frontiers in Genetics

MINI REVIEW
published: xx July 2020
doi: 10.3389/fgene.2020.00884

Forensic Autosomal Short Tandem Repeats and Their Potential Association With Phenotype

Nicole Wyner[1]*, *Mark Barash*[1,2] and *Dennis McNevin*[1]

[1] Centre for Forensic Science, School of Mathematical and Physical Sciences, Faculty of Science, University of Technology Sydney, Sydney, NSW, Australia, [2] Department of Justice Studies, San José State University, San Jose, CA, United States

Forensic DNA profiling utilizes autosomal short tandem repeat (STR) markers to establish identity of missing persons, confirm familial relations, and link persons of interest to crime scenes. It is a widely accepted notion that genetic markers used in forensic applications are not predictive of phenotype. At present, there has been no demonstration of forensic STR variants directly causing or predicting disease. Such a demonstration would have many legal and ethical implications. For example, is there a duty to inform a DNA donor if a medical condition is discovered during routine analysis of their sample? In this review, we evaluate the possibility that forensic STRs could provide information beyond mere identity. An extensive search of the literature returned 107 articles associating a forensic STR with a trait. A total of 57 of these studies met our inclusion criteria: a reported link between a STR-inclusive gene and a phenotype and a statistical analysis reporting a *p*-value less than 0.05. A total of 50 unique traits were associated with the 24 markers included in the 57 studies. TH01 had the greatest number of associations with 27 traits reportedly linked to 40 different genotypes. Five of the articles associated TH01 with schizophrenia. None of the associations found were independently causative or predictive of disease. Regardless, the likelihood of identifying significant associations is increasing as the function of non-coding STRs in gene expression is steadily revealed. It is recommended that regular reviews take place in order to remain aware of future studies that identify a functional role for any forensic STRs.

Keywords: short tandem repeat, phenotype, forensic marker, DNA profiling, junk DNA, non-coding STRs

OPEN ACCESS

Edited by:
Kenneth K. Kidd, School of Medicine, Yale University, United States

Reviewed by:
Jianye Ge, The University of North Texas Health Science Center at Fort Worth, United States
Adrian Matthew Linacre, Flinders University, Australia

***Correspondence:**
Nicole Wyner
nicole.wyner@alumni.uts.edu.au

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *19 May 2020*
Accepted: *17 July 2020*
Published: *xx July 2020*

Citation:
Wyner N, Barash M and McNevin D (2020) Forensic Autosomal Short Tandem Repeats and Their Potential Association With Phenotype. Front. Genet. 11:884. doi: 10.3389/fgene.2020.00884

INTRODUCTION

Short tandem repeats (STRs) are short repeated sequences of DNA (2–6 bp) that account for approximately 3% of the human genome (Lander et al., 2001). The number of repeat units is highly variable among individuals, which offers a high power of discrimination when analyzed for identification purposes. It is a widely accepted notion that STRs are non-coding in nature and are therefore not implicated in gene expression (Tautz and Schlotterer, 1994; Ramel, 1997; Butler, 2006; Biscotti et al., 2015). There is increasing evidence, however, that non-coding DNA sequences such as STRs may be involved in gene regulation via various mechanisms, hence being associated with phenotype (Sawaya et al., 2013; Chen et al., 2016).

The first STR markers used in forensic casework were selected in 1994 by the Forensic Science Service (FSS) in the United Kingdom for a quadruplex amplification system consisting of four

tetranucleotide STRs—TH01, vWA, FES/FPS, and F13A1 (Kimpton et al., 1994). These markers were deemed suitable for PCR amplification due to their simple repeat sequences and their propensity to display regularly spaced alleles differing by four bases; however, the quadruplex system did not offer a high level of discrimination. In 1997, the Federal Bureau of Investigation (FBI) nominated 13 autosomal STR loci to form the core of the Combined DNA Index System (CODIS), a database consisting of profiles contributed by federal, state, and local forensic laboratories. Two of the markers initially selected by the FSS (vWA and TH01) were included within the core CODIS set, whereas FES/FPS and F13A01 were eventually discarded due to low levels of polymorphism. The core set was reviewed in 2010 with an additional seven STRs being implemented from January 1, 2017. The majority of commercially available DNA profiling kits are manufactured to include the core CODIS STR loci (Butler, 2006). In accordance with the DNA Identification Act of 1994, CODIS is bound by stringent privacy protection protocols, in that the stored DNA samples and subsequent analyses be used strictly for law enforcement identification purposes. The DNA Analysis Backlog Elimination Act of 2000 reaffirms that the markers used for forensic applications were specifically selected because they are not known to be associated with any known physical traits or medical characteristics.

The markers nominated for CODIS were specifically chosen due to their location within non-coding regions of the genome; however, claims that non-coding regions play no functional role have been contested in recent years (Cole, 2007; Kaye, 2007; Sarkar and Adshead, 2010). There is increasing evidence that there may be associations between certain STR alleles and medical conditions (von Wurmb-Schwark et al., 2011; Meraz-Rios et al., 2014). This should not be confused with situations where alleles or loci are diagnostic for medical conditions (e.g., trisomy). Additionally, the ability to infer biogeographical ancestry (BGA) from forensic STRs is possible (Graydon et al., 2009; Algee-Hewitt et al., 2016) with investigators using population-specific STR data as intelligence to guide enquiries (Lowe et al., 2001). BGA is correlated with some phenotypes such as blue eye color in Europeans (Gettings et al., 2014) and lighter skin color with increasing distance from the equator (Relethford, 1997). However, the STR genotype *per se* is not causative of BGA phenotype in any direct sense and is mostly associated with BGA as a result of genetic drift (as STRs for forensic use have been selected to exhibit Hardy Weinberg equilibrium). In the event that any CODIS markers are in future found to be linked to a medical condition or physical trait, the analysis of the DNA sample must still be used only for identification purposes pursuant to the DNA Identification Act of 1994.

