European Journal of Human Genetics (2018) 26:1060–1064
https://doi.org/10.1038/s41431-018-0123-5

BRIEF COMMUNICATION

Molecular-genetic characterization of common, noncoding *UBASH3A* variants associated with type 1 diabetes

Yan Ge[1,2] · Patrick Concannon[1,2]

Received: 30 April 2017 / Revised: 27 January 2018 / Accepted: 6 February 2018 / Published online: 28 February 2018

Abstract

Genome-wide association and fine-mapping studies have identified over 40 susceptibility regions for type 1 diabetes (T1D), a common autoimmune disease; however, most of the disease-associated variants are noncoding, and it remains a challenge to understand their biological contributions to T1D pathogenesis. One identified T1D risk locus is located at chromosome 21q22.3 where the most likely candidate gene is *UBASH3A*, a negative regulator of NF-κB signaling. Various noncoding variants in *UBASH3A* have been shown to be associated with T1D or other autoimmune diseases. Here we investigated four such SNPs—rs11203202, rs80054410, rs11203203, and rs1893592. We discovered a novel role for rs1893592 in T1D and showed that its minor allele protects against T1D. Our haplotype analysis identified three T1D-associated *UBASH3A* haplotypes, and revealed that risk for T1D is affected by additive effects of these four *UBASH3A* variants. In human primary $CD4^+$ T cells, upon T-cell receptor stimulation, the minor allele of rs1893592 was associated with both a significant reduction in the overall mRNA levels of *UBASH3A*, and an increase in the proportion of a normally occurring, but low-abundant, *UBASH3A* transcript that retains intron-9 sequences and cannot produce full-length UBASH3A protein. This reduction in *UBASH3A*, as a consequence of the minor allele at rs1893592, resulted in increased secretion of IL-2, a key cytokine that is required for T-cell activation and function but is deficient in some T1D subjects. Our study provides new mechanistic insights into how rs1893592 affects T1D and autoimmunity, and how interactions between multiple T1D-associated, noncoding variants influence the disease risk.

Introduction

Type 1 diabetes (T1D) is a common, complex autoimmune disease arising from the destruction of the insulin-producing pancreatic β cells. Multiple genetic and environmental risk factors contribute to T1D. Although genome-wide association studies (GWAS) have discovered over 40 chromosomal regions where there is significant statistical evidence of association with T1D [1, 2], it remains challenging to pinpoint the causative genes and variants located in most of these regions. Furthermore, the majority of lead GWAS single-nucleotide polymorphisms (SNPs) for autoimmune diseases, including T1D, are noncoding, and their functions in disease pathogenesis are largely unknown. To tackle these problems, we herein investigated a group of disease-relevant, common noncoding SNPs in *UBASH3A*, the most likely candidate gene for the T1D risk locus on chromosome 21q22.3 [1–5].

UBASH3A is expressed primarily in T cells, and plays a broad role in autoimmunity; five different common autoimmune diseases have significant reported associations with SNPs in *UBASH3A* [1–11]. We previously showed that UBASH3A inhibits T-cell receptor-induced nuclear factor-κB (NF-κB) signaling and thereby suppresses *IL2* expression, a target gene of NF-κB, and that the minor alleles of rs80054410 (GRCh37 chr21:g.43836010T>C) and rs11203203 (GRCh37 chr21:g.43836186G>A)—two credible causative variants located in a potential enhancer in intron 4 of *UBASH3A* [2]—enhance *UBASH3A* expression in human primary T cells while diminishing *IL2* expression upon T-cell receptor stimulation [12]. These findings

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s41431-018-0123-5) contains supplementary material, which is available to authorized users.

