of the polygenic signal (Bulik-Sullivan et al., 2015). As such – even when only a limited number of SNPs reach this stringent significance level, there is signal in the other association tests. The weighted effects of all the genetic variants involved in aggression could produce a polygenic risk score with a certain predictive value (Beaver, Connolly, Nedelec, & Schwartz, 2018).

Many reviews discussed a whole genome approach to understanding aggression, but only three have done so in a systematic manner (Fernàndez-Castillo & Cormand, 2016; Veroude, Zhang-James, et al., 2016; Waltes et al., 2015). We will summarize findings for genes harbouring, or in close proximity to, variants that reached genome-wide ($P \leq 5.0E^{-08}$) or nominal ($P \leq 1.0E^{-05}$) significance levels in all GWAS of aggression phenotypes to date. These include aggression-related phenotypes, i.e., anger, hostility dimensions, aggressive behaviour, physical aggression, ASB, violent offending, CD, ODD, and ASPD.

To provide a complete picture of the GWAS literature available, we chose to include phenotypes which clearly include aggression, but are sometimes conflated with other antisocial behaviours (e.g., rule breaking) or personality characteristics (e.g., being suspiciousness, being loud). These phenotypes can be found in Supplement S4. Most GWASs on aggression were performed in child and adolescent samples that were assessed using rating scales and were done in samples of European ancestry (see Table 3).

GWAS studies have mainly resulted in nominal associations between genetic variants and aggression-related traits and disorders. Collectively these studies reported 10 genome-wide significant findings (Dick et al., 2011; Montalvo-Ortiz et al., 2018; Rautiainen et al., 2016; Tielbeek et al., 2017). Five of these variants are located inside or close to four genes: *LINC00951* (long intergenic non-protein coding RNA 951) (Rautiainen et al., 2016), *C1QTNF7* (C1q tumor necrosis factor-related protein 7) (Dick et al., 2011), *PSMD1* (proteasome 26S subunit, non-ATPase 1) and *HTR2B* (5-hydroxytryptamine receptor 2B) (Montalvo-Ortiz et al., 2018). Lastly, the five remaining significant SNPs are located on chromosomes 11 (Dick et al., 2011; Tielbeek et al., 2017), 13 (Dick et al., 2011), 1 and X (Tielbeek et al., 2017).

In a mixed sample of subjects from European and African-American ancestry, three SNPs inside *C1QTNF7* were significantly associated with Conduct Disorder (CD) symptoms in adults with substance dependence (Dick et al., 2011). When the sample was split on the basis of ancestry, no SNPs reached suggestive levels in the European-American sample. In the African-American sample one out of the three SNPs reached suggestive levels (minimum $p = 4.35E^{-06}$), along with two additional suggestive findings (minimum $p = 2.67E^{-07}$). *C1QTNF7* is less expressed in the brain, compared to such tissues as endometrium, gall bladder, lungs, ovaries and 18 other tissues, and has a potential role in maintaining energy balance (Kaye et al., 2017).

In a study focusing on antisocial personality disorder in Finnish criminal offenders, Rautiainen and colleagues (2016) found one hit (rs4714329, $p = 1.6E^{-09}$) in the cross-sex meta-analysis. This variant is in close proximity to *LINC00951* (long intergenic non-protein coding RNA 951). The same SNPs returned suggestive associations in the male-specific GWAMA of ASPD ($p = 1.38E^{-07}$). The signal from these variants was specific for ASPD, and did not cover a broader range of criminal behaviour. Montalvo-Ortiz and colleagues (2018) found that SNPs located in the *HTR2B* ($p = 2.16E^{-08}$) and *PSMD1* ($p = 1.79E^{-08}$) genes were significantly associated with cannabis-related physical aggression in African-Americans, but these SNPs did not reach even suggestive significance in European-Americans. Cannabis use has been associated with greater impulsive decision-making and increased aggressive behaviour. Notably this is the only GWAS study which focused purely on physical aggression.