Katsanis and Wagner (2013) assessed 24 CODIS loci for phenotypic associations, but found no evidence to support the disclosure of any biomedically relevant information. For example, despite the fact that the locus TH01 was associated with as many as 18 traits: from alcoholism to spinocerebellar ataxia, the authors state that association with these traits does not necessarily imply that individual genotypes are causative or predictive of a particular trait. Following this, a statement issued by the Scientific Working Group of DNA Analysis Methods [SWGDAM] (2013) restated that although alternate discoveries may be made in the future, current understanding is that the CODIS loci do not reveal any information beyond identity. There has only been one STR to date that has been removed from consideration as a marker used in human identity testing (Szibor et al., 2005). The STR locus HumARA is located within a coding region on the X-chromosome and has been linked to muscular dystrophy. HumARA is a trinucleotide repeat and these are known to be more prone to disease-causing expansions than tetranucleotide repeats (Orr and Zoghbi, 2007; Castel et al., 2010; Hannan, 2018).

MATERIALS AND METHODS

A systematic search of the literature was conducted across three databases (Web of Science, PubMed, and Google Scholar) between August and December 2018. Population data studies, allele frequency studies, validation studies, technique developments, single case reports, mutation analyses, off-ladder allele identification, loss of heterozygosity studies, and locus characterizations were excluded. Additional papers were located by back referencing relevant or similar studies. Following the literature search, each STR was analyzed in the University of California Santa Cruz (UCSC) Genome Browser (Human GRCh38/hg38 Assembly) using the following tracks: Mapping and Sequencing—Base Position-dense; STS Markers-full, Gene and Gene Prediction—GENCODE v29-full; NCBI RefSeq-pack, Phenotype and Literature—OMIM Alleles-full; OMIM Pheno Loci-full; OMIM Genes-full; HGMD Variants-full; GWAS Catalog-full, Regulation—ENCODE Regulation-show; RefSeq Func Elems-full, Variation—Common SNPs(151)-full; FlaggedSNPs(151)-full, Repeats—Microsatellite-full; Simple Repeats-full. The STRs investigated included the 20 CODIS core loci used by the FBI, three extra loci currently used in Australia (Penta E, Penta D, D6S1043), and SE33 which is a core STR in the German national database and has subsequently been incorporated into several European kits.

RESULTS AND DISCUSSION

A total of 57 association studies sourced from three databases met our inclusion criteria: a reported link between a STR-inclusive gene and a phenotype and a statistical analysis reporting a p-value less than 0.05. Fifty unique traits were identified across the 24 markers (**Supplementary Table 1**). Schizophrenia was the trait most frequently described with a total of 11 studies reporting data on 14 different polymorphisms potentially associated with eight loci. Two separate articles investigated the allelic frequency amongst people who attempted suicide and reported a significantly higher frequency amongst 10 different alleles of seven forensic loci. The intronic STR TH01 had the greatest number of studies with 26 reports describing 27 traits potentially linked to 40 different genotypes. Five of these studies were investigating a link to schizophrenia, reporting five polymorphisms that are possibly associated with the disease.

No studies associating alleles or genotypes with phenotype were found for Penta E, Penta D, D3S1358, SE33, or D10S1248; however, one study by Shi et al. (2012) investigated the method of diagnosing Down syndrome by testing for a trisomy at the Penta D locus as it is located on chromosome 21. Similarly, six of the 10 articles included for D21S11 were investigating the marker's efficiency in genetic tests for Down syndrome.

Of the 57 articles proposing an association between a forensic STR and a phenotype, none of them confirmed any particular genotype to be solely causative of a phenotype. Despite 13 of the STRs being located within a functional gene, there were no entries in the Online Mendelian Inheritance in Man (OMIM) database relating any STR-inclusive regions of these genes with a disease. A stand-out result is the number of studies reporting an association between a phenotype with polymorphisms at the TH01 locus.

TH01

TH01 is located within the first intron of the tyrosine hydroxylase (TH) gene and is commonly characterized by the repeat motif $[AATG]_n$ or alternatively by the $[TCAT]_n$ motif, according to GenBank top strand nomenclature. TH is the rate-limiting enzyme involved in the biosynthesis of the catecholamines dopamine, epinephrine, and norepinephrine. Catecholamines act as both neurotransmitters and hormones that assist in maintaining homeostasis (Eisenhofer et al., 2004). As such, a strong relationship has been reported in the literature (Eisenhofer et al., 2004; Ng et al., 2015) between variations in the expression of TH and the development of neurological, psychiatric, and cardiovascular diseases.

Previous studies (McEwen, 2002; Antoni et al., 2006; Bastos et al., 2018) have shown that increased levels of epinephrine and norepinephrine are expressed in individuals experiencing acute or chronic stress. Wei et al. (1997) found that individuals carrying the TH01-9 allele showed the highest levels of serum norepinephrine amongst a population of unrelated healthy adults, whereas carriers of the TH01-7 allele showed the lowest. Barbeau et al. (2003) investigated the relationship between the number of TH01 repeats and hemodynamic parameters in subjects at rest and in response to applied stressors. The results of this study indicate that the 6 and 9.3 TH01 alleles are associated with a decrease in the hemodynamic responses to stress, offering a protective effect to individuals carrying those alleles. Carriers of the TH01-6 allele displayed a lower heart rate reactivity when exposed to stressors with increasing age than those without the TH01-6 allele. Furthermore, individuals carrying TH01-9.3 showed no increase in systolic blood pressure in response to stress, whereas those not possessing the TH01-9.3 allele demonstrated a significant increase in systolic blood pressure reactivity with increasing age. Conversely, the TH01-7 allele was found to be detrimental to blood pressure in those with a greater body mass index (BMI). Subjects carrying TH01-7 displayed a higher resting systolic blood pressure as BMI increases and increased heart rate reactivity in response to stressors with increasing BMI.