✉ Patrick Concannon
patcon@ufl.edu

[1] Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL, USA

[2] Genetics Institute, University of Florida, Gainesville, FL, USA

Table 1 Single-marker family-based association tests of noncoding variants in the *UBASH3A* gene

SNP	Intron	cDNA change	MAF	Family (*n*)	*Z* score	*P*-value
rs11203202	1	c.114-1060C>G	0.36	1367	3.644	2.69E-04
rs80054410	4	c.554-541T>C	0.40	1411	3.521	4.30E-04
rs11203203	4	c.554-365G>A	0.39	1410	3.542	3.97E-04
rs1893592	9	c.1279 + 3A>C	0.25	1171	−4.207	2.60E-05

cDNA changes and intron numbers are based on the RefSeq sequences NM_001001895.2 and NG_029750.1. *Z* scores and *P*-values are calculated for the minor alleles. Positive and negative Z scores indicate risk and protection, respectively. *MAF* minor allele frequency, *n* number of informative families

provide a mechanistic explanation for the association of *UBASH3A* with autoimmunity, as interleukin-2 (IL-2) is a crucial cytokine for the differentiation and function of T cells.

In the current study, we examine the role of another noncoding SNP in *UBASH3A*, rs1893592 (GRCh37 chr21: g.43855067A>C). The minor allele of rs1893592 protects against three autoimmune diseases, rheumatoid arthritis, celiac disease, and primary sclerosing cholangitis [8, 10, 11, 13], and it has been proposed to affect splicing, rather than expression, of *UBASH3A* [13, 14]. Our study reconciles these disparate mechanistic roles for autoimmune-associated SNPs in *UBASH3A* and highlights the combinatorial effects of noncoding *UBASH3A* variants on T1D risk.

Materials and methods

Frozen peripheral blood mononuclear cells from healthy subjects of European ancestry were obtained from the Type 1 Diabetes Genetics Consortium and from STEMCELL Technologies (Vancouver, BC, Canada). Genotype and phenotype data used in this study are available in dbGaP (www.ncbi.nlm.nih.gov/gap; Study Accession: phs000911.v1.p1). All biospecimens and data were represented by only non-identifying codes. This study was approved by the University of Florida Institutional Review Board.

Quantitative PCR and isolation and stimulation of human primary $CD4^+$T cells were performed as previously described [12]. Primers used for quantitative PCR are provided in Supplementary Table 1.

IL-2 in culture supernatants was measured using the LEGENDplex kit (BioLegend, San Diego, CA, USA), following the manufacturer's protocol.

The Family-Based Association Test program (v2.0.3) was used for single-marker association tests and haplotype analyses [15]. Conditional analysis was conducted using the UNPHASED program (v3.1.7). The Prism software (Graphpad v7.0) was used to perform Student's *t*-tests and linear regression analysis.

Results

The minor allele of rs1893592 protects against T1D

To examine the effect of rs1893592 on risk of T1D, we conducted single-marker family-based association tests using data from 10,796 individuals from 2732 T1D-affected sibling pairs and trio families of European ancestry. The minor allele of rs1893592 was associated with reduced risk for T1D ($P = 2.6 \times 10^{-5}$, Table 1). rs11203202 (GRCh37 chr21:g.43825357C>G) was previously identified as the index SNP for the 21q22.3 risk locus in a large-scale fine-mapping study for T1D [2]. As expected, the minor alleles of rs11203202 ($P = 2.7 \times 10^{-4}$), rs80054410 ($P = 4.3 \times 10^{-4}$), and rs11203203 ($P = 4.0 \times 10^{-4}$) were each found to be associated with increased risk for T1D (Table 1) [2, 12].