Table 3 Overview of genome-wide suggestive and significant associations with aggression-related traits at P≤1E-05 per GWASs

Study	Sample	Phenotype	$N_{variants}$	N_{genes}
Sonuga-Barke et al. (2008)	N=909 probands in trios (~87% males)z ~99% had ADHD diagnosis Age range: 5-17 years European Caucasian ancestry	CD using PACS	18	7
Anney et al. (2008)	N=938 probands in trios (~87% males) ~99% had ADHD diagnosis Age range: 5-17 years European Caucasians ancestry	CD using DSM-IV criteria for CD, PACS and CPRS-R:L, gathered the symptom on a less severe behavioural characteristic of an oppositional defiant individual.	54	41
Viding et al. (2010)	N=600 (69% males) from twin cohort (high- and low-scoring of AB) Replication N=586 (71%males) Age=7 years Caucasian ancestry	ASB/CU: Teacher-rated conduct problems and CU traits using SDQ; 3-point scale	0	0

Genes	Summary of main findings
GxE interaction with "mother's criticism" *PPM1K, ZBTB16* GxE interaction with "mother's warmth" *RBFOX1(A2BP1), ADH1C (proximal), MFHAS1, SLC6A1, RIT1 (proximal)*	Suggestive GxE interactions were reported for 18 SNPs, of which 3 SNPs also showed a suggestive main effect. For both the main and interaction effects, no SNP reached genome-wide significance.
LIG4 (proximal), ABHD13 (proximal); AMOLT1 (proximal), CWD15 (proximal), KDM4D (JMJD2D) (proximal); FLJ16077; RXFP1 (proximal); PAWR; LOC729257; SPATA8 (proximal); YWHAZ (proximal); FLJ31818, GPR85 (proximal); KIRREL3; PRPRD (proximal); ATP8B1 (proximal); MYRFL (c12orf28); LIG4 (proximal), ABHD13 (proximal); PKD1L2; c16orf46 (proximal); PKD1L3; KIAA0174 (proximal),; DHODH (proximal); c5orf16 (proximal); c5orf15 (proximal); FLJ39064; FZD10 (proximal); FLJ39063; FZD9 (proximal); FLJ39062; FZD8 (proximal); ILVBL(FLJ39061); FZD7 (proximal); ETV3L (proximal), ETV3 (proximal); FLJ17340; GSX1 (proximal), PDX1 (proximal); PITRM1 (proximal); RBFOX1(A2BP1); GLT25D2 (proximal); RGL1	Suggestive associations were reported for 54 SNPs. These SNPs were located in 11 genes and/or were within a 200kb window of 23 additional genes. The top five association signals were observed on Chr 13, 21, 11, 4, and 12.
Suggestive in replication (p = 4,77E^{-05}) *KCNMA1*	In both the discovery and replication study, no SNP reached genome-wide significance. Several top SNPs were located near neurodevelopmental genes such as ROBO2 (p = 4.61E^{-03})

Table 3 Continued.

Study	Sample	Phenotype	$N_{variants}$	N_{genes}
Dick et al. (2011)	N=3,963 (N_{cases}=872, $N_{controls}$=3091) Age range: 18-77 years European Americans, African Americans	CD: retrospective report of DSM-IV CD symptoms, natural log as primary CD measure.	29	10
Merjonen et al. (2011)	N=2443 (46% males) Age range: 15-30 years Followed up for 15 years European Caucasians ancestry (Finnish population)	Anger in hostility dimensions measured by the Irritability Scale of the Buss-Durkee Hostility Inventory in four time points over a 15-year interval	20	2
Mick et al. (2011)	N = 341 (64% males) ADHD offspring from 339 ADHD affected trio families Age range: 6-17 years Ancestry: NA	CBCL dysregulation subscale (anxiety/depression, aggression, attention problems subscale)	9	5
Tielbeek et al. (2012)	Combined sample N=4816 (41% males) 298 cases, 4518 controls Age range cases: 18-74 years Age range controls: 18-77 years Australians	ASB according to DSM-IV for CD Cohort 1: non-diagnostic measure covering seven items related to antisocial behaviour, case status was 3 symptoms or more Cohort 2: Diagnostic measure of ASPD, cases had a diagnoses of ASPD except for criterion D (the occurrence of antisocial behaviour is not exclusively during the course of schizophrenia or a manic episode)	22	12

Genes	Summary of main findings
In a sample with mixed ancestry *C1QTNF7*; PDE10A; SELPLG; TOX2; LOC343052; ERCC4*	European sample: were only reported for the top 20 SNPs that came out as suggestive/significant for the mixed analysis. None of the SNPs were suggestively associated with either phenotype within the European sample. Mixed sample with European and African ancestry: 4 SNPs reached genome-wide significance level for CD_{symp} – but not for CD_{cc} – two of which were located inside *C1QTNF7*. The other two significant SNPs were not located near any gene.
SHISA6; PURG	One SNP reached significance $p < 9E^{-8}$: Chr 17: rs11656526, closest gene *SHISA6*. Many associations with anger approached significance, among them SNPs located close to genes *PURG*.
FERMT3; LRRC7; STIP1; TRPT1; SEMA3A	Only results for top 50 SNPs were reported. No SNP reached genome-wide significance, but 9 were suggestively associated with DP. Out of these 9, 7 were located within 4 genes. Suggestive evidence for developmentally expressed genes operant in hippocampal dependent memory and learning associated with CBCL-DP is found.
DYRK1A; AL590874.1; CIB1; SEMA4B; TTC7B; IMMT; CSMD1; REEP1; RP11; BAZ2B; STK32A; VRK1	Sample was pooled together from two studies. Suggestive levels of significance were reached by 22 SNPs, located inside 12 genes. The gene with the strongest association was *DYRK1A*, previously related to abnormal brain development and mental retardation.