TH01-7 was also reported to be significantly more prevalent in patients prone to depression (Chiba et al., 2000). The TH01-8 allele was found more frequently in suicide attempters (Persson et al., 1997), individuals with depression (Serretti et al., 1998), and individuals with delusional disorder (Morimoto et al., 2002). Persson et al. (2000) investigated the influence of the number of TH01 repeats on 30 personality dimensions. Subjects possessing the TH01-8 allele scored higher in the neuroticism facets with significant differences observed between individuals displaying anger, hostility and vulnerability (Persson et al., 2000), compared to non-TH01-8 allele carriers. Nine repeats at the TH01 locus were associated with delusional disorder (Morimoto et al., 2002) and extraversion (Tochigi et al., 2006). Furthermore, Yang et al. (2011) conducted a number of association studies in China and reported that the frequency of TH01-9.3 was higher in those displaying suicidal behavior, and TH01-10 was significantly overrepresented in individuals demonstrating violent behavior including sexual assaults (Yang et al., 2010) and in males with impulsive violent behavior (Yang et al., 2013). TH01 was also linked to various disease states such as schizophrenia (Jacewicz et al., 2006b), predisposition to malaria (Gaikwad et al., 2005; Alam et al., 2011), sudden infant death syndrome (SIDS) (Klintschar et al., 2008; Courts and Madea, 2011), and Parkinson's disease (Sutherland et al., 2008).

As previously mentioned, TH catalyzes the conversion of tyrosine to levodopa (L-DOPA) which is then converted to dopamine. Dopamine can be further converted into norepinephrine and epinephrine. *In vitro* experiments have previously demonstrated that TH01 can regulate TH gene transcription, displaying a quantitative silencing effect (Albanèse et al., 2001). TH01 alleles inhibited transcription proportionally to the number of repeats. Given that so many vital functions rely on the presence of dopamine and its metabolites (Wei et al., 1997; Meiser et al., 2013), malfunctions of dopaminergic pathways have been associated with the development of numerous psychological diseases (Meiser et al., 2013), and in this review, TH01 was largely connected with schizophrenia (Kurumaji et al., 2001) and Parkinson's disease (Meiser et al., 2013). The longer TH01-9.3 and TH01-10 alleles, predicted to yield less dopamine, were found more frequently in individuals displaying traits indicative of dopaminergic dysfunction such as impulsive violent behavior (Yang et al., 2013), sexual assault (Yang et al., 2010), and addiction (Sander et al., 1998; Anney et al., 2004).

Some contradictory associations were observed between TH01 and certain phenotypes. For instance, De Benedictis et al. (1998) reported a significant association of >9 TH01 repeats with longevity in male Italian centenarians. Contrariwise, von Wurmb-Schwark et al. (2011) were unable to replicate this result when using the same study design on a German population, just as Bediaga et al. (2015) were also unable to confirm an association in a northern Spanish population. Similarly, there are conflicting reports on the association of TH01-9.3 with SIDS across European populations. In 2008, Klintschar et al. (2008) found that the frequency of the TH01-9.3 allele was significantly higher in SIDS patients than in controls in a German population. This association was further confirmed by Courts and Madea (2011). On the contrary, Studer et al. (2014) were unable to replicate this result in a Swiss population. Further

population-based association studies are needed to confirm the existence of associations between TH01 and these phenotypes.

None of the studies investigating TH01 have identified any of the associated genotypes as being causative of disease; therefore, the associations mentioned should only be considered as possible or potential. Many of the traits reported to be associated with TH01 are multifactorial, meaning they are affected by both genes and the environment, such as in the case of Parkinson's disease (Meiser et al., 2013) and schizophrenia (Zhuo et al., 2019).

Potential Associations of Other STR Markers

Schizophrenia is a complex heritable mental health disorder characterized by delusions, hallucinations, and impaired social cognition. It is understood that schizophrenia is polygenic with disease burdening alleles being distributed across multiple loci (Giusti-Rodríguez and Sullivan, 2013; Zhuo et al., 2019). Consistent with this notion, our study revealed that schizophrenia was associated with the greatest number of STRs: FGA, TH01, vWA, D2S441, D2S1338, D8S1179, D16S539, and D18S51. One study (Jacewicz et al., 2006a) found that longer repeats in D18S51 and D2S1338 were significantly more frequent in patients than in controls. This trend is consistent with the expansion of trinucleotide repeats in other major psychiatric disorders. Although the inherent complexity of the disease has posed a challenge to researchers, neurotransmitter abnormalities have long been acknowledged as a major contributing factor in the pathogenesis of schizophrenia (Mäki et al., 2005; Modai and Shomron, 2016).

Genetic mutations alone are not enough to trigger the onset and development of schizophrenia; therefore, further research is required in order to explore how genetic risk factors interact with environmental risk factors in the development, onset, and progression of the condition.

Venous thromboembolism (VTE) is a disorder defined by the occurrence of deep vein thrombosis and/or pulmonary embolism. vWF is a glycoprotein that plays a role in platelet adhesion during coagulation; therefore, it is understood that alterations in serum levels of vWF can contribute to thrombosis disorders (Laird et al., 2007). Meraz-Rios et al. (2014) found that vWA-18, TPOX-9, and TPOX-12 were observed more frequently in individuals with venous thrombosis in the Mexican mestizo population. Furthermore, vWA and TPOX have been associated with chronic myeloid leukemia (Wang et al., 2012).