To explore the relationship of rs11203202, rs80054410, rs11203203, and rs1893592, we performed family-based haplotype association analysis and identified three T1D-associated haplotypes—H2, H3, and H9 (Table 2). The common H3 haplotype carrying the minor, protective allele of rs1893592 and the major alleles of rs11203202, rs80054410, and rs11203203 were associated with reduced risk for T1D ($P = 3.8 \times 10^{-6}$). The common H2 haplotype carrying the major allele of rs1893592 and the minor, risk alleles of rs11203202, rs80054410, and rs11203203 conferred risk for T1D ($P = 0.0018$). The rare H9 haplotype (frequency = 0.1%) carrying the minor alleles of all four SNPs also conferred risk for T1D ($P = 0.012$, Table 2). Even after conditioning on rs80054410 and rs11203203, rs1893592 exhibited a significant association with T1D ($P = 0.026$).

rs1893592 modulates the expression of *UBASH3A* and *IL2* in human $CD4^+$ T cells upon stimulation

rs1893592 is located at the +3 position of intron-9 of *UBASH3A* (NM_001001895.2 and NG_029750.1), and could affect splicing of *UBASH3A*. Our prior RNA-seq analysis of human primary T cells uncovered an alternative *UBASH3A* transcript that runs into intron-9 of *UBASH3A*

Table 2 Family-based haplotype association analysis of *UBASH3A*

Haplotype	rs11203202 (C>G)	rs80054410 (T>C)	rs11203203 (G>A)	rs1893592 (A>C)	Frequency	Family (*n*)	*Z* score	*P*-value
H1	C	T	G	A	0.362	1337	−0.35	0.73
H2	G	C	A	A	0.292	1264	3.12	0.0018
H3	C	T	G	C	0.175	958	−4.62	3.76E-06
H4	C	C	A	A	0.087	571	1.29	0.20
H5	G	T	G	C	0.055	405	0.21	0.83
H6	C	C	A	C	0.015	134	−1.04	0.30
H7	G	T	G	A	0.009	136	1.55	0.12
H8	C	C	G	C	0.005	37	0.10	0.92
H9	G	C	A	C	0.001	71	2.53	0.0116
H10	C	C	G	A	<0.001	2	−1.00	0.32
H11	G	C	G	A	<0.001	1	0	1.00

Positive and negative *Z* scores indicate risk and protection, respectively. *n* number of informative families

and contains a premature stop codon [16]. We verified this transcript (ENST00000473381.1, hereafter referred to as intron-9 *UBASH3A*) by RT-PCR and Sanger sequencing.

We performed quantitative PCR to further characterize intron-9 *UBASH3A*. Our results revealed that human primary $CD4^+$ T cells, regardless of their genotype at rs1893592, expressed intron-9 *UBASH3A*, albeit at low levels (Fig. 1a). In addition, the amount of intron-9 *UBASH3A* mRNA was positively correlated with that of total *UBASH3A* mRNA in both stimulated and mock-stimulated $CD4^+$ T cells (Fig. 1a).

To assess the effect of rs1893592, we controlled for genotypes at rs80054410 and rs11203203 by selecting subjects homozygous for the major alleles of these two SNPs. We found that in stimulated $CD4^+$ T cells, the minor, protective C allele of rs1893592 was associated with a significant reduction in the total amount of *UBASH3A* mRNA but not the amount of intron-9 *UBASH3A* transcript (Fig. 1b). As a result, the proportion of intron-9 *UBASH3A* mRNA relative to total *UBASH3A* mRNA was twice as large in rs1893592 A/C subjects as in rs1893592 A/A subjects (Fig. 1c). In addition, $CD4^+$ T cells from rs1893592 A/C subjects, upon stimulation, produced more *IL2* transcripts than those from rs1893592 A/A subjects (Fig. 1d).

To further test the hypothesis that the minor allele of rs1893592 protects against T1D by enhancing IL-2 production, we measured IL-2 secretion by primary $CD4^+$ T cells from healthy subjects upon 24-h stimulation with anti-CD3 plus anti-CD28, without controlling for genotypes at other T1D-associated SNPs in *UBASH3A* or other risk loci. We found that $CD4^+$ T cells from rs1893592 A/C subjects secreted five times more IL-2 than rs1893592 A/A subjects (Fig. 1e).