Table 3 Continued.

Study	Sample	Phenotype	$N_{variants}$	N_{genes}
McGue et al. (2013)	N=7,188 (46% males) Age: adults Caucasian ancestry	Behavioral Disinhibition; composite score consisting of five symptom counts for CD, ASB, Dissocial behavior, Delinquent Behavior Inventory, Aggressive underscore	4	1
Tiihonen et al. (2015)	Violent offending N_{cases}= 360 (94% males) Extreme violent offending N_{cases}= 56 (97% males) $N_{controls}$=5983 (57% males) Age (mean±s.d.) = 29.4±8.2 Finnish population	Violent offending; at least one sentence for violent offence. Extreme violent offending; 10 or more sever violent crimes	14	9
Mick et al. (2014)	N = 8,747 (47% males) From Atherosclerosis Risk in Communities Study. Age range: 45-64 years European ancestry	Angry temperament and angry reaction measured by SSTAS.	8	5
Salvatore et al. (2015)	Discovery N=1,379 (54% males) with alcohol dependency Age range: 18-79 years Replication N=1796 (46% males) Age range: 18-88 years European ancestry	ASB. Symptoms of DSM-IV ASPD. SSTAS	75	NA

Genes	Summary of main findings
GLIS1	Genome-wide suggestive levels were reached by 4 SNPs, tagging 1 gene.
Violent behaviour *SPIN1; NTM; ATP10B (proximal); PRMD2 (proximal); PLCB1; NXPH1 (proximal)* Extremely violent behaviour *CDH13; PRUNE2; LOC101928923*	Genome-wide suggestive levels for violent behavior were reached by 10 SNPs, mapping to 6 genes. Additionally, 4 suggestive SNPs (3 genes) were reported for extreme violent behavior.
Angry temperament *FYN (proximal), IYD (proximal), ZNFX1 (proximal), STAU1 (proximal), DDX27 (proximal)* Angry reaction ($p < 6E^{-03}$) *PHEX (proximal), SLC39A8 (proximal), MBOAT1(proximal), PLEK (proximal)*	p-values results from phenotypes adjusted for principal components representing genetic structure were used. Four SNPs reached suggestive levels of significance for angry temperament. Five SNPs reached suggestive levels for angry reaction $p < 6E^{-03}$, tagging four genes. Both scales were also dichotomized and treated as case-control phenotype, for which no SNP returned suggestive results. Results were only reported for SNPs with $P \leq 5E^{-06}$. 75 SNPs reached genome-wide suggestive levels. The top suggestive SNP on Chr 7, rs4728702, was in the *ABCB1* gene, which encodes a transporter protein. This suggestive association did not replicate in the replication sample. Found enrichment of several immune-related canonical pathways and gene ontologies, suggesting that immune and inflammatory pathways are associated with externalizing spectrum behaviours.

Table 3 Continued.

Study	Sample	Phenotype	$N_{variants}$	N_{genes}
Pappa et al. (2016)	N=18,988 9 cohorts Age range: 3-15 years North European ancestry	Predominantly parent-reported child aggressive behaviour. SDQ, CBCL, and other (parent rated questionnaires) in different cohort	76	16
Rautiainen et al. (2016)	Discovery N=6,220 (59% males) 370 ASPD, 5850 controls Age (mean±s.d.)$_{ASPD}$ = 34.5±8.0 Age $_{Controls}$ = 55.0±13.2 Replication N = 3939 (43% males) 173 ASPD, 3766 controls Age (mean±s.d.)$_{ASPD}$ = 34.2±9.2 Age $_{controls}$ = 55.0±17.0 Finnish population	ASPD (violent criminals, substance abuse, maltreatment). ASPD diagnoses, SCID-II items for DSM-IV	6	1
Aebi et al. (2016)	N=750 (87.8% males) with available ODD Age range: 5-18 years European Caucasian ancestry	ODD. CPRS-R: L. Continuous: defiant/ vindicative; irritable Case-control: low/moderate OPP vs irritable /severe OPP	53	14
Brevik et al. (2016)	N adults=1060 patients with ADHD N children= 750 with ADHD European Caucasian ancestry	Childhood aggressiveness in adult ADHD Adult sample: retroactive measure of childhood symptoms of ADHD. Child sample: CPRS-R:L, subdivided in defiant/ vindictive and irritable dimension	65	20