Trisomys

Down syndrome, or Trisomy-21, can be diagnosed by the presence of a third allele at chromosome 21. This trisomy can be present at any polymorphic marker found on chromosome 21, and there are several studies evaluating the use of D21S11 and Penta D as effective markers in Down syndrome detection (Yoon et al., 2002; Liou et al., 2004; Shi et al., 2012; Guan et al., 2013). Similarly, D18S51 and D13S317 can be used as genetic markers to diagnose the presence of Edwards syndrome (Trisomy-18) and Patau syndrome (Trisomy-13), respectively. Trisomys are an example of a causal association as all individuals with three chromosomes will be affected. While the presence of an extra allele at chromosomes 13, 18, or 21 does not reveal a medical condition unknown to the donor, it does provide additional identifiable information to investigators.

Cancer

Forensic STRs have been used as genetic markers in several studies to screen for cancer-related alleles. Hui et al. (2014) found that two pairs of alleles (D8S1179-16 with D5S818-13 and D2S1338-23 with D6S1043-11) were found more frequently in gastric cancer patients. Furthermore, a study from China identified a significant association between homozygous alleles at D6S1043 and an increased risk of invasive cervical cancer (Wu et al., 2008). Loss of heterozygosity (LOH) is a genetic mutation that results in the loss of one copy of a heterozygous gene, often resulting in cancer due to loss of functional tumor suppressor genes. LOH in different cancer tissues have been observed at a number of forensic loci such as CSF1PO, FGA, vWA, D3S1358, D5S818, D8S1179, D13S317, and D18S51 in patients with laryngeal cancer (Rogowski et al., 2004). LOH may alter the results of a DNA profile and should be taken into consideration in cases where only cancerous tissue is available for analysis (Peloso et al., 2003; Zhou et al., 2017).

Qi et al. (2018) conducted a study investigating the possibility of using genetic markers rather than related genes to screen for predisposition to lung and liver cancer. This study used CODIS markers to examine the theory of programed onset which hypothesizes that the occurrence of a chronic disease is independent of age and may instead depend on a programed onset pattern. The results showed a significant difference in the occurrence of lung cancer between those who carried the D18S51-20 allele and those who did not, and the incidence of liver cancer between those carrying D21S11-30.2 and D6S1043-18 alleles and those who did not. While these results demonstrate CODIS markers being used to predict an individual's predisposition to cancer, there are an extensive number of cancer-related genes in the genome; therefore, the risk of breaching genetic privacy with this information remains low.

Y and X STRs

The Y chromosome has accumulated male advantage and fertility genes (Lahn and Page, 1997; Graves, 2006) and so it is possible that phenotypes associated with maleness are associated with Y STRs. X-linked phenotypes (as a result of recessive genes on the X chromosome) are more prevalent in males (because there is no dominant Y chromosome homolog) so there may also be associations with X STRs. In fact, X-linked genes have recently been shown to influence male fertility and sex ratio of offspring in mice (Kruger et al., 2019).

Association Versus Causation

The association of a STR with a trait or disease does not infer causation. Moreover, some alleles seem to have opposite effects: TH01 allele 9.3 may help with stress (Zhang et al., 2004) but also has a potential link with suicide (Persson et al., 1997; Yang et al., 2011). A genetic variant is considered causative when it is known that the presence of the variant will produce an effect that in turn

causes disease (Hu et al., 2018). None of the associations reported in this study offer proof of causation (except for trisomys), rather they propose a general relationship between some STRs used in forensic applications and a phenotype. These relationships may also be explained by confounding variables, bias, or by chance in cases where a significant finding is unable to be replicated by another study. In fact, this review could be seen as a reflection of the broader so-called "replication crisis" in science (Schooler, 2014). Many of the studies reported in this review may not have sufficiently mitigated against the "multiple comparison problem" where a number of comparisons will be significant by chance. By setting our *p*-value threshold to 0.05, we run the risk that 5% of significant results are significant by chance.

Many of the traits that can be predicted by genetic analysis are the result of epistatic interactions between genes and environmental factors. When considering the associations in this review, it is not reasonable to suggest that an individual possessing the more frequently observed allele associated with a trait will express a specific phenotype. There are many underlying mechanisms involved in the development of complex diseases and while the risk of forensic STRs being found to expose revealing medical information is minimal, the presence of a particular allele may indicate heightened potential or risk for a phenotype.

Molecular Mechanisms

While it remains true that forensic markers are located within non-coding regions, there is growing evidence that STRs in introns and up- or down-stream of genes may affect phenotype. STR mutations in the 5′ untranslated region (UTR) are known to modify gene expression, probably because they serve as protein binding sites (Li et al., 2004). Mutations in the 3′ UTR result in extended mRNA which can be toxic to the cell (Li et al., 2004; La Spada and Taylor, 2010). There are 13 CODIS STRs located in introns (**Supplementary Table 2**). Mutations in introns can affect mRNA splicing which can result in gene silencing or loss of function (Li et al., 2004; La Spada and Taylor, 2010). The TCAT repeat in the first intron of TH01 acts as a transcription regulatory element *in vitro* (Meloni et al., 1998). Albanèse et al. (2001) reported a reduction in transcriptional activity of TH as the TCAT repeat number varied from three to eight. STRs are also found at high density in promoter regions and it is highly likely that some are implicated in gene expression by modulating spacing of regulatory elements (Gemayel et al., 2012; Sawaya et al., 2013; Gymrek et al., 2016; Quilez et al., 2016; Gymrek, 2017).

There is now etiological support for STRs as causative agents for disease in that they are quite plausibly epigenetic regulators for gene expression when located in introns or up- or down-stream of genes. This may increase prior support for the hypotheses of association and thus reduce the required significance level, as described by Kidd (1993), which is a counter to the "multiple comparison problem" discussed earlier.