Discussion

In this study, we sought to define how the minor allele of the intronic SNP rs1893592 in *UBASH3A* protects against T1D in the presence of risk alleles at other SNPs in the same gene. The results of our haplotype analysis suggest an additive model, at least at the population level, where T1D risk results from the sum of the individual effects of the four noncoding *UBASH3A* SNPs. This is consistent with the notion—predicted by quantitative genetics theory and supported by empirical data—that most genetic variance in a population is additive [17]. Furthermore, this model is supported by our mechanistic findings that these *UBASH3A* SNPs affect, albeit in different directions, the same molecular pathway—the expression of *UBASH3A*, and thereby *IL2* expression via the NF-κB signaling pathway [12].

UBASH3A is primarily expressed in T cells with little or no expression in other cell types, including B cells. Previous studies that assayed *UBASH3A* mRNA in human whole blood or B-lymphoblastoid cell lines concluded that rs1893592 was a splice QTL whose minor allele promoted the retention of intron-9 of *UBASH3A* [13, 14]. In contrast, Raj et al. [18] identified rs1893592 as the most significant cis-eQTL in *UBASH3A* in primary $CD4^+$ T cells from European Americans, and the minor allele was found to be associated with decreased expression of *UBASH3A*. Our results in stimulated human $CD4^+$ T cells both reconcile and extend these prior findings in several important ways. First, our study reveals that intron-9 *UBASH3A* is a normally produced transcript. Second, the impact of rs1893592 is two-fold: its minor allele reduces overall transcript levels of *UBASH3A*, while increasing alternative splicing, favoring the production of intron-9 *UBASH3A* (Fig. 1c). Third,