Genes	Summary of main findings
Overall *LRRTM4 (proximal)*;PDSS2; TRIM27 (proximal); MRC1; MECOM; CASC17 (proximal)* Early childhood *COL13A1; SDK1 (proximal); LOC101928923; TSG1 (proximal); LOC727982 (proximal)* Middle childhood/early adolescence *LRRTM4 (proximal); LOC101927797 (proximal); OPCML; COL13A1; GRIA1; ASBA; CNTN4*	Meta-analysis of nine cohorts reported one genome-wide significant hit. N35 SNPs reached suggestive levels for the overall GWAMA. These SNPs are located inside three genes and near three others.10 and 31 SNPs reached suggestive levels for GWAMA on early and middle childhood/ early adolescence AGG, respectively. Some of these SNPs overlap with the top hits reported in the overall GWAMA. In total suggestive associations were reported for 76 SNPs (66 unique) located in or around 16 genes.
Cross-sex *LINC00951 (proximal)** Males only *LINC00951 (proximal)*	Results based on meta-analysis across discovery and replication reported that for the cross-sex GWAMA, 1 SNP reached genome-wide significance while another SNP ~10Kbp away reached suggestive levels. The closest gene to these SNPs is *LINC00951*. In the male-specific GWAMA, four SNPs reached suggestive levels, two of which are the same ones as the SNPs reported in the overall GWAMA. The other two SNPs are within ~50Kbp
ADAM12; MYLK2 (proximal); OR2AG1 (proximal), OR2AG2 (proximal); BCL2L1; TPX2; DDX24 (proximal), ASB2 (proximal); RARB; RUNX1T1; FOXS1 (proximal); TTLL9 (proximal); COX4I2; SOX5; MYLK2	Results based on multivariate GWAS only reported that 53 SNPs reached genome-wide suggestive levels, which are located inside and/or near 14 unique genes.
NTM; CSMD1; KRT18P42 (proximal); TEPP; CPNE4; MICAL2 (proximal); LOC101929236; LOC101927464; NR_110053.1; H3F3A; LOC105370057; ACBD3 (proximal); LOC101929156; LOC105376469 (proximal); LOC105373223 (proximal); SPINK2; PHLPP1; UFM1	Results based on meta-analysis across adult and children samples reported that 65 SNPs – located in or near 20 genes – reached suggestive levels of associations. The strongest signal was observed at rs10826548 on Chr 10 located within the transcript of a long noncoding RNA (p = 1.07E^{-06}), closely followed by rs35974940 in *NTM* (p = 1.26E^{-06}).

Table 3 Continued.

Study	Sample	Phenotype	$N_{variants}$	N_{genes}
Tielbeek et al. (2017)	N=16,400 (47% males) Replication N=9,381 Mean age range across cohorts=6.7-56.1 years Ancestry: Mixed	Broad-spectrum ASB. Development and well-being assessment, conduct disorder scale, count of the number of APD criteria, rule-breaking behaviour, Teacher report Form, Antisocial Process Screening Device, Retrospective CD, SCID-II for DSM-IV disorders, CBCL: conduct problems (reported by mother), DSM-IV CD criteria	80	NA
Montalvo-Ortiz et al. (2018)	N=2,185 African Americans (~61% males) N=1,362 European Americans (~64% males) Replication N=89 African Americans (49% males) Exposed to cannabis use Age mean ~ 37-45 (in different cohorts European Americans, African Americans	Cannabis related physical aggression assessed with the question, "Did you ever get into physical fights while using marijuana?"	280**	43

From left to right, columns indicate (1) study, (2) sample description, (3) phenotype description, (4) number of (unique) associated SNPs/variants, (5) number of (unique) genes, (6) gene names, and (7) summary of main findings

Selection of associated with aggressive behaviour genes presented in the table is done on the base of associated SNP at $p < 1E^{-05}$ (nominally significant). Genes are sorted by ascending p in SNPs (the lowest level if gene is associated with several SNPs). When gene name has a new name in HUGO, the old name used in the study is given in brackets. The nearby location of nominally significant SNP is given in brackets (proximal), in other cases the location is intragenic.