CONCLUSION

While the results of this study did indicate a large number of phenotypic traits associated with forensic STRs, none were found to be independently causative or predictive of disease. Nevertheless, as there are numerous reported instances of tetranucleotide repeats being implicated in disease and molecular mechanisms have been demonstrated, there remains a strong chance that this inference may change in the near future. One limitation of this study was the sole use of the UCSC genome browser. Future studies may benefit from using a wider range of resources and investigating additional markers such as SNPs in flanking regions, mtDNA and Y-STRs. In the event that a statistically significant association, causal or predictive relationship is discovered, it is not necessarily a valid cause for removal from STR panels, but additional protective measures, such as tightening legislation surrounding genetic privacy, may need to be considered to prevent abuse of this information.

AUTHOR CONTRIBUTIONS

NW designed the study, performed the literature review and wrote the manuscript. MB conceived the project, designed the study, and reviewed and edited the manuscript. DM conceived and managed the project, designed the study, and reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00884/full#supplementary-material

REFERENCES

Alam, S., Ferdous, A., Ali, M. E., Ahmed, A., Naved, A. F., and Akhteruzzaman, S. (2011). Forensic microsatellite TH01 and malaria predisposition. *Dhaka Univ. J. Biol. Sci.* 20, 1–6. doi: 10.3329/dujbs.v20i1.8831

Albanèse, V., Biguet, N. F., Kiefer, H., Bayard, E., Mallet, J., and Meloni, R. (2001). Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite. *Hum. Mol. Genet* 10, 1785–1792. doi: 10.1093/hmg/10.17.1785

Algee-Hewitt, B. F. B., Michael, E. D., Kim, J., Li, J. Z., and Rosenberg, N. A. (2016). Individual identifiability predicts population identifiability in forensic microsatellite markers. *Curr. Biol.* 26, 935–942. doi: 10.1016/j.cub.2016.01.065

Anney, R. J. L., Olsson, C. A., Lotfi-Miri, M., Patton, G. C., and Williamson, R. (2004). Nicotine dependence in a prospective population-based study of adolescents: the protective role of a functional tyrosine hydroxylase polymorphism. *Pharmacogenetics* 14, 73–81. doi: 10.1097/01.fpc.0000054157.92680.b6

Antoni, M. H., Lutgendorf, S. K., Cole, S. W., Dhabhar, F. S., Sephton, S. E., McDonald, P. G., et al. (2006). The influence of bio-behavioural factors on tumour biology: pathways and mechanisms. *Nat. Rev. Cancer* 6, 240–248. doi: 10.1038/nrc1820

Barbeau, P., Litaker, M. S., Jackson, R. W., and Treiber, F. A. (2003). A tyrosine hydroxylase microsatellite and hemodynamic response to stress in a multi-ethnic sample of youth. *Ethn. Dis.* 13, 186–192.

Bastos, D. B., Sarafim-Silva, B. A. M., Sundefeld, M., Ribeiro, A. A., Brandao, J. D. P., Biasoli, E. R., et al. (2018). Circulating catecholamines are associated with biobehavioral factors and anxiety symptoms in head and neck cancer patients. *PLoS One* 13:e0202515. doi: 10.1371/journal.pone.0202515

Bediaga, N. G., Aznar, J. M., Elcoroaristizabal, X., Alboniga, O., Gomez-Busto, F., Artabe, I. A., et al. (2015). Associations between STR autosomal markers and longevity. *Age* 37:95. doi: 10.1007/s11357-015-9818-5

Biscotti, M. A., Olmo, E., and Heslop-Harrison, J. S. (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Res* 23, 415–420. doi: 10.1007/s10577-015-9499-z

Butler, J. M. (2006). Genetics and genomics of core short tandem repeat loci used in human identity testing. *J. Forensic Sci.* 51, 253–265. doi: 10.1111/j.1556-4029.2006.00046.x

Castel, A. L., Cleary, J. D., and Pearson, C. E. (2010). Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. Cell Biol.* 11, 165–170. doi: 10.1038/nrm2854

Chen, H. Y., Ma, S. L., Huang, W., Ji, L., Leung, V. H. K., Jiang, H., et al. (2016). The mechanism of transactivation regulation due to polymorphic short tandem repeats (STRs) using IGF1 promoter as a model. *Sci. Rep.* 6:38225. doi: 10.1038/srep38225

Chiba, M., Suzuki, S., Hinokio, Y., Hirai, M., Satoh, Y., Tashiro, A., et al. (2000). Tyrosine hydroxylase gene microsatellite polymorphism associated with insulin resistance in depressive disorder. *Metabolism* 49, 1145–1149. doi: 10.1053/meta.2000.8611

Cole, S. A. (2007). Is the "Junk". DNA designation bunk? *Nw. UL Rev. Colloquy* 102, 54–63.

Courts, C., and Madea, B. (2011). Significant association of TH01 allele 9.3 and SIDS. *J. Forensic Sci.* 56, 415–417. doi: 10.1111/j.1556-4029.2010.01670.x

De Benedictis, G., Carotenuto, L., Carrieri, G., De Luca, M., Falcone, E., Rose, G., et al. (1998). Gene/longevity association studies at four autosomal loci (REN, THO, PARP, SOD2). *Eur. J. Hum. Genet.* 6, 534–541. doi: 10.1038/sj.ejhg.5200222

Eisenhofer, G., Kopin, I. J., and Goldstein, D. S. (2004). Catecholamine metabolism: a contemporary view with implications for physiology and medicine. *Pharmacol.Rev.* 56, 331–349. doi: 10.1124/pr.56.3.1

Gaikwad, S., Ashma, R., Kumar, N., Trivedi, R., and Kashyap, V. K. (2005). Host microsatellite alleles in malaria predisposition? *Malar. J.* 4, 331–349. doi: 10.1186/1475-2875-4-50