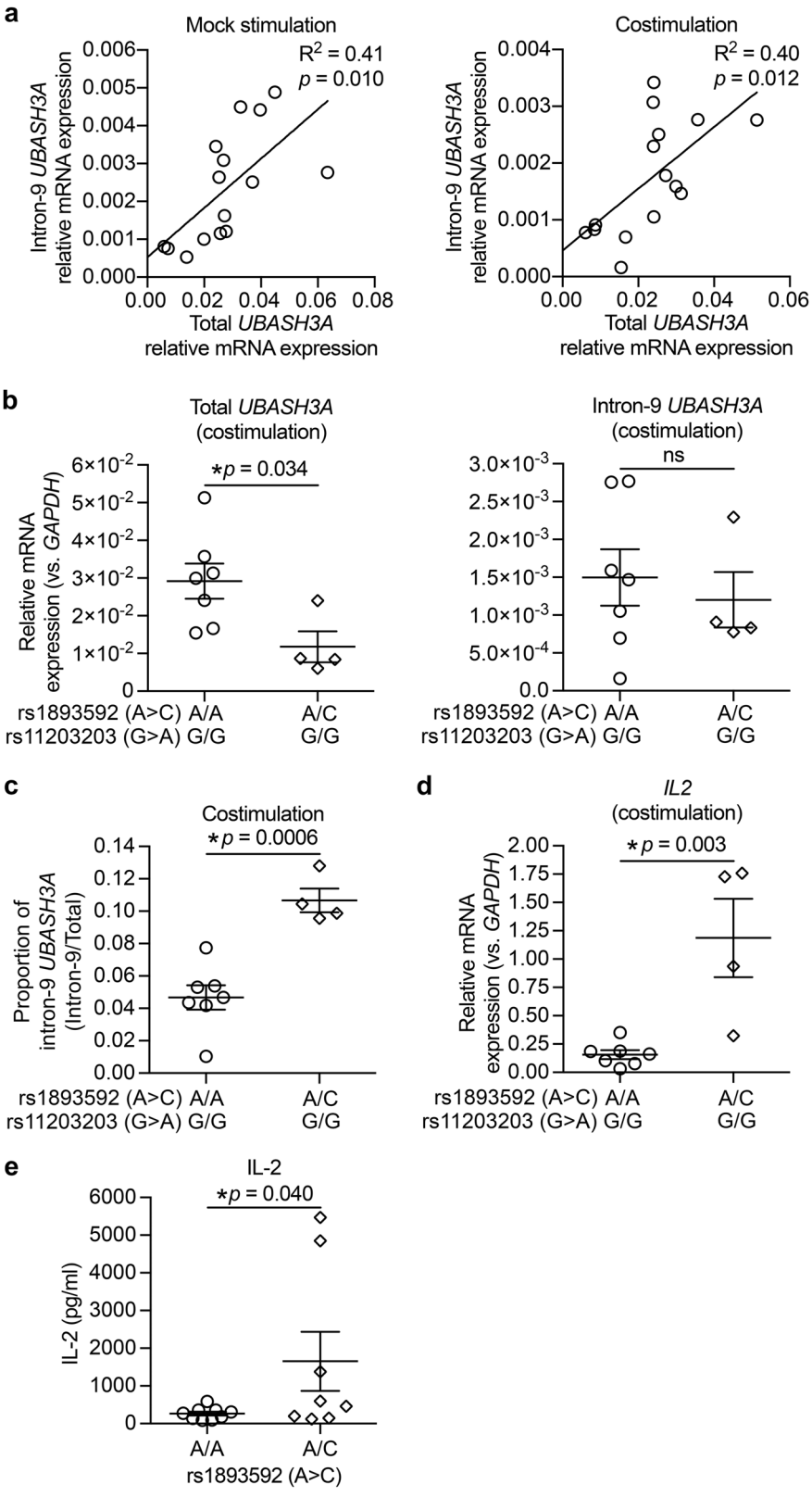

Fig. 1 Effects of rs1893592 on expression of *UBASH3A* and *IL2* in human primary CD4$^+$ T cells upon stimulation. **a–d** Human primary CD4$^+$ T cells from healthy subjects of European ancestry were stimulated with anti-CD3 plus anti-CD28, or with culture medium, for 6 h. Each data point represents one subject, and the data are pooled from four quantitative PCR experiments. In **a**, the best-fit line generated by linear regression is represented by solid line, and statistics for correlation and significance are shown. In **b–d**, all subjects are homozygous for the major alleles of rs80054410 (A) and rs11203203 (G). The mean and SEM values are indicated by solid lines and error bars, respectively. *P*-values from unpaired two-tailed Student's *t*-tests are shown. In **b**, the relative mRNA levels of total *UBASH3A* and intron-9 *UBASH3A* are shown. ns, not significant. In **c**, the ratio of intron-9 *UBASH3A* mRNA to total *UBASH3A* mRNA is shown. In **d**, the relative mRNA level of *IL2* is shown. **e** Human primary CD4$^+$ T cells from healthy subjects of European ancestry were stimulated with anti-CD3 plus anti-CD28 for 24 h. IL-2 concentrations in the culture supernatants are shown, and the data are pooled from two LEGENDplex assays. Each data point represents one subject, and *P*-value from unpaired one-tailed Student's *t*-test is shown

since rs1893592 does not significantly alter the absolute levels of intron-9 *UBASH3A*, the most likely mechanism whereby variation at rs1893592 affects T1D risk is via its effects on overall *UBASH3A* mRNA levels rather than through the production of a novel protein isoform from the intron-9 *UBASH3A* transcript.

We showed that a significant outcome of reduced total amount of *UBASH3A* mRNA, resulting from the minor allele of rs1893592, is enhanced transcription and production of IL-2 in stimulated $CD4^+$ T cells. This explains, at least partially, the protective effect of the minor allele of rs1893592, as IL-2 production is impaired in some subjects with T1D or rheumatoid arthritis [19, 20]. We noticed that the amounts of secreted IL-2 by $CD4^+$ T cells varied markedly among rs1893592 A/C subjects (Fig. 1e). This is likely due to the facts that we grouped the subjects based only on their genotype at rs1893592, and that IL-2 production can be affected by variants in genes other than *UBASH3A*, such as those in *IL2RA* and *PTPN2*, two T1D-associated genes involved in the IL-2 pathway [2]. Nevertheless, our results show that rs1893592 has a considerable impact on IL-2 production, which may facilitate future IL-2 clinical trials for T1D.