Genes for SNPs with genome-wide significance ($p < 5.0E^{-08}$) are indicated with *

ASB=antisocial behaviour, CD=conduct disorder, CU=callous-unemotional, ASPD=antisocial personality disorder, ODD = oppositional defiant disorder, DP=dysregulation profile

Genes	Summary of main findings
	GWAMA across five cohorts. Only independent signals are reported. The cross-sex GWAMA reports 20 suggestive associations, of which 2 are InDels. Two significant associations were found for the female-specific GWAMA. These two SNPs are located on Chr 1 and 11, respectively. The male-specific GWAMA returned one significant association on the X-chromosome. The female- and male-specific GWAMAs returned 37 and 20 suggestive associations, respectively. In total 80 unique variants (64 SNPs) were associated with ASB. ASB has potential heterogeneous genetic effects across sex.
European ancestry *LPPR1; ARHGEF3; RARB; TMEM92; ERBB4; CCDC171; ATP10A; UST; GPRC5B; CDH13; GRIN2B* African ancestry *PSMD1*; HTR2B*; CCDC157; TBC1D10A; GSG1L; THSD7B; BRINP1; CNTN3; NSG2; SF3A1; SOD3; ADGRV1 (GPR98); KLHL3; SEC31A; ABR; TSPEAR; TMEM53; CCDC141; STAB2; RTN1; CDYL; UBE2H; LRMDA (C10orf11); ANO4; STRC; TASOR2 (FAM208B); SERTAD1; ARMH1 (C1orf228); CEP126 (KIAA1377); ABCA13; SLC17A6; LRRC4C*	European-American sample: suggestive associations were found for 76 variants, of which 7 were structural variants. The 76 variants implicate 11 genes African-American sample: the top SNPs included rs35750632 in *PSMD1* and rs17440378 in *HTR2B*. Based both on its demonstrated contribution to aggressive behaviour and functional annotation analysis, *HTR2B* is suggested to be the relevant gene.

Chr=chromosome, GWS = genome-wide significant, NA= not available, GWAMA=genome-wide association meta-analysis
PACS=Parental Account of Childhood Symptoms; CPRS-R: L=long version of the Conners Parent Rating Scale; CBCL=Child Behavioural Checklist; SCID-II=Structured Clinical Interview Axis II; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders; SSTAS=Spielberger State-Trait Anger Scale; BDHI= Buss-Durkee Hostility Inventory; SDQ=Strengths and Difficulties Questionnaire; SCID-II=Structured Clinical Interview Axis II
**for Montalvo-Ortiz et al. (2018) SNPs, variants and genes are included at $p < 1E^{-06}$

Anney and colleagues (2008) listed 54 SNPs nominally associated with conduct problems. These SNPs tagged 41 genes 3 of which are with known functions and are involved in the regulation of dopamine receptor D2 signalling (*PAWR* (pro-apoptotic WT1 regulator)), synaptic plasticity (*KIRREL3* (kirre like nephrin family adhesion molecule 3)) and neuronal development (*RBFOX1* (ral guanine nucleotide dissociation stimulator like 1)). Sonuga-Barke and colleagues (2008), analysed interactions between CD symptoms and maternal warmth. Nominal effects were found for SNPs located in genes involved in brain maturation, neurotransmission, neuronal development and regeneration. Viding and colleagues (Viding et al., 2010) examined teacher-reported conduct problems in children and found no suggestive SNPs (minimum $p = 4.6E^{-05}$).

For adult ASB (Tielbeek et al., 2012) the strongest signal was for a SNP (rs346425; $p = 2.51E^{-07}$) located on chromosome 5. Salvatore and colleagues (2015) in an adult ASB sample observed the strongest association for rs4728702 ($p = 5.77E^{-07}$), located in *ABCB1* (ATP binding cassette subfamily B member 1) on chromosome 7 that may confer general risk across a wide range of externalizing behaviours. Enrichment analyses further indicated involvement of immune-related pathways. Two genome-wide association studies compared cohorts of Finnish violent offenders to the general population (Rautiainen et al., 2016; Tiihonen et al., 2015), and obtained association signals at genes involved in neuronal development (Tiihonen et al., 2015) and adaptive immunity (Rautiainen et al., 2016).

Aebi and colleagues (2016) hypothesized that *BCL2L1* (BCL2 like 1) is likely associated with oppositional behaviour, because of its influence on presynaptic plasticity through regulation of neurotransmitter release and retrieval of vesicles in neurons. Brevik and colleagues (2016) applying gene-based tests observed *NTM* (neurotrimin) as the top gene, that is differentially expressed in aggression-related structures of the amygdala and the prefrontal cortex in early stages of brain development.