Gemayel, R., Cho, J., Boeynaems, S., and Verstrepen, K. J. (2012). Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. *Genes* 3, 461–480. doi: 10.3390/genes3030461

Gettings, K. B., Lai, R., Johnson, J. L., Peck, M. A., Hart, J. A., Gordish-Dressman, H., et al. (2014). A 50-SNP assay for biogeographic ancestry and phenotype prediction in the U.S. population. *Forensic Sci. Int. Genet.* 8, 101–108. doi: 10.1016/j.fsigen.2013.07.010

Giusti-Rodríguez, P., and Sullivan, P. F. (2013). The genomics of schizophrenia: update and implications. *J. Clin. Invest.* 123, 4557–4563. doi: 10.1172/JCI66031

Graves, J. A. (2006). Sex chromosome specialization and degeneration in mammals. *Cell* 124, 901–914. doi: 10.1016/j.cell.2006.02.024

Graydon, M., Cholette, F., and Ng, L.-K. (2009). Inferring ethnicity using 15 autosomal STR loci—Comparisons among populations of similar and distinctly different physical traits. *Forensic Sci. Int. Genet.* 3, 251–254. doi: 10.1016/j.fsigen.2009.03.002

Guan, L., Ren, C., Li, H., Gao, L., Jia, N., and Guan, H. (2013). [Practicality of rapid prenatal screening for Down syndrome with PCR-short tandem repeat method]. *Chinese J. of Med. Gen.* 30, 277–282. doi: 10.3760/cma.j.issn.1003-9406.2013.03.006

Gymrek, M. (2017). A genomic view of short tandem repeats. *Curr. Opin. Genet. Dev.* 44, 9–16. doi: 10.1016/j.gde.2017.01.012

Gymrek, M., Willems, T., Guilmatre, A., Zeng, H., Markus, B., Georgiev, S., et al. (2016). Abundant contribution of short tandem repeats to gene expression variation in humans. *Nat. Genet.* 48, 22–29. doi: 10.1038/ng.3461

Hannan, A. J. (2018). Tandem repeats mediating genetic plasticity in health and disease. *Nat. Rev. Genet.* 19, 286–298. doi: 10.1038/nrg.2017.115

Hu, P., Jiao, R., Jin, L., and Xiong, M. (2018). Application of causal inference to genomic analysis: advances in methodology. *Front. Genet.* 9:238. doi: 10.3389/fgene.2018.00238

Hui, L., Liping, G., Jian, Y., and Laisui, Y. (2014). A new design without control population for identification of gastric cancer-related allele combinations based on interaction of genes. *Gene* 540, 32–36. doi: 10.1016/j.gene.2014.02.033

Jacewicz, R., Babol-Pokora, K., Berent, J., Pepinski, and Szram, S. (2006a). Are tetranucleotide microsatellites implicated in neuropsychiatric diseases? *Int. Congr. Ser.* 1288, 783–785. doi: 10.1016/j.ics.2005.09.101

Jacewicz, R., Szram, S., Gałecki, P., and Berent, J. (2006b). Will genetic polymorphism of tetranucleotide sequences help in the diagnostics of major psychiatric disorders? *Forensic Sci. Int.* 162, 24–27. doi: 10.1016/j.forsciint.2006.06.024

Katsanis, S. H., and Wagner, J. K. (2013). Characterization of the standard and recommended CODIS markers.*J. Forensic Sci.* 58(Suppl. 1), S169–S172. doi: 10.1111/j.1556-4029.2012.02253.x

Kaye, D. H. (2007). Please, Let's Bury the Junk: the CODIS Loci and the Revelation of Private Information. *Nw. UL Rev. Colloquy* 102:70.

Kidd, K. K. (1993). Associations of disease with genetic markers: déjà vu all over again. *Am. J. Med. Genet.* 48, 71–73. doi: 10.1002/ajmg.1320480202

Kimpton, C., Fisher, D., Watson, S., Adams, M., Urquhart, A., Lygo, J., et al. (1994). Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. *Int. J. Legal Med.* 106, 302–311. doi: 10.1007/BF01224776

Klintschar, M., Reichenpfader, B., and Saternus, K. S. (2008). A functional polymorphism in the tyrosine hydroxylase gene indicates a role of noradrenalinergic signaling in sudden infant death syndrome. *J. Pediatr.* 153, 190–193. doi: 10.1016/j.jpeds.2008.02.032

Kruger, A. N., Brogley, M. A., Huizinga, J. L., Kidd, J. M., de Rooij, D. G., Hu, Y. C., et al. (2019). A neofunctionalized X-linked ampliconic gene family is essential for male fertility and equal sex ratio in mice. *Curr. Biol.* 29, 3699.e5–3706.e5. doi: 10.1016/j.cub.2019.08.057

Kurumaji, A., Kuroda, T., Yamada, K., Yoshikawa, T., and Toru, M. (2001). An association of the polymorphic repeat of tetranucleotide (TCAT) in the first intron of the human tyrosine hydroxylase gene with schizophrenia in a Japanese sample. *J. Neural Transm.* 108, 489–495. doi: 10.1007/s007020170069

La Spada, A. R., and Taylor, J. P. (2010). Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.* 11, 247–258. doi: 10.1038/nrg2748

Lahn, B. T., and Page, D. C. (1997). Functional coherence of the human Y chromosome. *Science* 278, 675–680. doi: 10.1126/science.278.5338.675

Laird, R., Schneider, P. M., and Gaudieri, S. (2007). Forensic STRs as potential disease markers: a study of VWA and von Willebrand's Disease. *Forensic Sci. Int. Genet.* 1, 253–261. doi: 10.1016/j.fsigen.2007.06.002

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. doi: 10.1038/35057062

Li, Y. C., Korol, A. B., Fahima, T., and Nevo, E. (2004). Microsatellites within genes: structure, function, and evolution. *Mol. Biol. Evol.* 21, 991–1007. doi: 10.1093/molbev/msh073

Liou, J. D., Chu, D. C., Cheng, P. J., Chang, S. D., Sun, C. F., Wu, Y. C., et al. (2004). Human chromosome 21-specific DNA markers are useful in prenatal detection of Down syndrome. *Ann. Clin. Lab. Sci.* 34, 319–323.