Acknowledgements This research utilizes resources from the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF). This study was supported by NIH grants DK106718 and DK085678 (to PC), and by the JDRF Advanced Postdoctoral Fellowship award 3-APF-2016-177-A-N (to YG).

Funding: NIH grants DK106718 and DK085678 (to PC), and the JDRF Advanced Postdoctoral Fellowship award 3-APF-2016-177-A-N (to YG).

Compliance with ethical standards

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

References

1. Barrett JC, Clayton DG, Concannon P, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009;41:703–7.
2. Onengut-Gumuscu S, Chen W-M, Burren O, et al. Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. Nat Genet. 2015;47:381–6.
3. Concannon P, Onengut-Gumuscu S, Todd JA, et al. A human type 1 diabetes susceptibility locus maps to chromosome 21q22.3. Diabetes. 2008;57:2858–61.
4. Grant SFA, Qu H-Q, Bradfield JP, et al. Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes. Diabetes. 2009;58:290–5.
5. Johnson K, Wong R, Barriga KJ, et al. rs11203203 is associated with type 1 diabetes risk in population pre-screened for high-risk HLA-DR,DQ genotypes. Pediatr Diabetes. 2012;13:611–5.
6. Jin Y, Birlea SA, Fain PR, et al. Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. N Engl J Med. 2010;362:1686–97.
7. Zhernakova A, Stahl EA, Trynka G, et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. PLoS Genet. 2011;7:e1002004.
8. Trynka G, Hunt KA, Bockett NA, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet. 2011;43:1193–201.
9. Diaz-Gallo L-M, Sánchez E, Ortego-Centeno N, et al. Evidence of new risk genetic factor to systemic lupus erythematosus: the *UBASH3A* gene. PLoS One. 2013;8:e60646.
10. Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2014;506:376–81.
11. Kim K, Bang S-Y, Lee H-S. et al. High-density genotyping of immune loci in Koreans and Europeans identifies eight new rheumatoid arthritis risk loci. Ann Rheum Dis. 2015;74:e13
12. Ge Y, Paisie TK, Newman JRB, McIntyre LM, Concannon P. UBASH3A mediates risk for type 1 diabetes through inhibition of T-cell receptor-induced NF-κB signaling. Diabetes. 2017;66:2033–43.
13. Ji S-G, Juran BD, Mucha S, et al. Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease. Nat Genet. 2017;49:269–73.
14. Battle A, Mostafavi S, Zhu X, et al. Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. Genome Res. 2014;24:14–24.
15. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol. 2000;19: S36–S42.
16. Newman JRB, Conesa A, Mika M, et al. Disease-specific biases in alternative splicing and tissue-specific dysregulation revealed by multitissue profiling of lymphocyte gene expression in type 1 diabetes. Genome Res. 2017;27:1807–15.
17. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. PLoS Genet. 2008;4:e1000008.
18. Raj T, Rothamel K, Mostafavi S, et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. Science. 2014;344:519–23.
19. Roncarolo MG, Zoppo M, Bacchetta R, et al. Interleukin-2 production and interleukin-2 receptor expression in children with newly diagnosed diabetes. Clin Immunol Immunopathol. 1988;49:53–62.
20. Combe B, Pope RM, Fischbach M, Darnell B, Baron S, Talal N. Interleukin-2 in rheumatoid arthritis: production of and response to interleukin-2 in rheumatoid synovial fluid, synovial tissue and peripheral blood. Clin Exp Immunol. 1985;59:520–8.