Merjonen and colleagues (2011) saw suggestive associations for SNPs that lie inside genes involved in the maintenance of high frequency synaptic transmission at hippocampal synapses, and regulating synaptic activation [*SHISA6* (shisa family member 6) in a Finnish population sample]. Mick and colleagues (2011) found associations for SNPs that lie inside or close to multiple genes, including *LRRC7* (leucine rich repeat containing 7), involved in neuronal excitability and used as postsynaptic marker of hippocampal glutamatergic synapse integrity, and

STIP1 (stress induced phosphoprotein 1), involved in astrocyte differentiation and highly expressed in the brain. A second GWAS by Mick and colleagues (2014) observed a nominal association of proneness to anger with the gene, involved in calcium influx and release in the post-synaptic density, and in long-term potentiation (*FYN* (FYN proto-oncogene, Src family tyrosine kinase)). McGue et al. (2013) reported four SNPs associated with behavioural disinhibition including symptoms of CD and aggression, one of which (rs1368882; $p = 1.90E^{-06}$) was located inside the *GLIS1* (GLIS family zinc finger 1) gene responsible for a transcription factor that is involved in regulating the expression of numerous genes.

Recently, two larger studies attempted to identify genes associated with aggression or antisocial behaviour by increasing power through the inclusion of multiple cohorts. Pappa and colleagues (2016) collected a sample of 18,988 children 3 – 15 years for meta-analysis and reported a near genome-wide significant locus on chromosome 2p12 ($p = 5.3E^{-08}$). This locus is in close proximity to two genes: *LRRTM4* (leucine rich repeat transmembrane neuronal 4), which regulates excitatory synapse development, and *SNAR-H [small NF90 (ILF3) associated RNA H]*, which is implicated in the transcription process and is expressed in neurons. They found 19 genes nominally related to aggression from gene-based tests. Among them, *LRRTM4*, *PDSS2* (decaprenyl diphosphate synthase subunit 2), *TRIM27* (tripartite motif containing 27), *MRC1* (mannose receptor C-type 1), *MECOM* (MDS1 and EVI1 complex locus), and *CASC17* (cancer susceptibility 17).

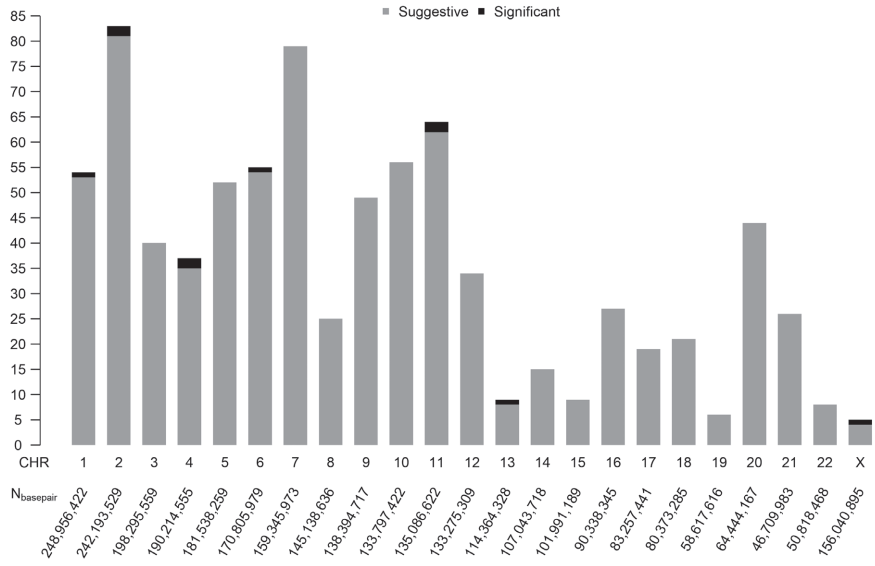
Another larger study by Tielbeek and colleagues (2017) focused on the broader antisocial behaviour phenotype in 16,400 individuals. The overall GWAMA found no hits, but sex-stratified GWAMAs returned three genome-wide significantly associated SNPs (minimum $p = 1.95E^{-08}$), but failed to identify significant genes. This suggested that there might be sex-specific genetic effects on antisocial behaviour and focusing on a more specific phenotype could improve chances of findings significant results.

Thus, nominal genome-wide associations ($p < 1E^{-05}$) have been found in genes involved in a wide variety of biological systems: the immune system, the endocrine system, pathways involved in neuronal development and differentiation and synaptic plasticity. These findings have not been replicated across GWASs, but some studies reported the same genes independently: *NTM* (Brevik et al.,

2016; Tiihonen et al., 2015) and *RBFOX1(A2BP1*) (Anney et al., 2008; Sonuga-Barke et al., 2008).