Lowe, A. L., Urquhart, A., Foreman, L. A., and Evett, I. W. (2001). Inferring ethnic origin by means of an STR profile. *Forensic Sci. Int.* 119, 17–22. doi: 10.1016/S0379-0738(00)00387-X

Mäki, P., Veijola, J., Jones, P. B., Murray, G. K., Koponen, H., Tienari, P., et al. (2005). Predictors of schizophrenia—a review. *Br. Med. Bull.* 73-74, 1–15. doi: 10.1093/bmb/ldh046

McEwen, B. S. (2002). Sex, stress and the hippocampus: allostasis, allostatic load and the aging process. *Neurobiol. Aging* 23, 921–939. doi: 10.1016/S0197-4580(02)00027-1

Meiser, J., Weindl, D., and Hiller, K. (2013). Complexity of dopamine metabolism. *Cell Commun. Signal.* 11:34. doi: 10.1186/1478-811X-11-34

Meloni, R., Albanese, V., Ravassard, P., Treilhou, F., and Mallet, J. (1998). A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element in vitro. *Hum. Mol. Genet.* 7, 423–428. doi: 10.1093/hmg/7.3.423

Meraz-Rios, M. A., Majluf-Cruz, A., Santana, C., Noris, G., Camacho-Mejorado, R., Acosta-Saavedra, L. C., et al. (2014). Association of vWA and TPOX polymorphisms with venous thrombosis in Mexican mestizos. *BioMed Res. Int.* 9:697689. doi: 10.1155/2014/697689

Modai, S., and Shomron, N. (2016). Molecular risk factors for schizophrenia. *Trends. Mol. Med.* 22, 242–253. doi: 10.1016/j.molmed.2016.01.006

Morimoto, K., Miyatake, R., Nakamura, M., Watanabe, T., Hirao, T., and Suwaki, H. (2002). Delusional disorder: molecular genetic evidence for dopamine psychosis. *Neuropsychopharmacology* 26, 794–801. doi: 10.1016/S0893-133X(01)00421-3

Ng, J., Papandreou, A., Heales, S. J., and Kurian, M. A. (2015). Monoamine neurotransmitter disorders—clinical advances and future perspectives. *Nat. Rev. Neurol.* 11, 567–584. doi: 10.1038/nrneurol.2015.172

Orr, H. T., and Zoghbi, H. Y. (2007). Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* 30, 575–621. doi: 10.1146/annurev.neuro.29.051605.113042

Peloso, G., Grignania, P., Rosso, R., and Previdere, C. (2003). "Forensic evaluation of tetranucleotide STR instability in lung cancer," in *Progress in Forensic Genetics 9*, eds B. Brinkman and A. Carracedo (Amsterdam: Elsevier Science Bv), 719–721. doi: 10.1016/s0531-5131(02)00500-9

Persson, M. L., Wasserman, D., Geijer, T., Jonsson, E. G., and Terenius, L. (1997). Tyrosine hydroxylase allelic distribution in suicide attempters. *Psychiatry Res.* 72, 73–80. doi: 10.1016/s0165-1781(97)00068-1

Persson, M. L., Wasserman, D., Jonsson, E. G., Bergman, H., Terenius, L., Gyllander, A., et al. (2000). Search for the influence of the tyrosine hydroxylase (TCAT)n repeat polymorphism on personality traits. *Psychiatry Res.* 95, 1–8. doi: 10.1016/s0165-1781(00)00160-8

Qi, X., Yu, Y. J., Ji, N., Ren, S. S., Xu, Y. C., and Liu, H. (2018). Genetic risk analysis for an individual according to the theory of programmed onset, illustrated by lung and liver cancers. *Gene* 673, 107–111. doi: 10.1016/j.gene.2018.06.044

Quilez, J., Guilmatre, A., Garg, P., Highnam, G., Gymrek, M., Erlich, Y., et al. (2016). Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. *Nucleic Acids Res.* 44, 3750–3762. doi: 10.1093/nar/gkw219

Ramel, C. (1997). Mini- and microsatellites. *Environ. Health Perspect.* 105, 781–789. doi: 10.2307/3433284

Relethford, J. H. (1997). Hemispheric difference in human skin color. *Am. J. Phys. Anthropol.* 104, 449–457. doi: 10.1002/(sici)1096-8644(199712)104:4<449::aid-ajpa2>3.0.co;2-n

Rogowski, M., Walenczak, I., Pepinski, W., Skawronska, M., Sieskiewicz, A., and Klatka, J. (2004). Loss of heterozygosity in laryngeal cancer. *Rocz. Akad. Med. Bialymst.* 49, 262–264.

Sander, T., Harms, H., Rommelspacher, H., Hoehe, M., and Schmidt, L. G. (1998). Possible allelic association of a tyrosine hydroxylase polymorphism with vulnerability to alcohol-withdrawal delirium. *Psychiatr. Genet.* 8, 13–17.

Sarkar, S. P., and Adshead, G. (2010). Whose DNA is it anyway? European court, junk DNA, and the problem with prediction. *J. Am. Acad. Psychiatry Law* 38, 247–250.