In summary, the 17 GWASs in our review show that genome-wide significant and/or suggestive associations between aggression-related traits and SNPs are found on all chromosomes (range: 1 - 63; see Supplement S5-6). As shown in Figure 3 near 55% of suggestive associations were found on chromosomes 1, 2, 5, 6, 7, 9, 10, and 11, with the majority of suggestive SNPs on chromosome 7 reported in the sample of African ancestry (Montalvo-Ortiz et al., 2018). The genome-wide significant associations are located on chromosomes 1, 2, 4, 6, 11, 13 and X.

Figure 3 Number of genetic variants associated with aggression-related traits at P≤1E-05 on different chromosomes reported the included GWAS studies



The X-axis shows chromosome number and length (in base pairs)
Nstudies = 17, Nvariants = 817

DISCUSSION

Aggression has a considerable genetic component, as indicated by decades of behaviour genetics research. However, no genomic variants have (yet) been identified. In our review covering GWASs on human aggression, only 4 out of 17 studies reported genome-wide significant hits in primary or replication samples (Dick et al., 2011; Rautiainen et al., 2016; Tielbeek et al., 2017; Montalvo-Ortiz et al., 2018). In the reviews on aggression and GWASs, several explanations are offered for the discrepancy between heritability estimates in behavioural and molecular genetic studies, for example the heterogeneous, context-dependent, and developmental nature of aggression, but foremost, small sample sizes. Fortunately, these limitations can be remedied, and provide future directions for research.

Most of the reviews covered, mention the often cited heritability estimates of 50% for aggression by Miles and Carey (1997), and 41% for antisocial behaviour by Rhee and Waldman (2002) and these estimates are confirmed in more recent empirical studies. Moderation, or any genotype x environment effects seem small, and most pronounced for non-aggressive antisocial behaviour (Burt et al., 2016).

How to address non-significant findings in GWAS studies on psychiatric problems is a pressing issue. Opinions are divided on what approach is most optimal to define phenotypes for GWAS analyses. Some believe that reduction of phenotypic heterogeneity could lead to more genome-wide significant findings (Anholt & Mackay, 2012; CONVERGE Consortium, 2015; Runions et al., 2019). This view is supported by the GWASs covered in this review that did find genome-wide significant hits. These relatively underpowered studies (Nrange = 2,185-6,220 participants) focus on individuals with severe antisocial behaviour and specific types of aggression: individuals with DSM-defined CD symptoms (Dick et al., 2011), cannabis-induced physical aggression (Montalvo-Ortiz et al., 2018), and criminal offenders with antisocial personality disorder (Rautiainen et al., 2016). Two studies were conducted in specific samples; exclusively male, with associations only in African-American subgroup (Montalvo-Ortiz et al., 2018), and predominantly male (89% of cases) and ethnically homogeneous (Rautiainen et al., 2016).

In contrast, other researchers propose a broader approach which includes more lenient phenotypes (Ormel, Hartman, & Snieder, 2019; Vassos, Collier, & Fazel, 2014). This lenient phenotyping approach has already achieved success in depression research, for example, although here the value of minimal versus

broader phenotyping is debated as well (Cai et al., 2019). The two largest GWASs on aggression that were covered by this review used broad, lenient measures of childhood aggression (Pappa et al., 2016) and antisocial behaviour (Tielbeek et al., 2017). Pappa and colleagues (2016) found no significant hits, but several promising loci on chromosomes 2, 3, 6 and 17 (minimum $p = 5.3E^{-08}$). Tielbeek and colleagues (2017) reported three significant hits for the sex-stratified GWAMAs.

Early linkage studies on aggression indicated chromosomes 1 (Criado et al., 2012), 2 and 19 (Dick et al., 2004) as potential loci. GWAS findings in our review confirm loci on chromosomes 1 and 2 which gave more associated variants and significant results. The X- and Y-chromosomes did not give evident results, even if one significant sign was reported in X-chromosome (Tielbeek et al., 2017).

In order to identify 80% of all causal SNPs, depending on the extent of SNP heritability, between 10^5 and 10^7 (100,000 – 10,000,000) independent subjects would be required (Holland et al., 2019). This means that, with sample sizes l0 time less than the lower bound, current GWASs were clearly underpowered. At present, several initiatives are under way to collaborate in achieving larger sample sizes. One example of a large collaborative project is the ACTION consortium (Aggression in Children: unraveling gene-environment interplay to inform Treatment and InterventiON strategies: http://www.action-euproject.eu/) which has brought together over 30 cohorts with childhood data on aggression for GWAS, EWAS and biomarker studies.

As mentioned, multiple reviews suggest that heterogeneity of aggression is a problem in research, with several reviews suggesting some kind of distinction between subtypes, subgroups, or developmental stages. Standardized phenotypic and environmental assessments are proposed as a solution (Craig & Halton, 2009). Although this standardization of assessment could be an option, recent advances in multivariate modelling allow for exploration of other potential avenues (e.g., Baselmans et al. 2019). This approach is also discussed in the meta-analyses of Zhang-James and Faraone (2016), in which aggression might be considered a multi-dimensional trait consisting of distinct, but related, constructs with shared aetiologies (Zhang-James and Faraone, 2016). In other words, although some individuals show different problem behaviours, including aggression, they all share a common genetic vulnerability. Taking a multivariate, approach would allow the inclusion of large cohorts with existing phenotypic (Bartels et al., 2018) and SNP data. However, the focus on ever broader phenotypes and bigger

samples, raises the question how to translate results into practice, to alleviate problems of individuals.

Future directions

We should recognize that the nature-nurture debate has moved on from the question whether aggressive behaviour is heritable to the discovery of the biological bases of aggression. This is currently achieved by investigating aggression's relation to genes, SNPs, and relevant biological pathways. It is expected that GWASs with larger or combined datasets will improve our understanding of the mechanisms of gene regulation of aggression. Individual GWASs on aggression and aggression-like traits are still limited in terms of explaining variation in the population, but ongoing GWASs and other efforts, e.g., in epigenetics and biomarker studies are likely provide insight into the aetiology of aggressive behaviour. Expansion of disease gene maps (Goh et al., 2007) by including aggression-related traits into, for example, OMIM datasets can help in future analyses of underlying cellular network-based relationships between genes and functional modules of aggressive behaviour, and future work should determine if genes mediating aggression pathways are enriched in the polygenic background of disorders associated with aggression.

Also, leveraging on Genotype-Tissue Expression (GTEx; (eGTAxProject, 2017)) GWAS findings can be annotated with additional information and thereby identify biologically relevant systems. One particularly interesting source of biological annotation revolves expression quantitative trait loci (eQTL), i.e., SNPs that have been associated with gene expression levels. Once genome-wide hits are found, overlapping these with known eQTLs could identify genes that are of biological interest (Gusev et al., 2016; Lowe & Reddy, 2015; Zhu et al., 2016).

Systematic reviews with automated functions

The workload on selection process of researchers in our systematic review was around 60 hours (screening and selecting relevant papers from list of 2,069 records). By using automated procedures to screen for relevant literature for inclusion in systematic reviews, it was possible to save 39.1% (23.5 hours) of reading/scanning time. The downside of automated methods is that relevant literature can be missed. On the other hand, even an expert reviewer might omit studies that the automated procedures include. Optimization of the expert reviewer is covered by education and training, whereas optimization of automated

selection is under active development (Borah, Brown, Capers, & Kaiser, 2017; Cohen, Hersh, Peterson, & Yen, 2006; Khabsa, Elmagarmid, Ilyas, Hammady, & Ouzzani, 2016). We opted for a recent approach that utilizes a machine learning algorithm to obtain a selection of papers that could be relevant for this systematic review.

Although the automated systematic review tool we applied is quite new and is still under active development, we found that applying the machine learning approach as implemented in the software hosted at https://github.com (*Automated systematic reviews by using Deep Learning and Active Learning*, 2019) could be indeed of considerable aid to the researcher performing a systematic review solving problems of missed literature in screening phase due to human errors or excluded by searching algorithms.

For the benefit of further developments in automated selection approaches aiding the review process, we advise review authors to supply their search results as additional information to their work. These results can then serve for further refinement of literature search models. This would avoid double work across research groups, create a comprehensive overview of aggression literature, and increase our understanding of the genetic nature of human aggression.

CONCLUSIONS

Aggression in humans is a heritable trait, whose genetic basis largely remains to be uncovered. No sufficiently large genome-wide association studies have been carried out yet. With increases in sample size, we expect aggression to behave like other complex human traits for which GWAS has been successful. There are several ongoing efforts to achieve genome-significant GWAS findings – merging samples in consortia, replication strategies, searching for close phenotypes from other domains associated with aggression for sample extension, developing new approaches of partitioning genetic heterogeneity and sample stratification. Automated tools for systematic review, which are based on machine learning, could be used to optimize the integration of research findings from different studies.