Sawaya, S., Bagshaw, A., Buschiazzo, E., Kumar, P., Chowdhury, S., Black, M. A., et al. (2013). Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. *PLoS One* 8:e54710. doi: 10.1371/journal.pone.0054710

Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. *Nature* 515:9. doi: 10.1038/515009a

Scientific Working Group of DNA Analysis Methods [SWGDAM] (2013). *SWGDAM Considerations for Claims That the CODIS Core Loci are 'Associated' With Medical Conditions/Diseases*. Quantico, VA: SWGDAM.

Serretti, A., Macciardi, F., Verga, M., Cusin, C., Pedrini, S., and Smeraldi, E. (1998). Tyrosine hydroxylase gene associated with depressive symptomatology in mood disorder. *Am. J. Med. Genet.* 81, 127–130. doi: 10.1002/(SICI)1096-8628(19980328)81:2<127::AID-AJMG1>3.0.CO;2-T

Shi, Y. F., Li, X. Z., Li, Y., Zhang, X. L., Zhang, Y., and Yue, T. F. (2012). [Diagnosis of Downs syndrome using short tandem repeat loci D21S11, D21S1440 and Penta D]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 29, 443–446. doi: 10.3760/cma.j.issn.1003-9406.2012.04.014

Studer, J., Bartsch, C., and Haas, C. (2014). Tyrosine hydroxylase TH01 9.3 allele in the occurrence of sudden infant death syndrome in swiss caucasians. *J. Forensic Sci.* 59, 1650–1653. doi: 10.1111/1556-4029.12526

Sutherland, G., Mellick, G., Newman, J., Double, K. L., Stevens, J., Lee, L., et al. (2008). Haplotype analysis of the IGF2-INS-TH gene cluster in Parkinson's disease. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 147B, 495–499. doi: 10.1002/ajmg.b.30633

Szibor, R., Hering, S., and Edelmann, J. (2005). The HumARA genotype is linked to spinal and bulbar muscular dystrophy and some further disease risks and should no longer be used as a DNA marker for forensic purposes. *Int. J. Legal Med.* 119, 179–180. doi: 10.1007/s00414-005-0525-0

Tautz, D., and Schlotterer. (1994). Simple sequences. *Curr. Opin. Genet. Dev.* 4, 832–837.

Tochigi, M., Otowa, T., Hibino, H., Kato, C., Otani, T., Umekage, T., et al. (2006). Combined analysis of association between personality traits and three functional polymorphisms in the tyrosine hydroxylase, monoamine oxidase A, and catechol-O-methyltransferase genes. *J. Neurosci. Res.* 54, 180–185. doi: 10.1016/j.neures.2005.11.003

von Wurmb-Schwark, N., Caliebe, A., Schwark, T., Kleindorp, R., Poetsch, M., Schreiber, S., et al. (2011). Association of TH01 with human longevity revisited. *Eur. J. Hum. Genet.* 19, 924–927. doi: 10.1038/ejhg.2011.43

Wang, Z. L., Dai, L., Li, S., Qiu, G. Q., and Wu, H. Q. (2012). [Comparison of allelic frequencies of 15 short tandem repeat loci between chronic myeloid leukemia patients and non-related healthy individuals]. *Chin. J. Med. Genet.* 29, 306–308. doi: 10.3760/cma.j.issn.1003-9406.2012.03.013

Wei, J., Ramchand, C. N., and Hemmings, G. P. (1997). Possible association of catecholamine turnover with the polymorphic (TCAT)n repeat in the first intron of the human tyrosine hydroxylase gene. *Life Sci.* 61, 1341–1347. doi: 10.1016/S0024-3205(97)00679-6

Wu, Y., Zhang, Q., Liu, B., and Yu, G. (2008). The analysis of the entire HLA, partial non-HLA and HPV for Chinese women with cervical cancer. *J. Med. Virol.* 80, 1808–1813. doi: 10.1002/jmv.21251

Yang, C., Ba, H., Gao, Z., Zhao, H., Yu, H., and Guo, W. (2013). Case-control study of allele frequencies of 15 short tandem repeat loci in males with impulsive violent behavior. *Shanghai Arch. Psychiatry* 25, 354–363. doi: 10.3969/j.issn.1002-0829.2013.06.004

Yang, C., Ba, H., and Zhao, H. (2011). Association study between the genetic polymorphism of 15 STR loci and the suicide behavior in Jiangsu province. *J. Psych.* 1:9.

Yang, C., Huajie, B., Gao, Z., Lin, Z., Zhao, H., Liu, B., et al. (2010). Association study between the genetic polymorphism of 15 STR loci and the crime of rape. *Chin. J. Behav. Med. Brain Sci.* 19, 421–424.

Yoon, H. R., Park, Y. S., and Kim, Y. K. (2002). Rapid prenatal detection of Down and Edwards syndromes by fluorescent polymerase chain reaction with short tandem repeat markers. *Yonsei Med. J.* 43, 557–566. doi: 10.3349/ymj.2002.43.5.557

Zhang, L., Rao, F., Wessel, J., Kennedy, B. P., Rana, B. K., Taupenot, L., et al. (2004). Functional allelic heterogeneity and pleiotropy of a repeat polymorphism in tyrosine hydroxylase: prediction of catecholamines and response to stress in twins. *Physiol. Genomics* 19, 277–291. doi: 10.1152/physiolgenomics.00151.2004

Zhou, S., Wang, H., Wang, Q. K., Wang, P., Wang, F., and Xu, C. (2017). Loss of heterozygosity detected at three short tandem repeat locus commonly used for human DNA identification in a case of paternity testing. *Legal Med.* 24, 7–11. doi: 10.1016/j.legalmed.2016.11.001

Zhuo, C., Hou, W., Li, G., Mao, F., Li, S., Lin, X., et al. (2019). The genomics of schizophrenia: shortcomings and solutions. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 93, 71–76. doi: 10.1016/j.pnpbp.2019.03.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.