

Identification of a novel proinsulin-associated SNP and demonstration that proinsulin is unlikely to be a causal factor in subclinical vascular remodelling using Mendelian randomisation

Strawbridge Proinsulin effects on carotid IMT

Rona J Strawbridge^{1,a,b}, Angela Silveira^a, Marcel den Hoed^c, Stefan Gustafsson^d, Jian'an Luan^e, Denis Rybin^f, Josée Dupuis^{g,h}, Ruifang Li-Gaoⁱ, Maryam Kavousi^j, Abbas Dehghan^{j,k}, Kadri Haljas^l, Jari Lahti^{l,m}, Jesper R. Gådin^a, Alexandra Bäcklund^a, Ulf de Faire^{n,o}, Karl Gertow^a, Phillipe Giral^p, Anuj Goel^{q,r}, Steve E Humphries^s, Sudhir Kurl^t, Claudia Langenberg^e, Lars L Lannfelt^u, Lars Lind^{e,v}, Cecilia CM Lindgren^q, Elmo Mannarino^w, Dennis O Mook-Kanamori^{i,x}, Andrew P. Morris^{q,y}, Renée de Mutsertⁱ, Rainer Rauramaa^{z,a1}, Peter Saliba-Gustafsson^a, Bengt Sennblad^{a,b1}, Andries J Smit^{c1,d1}, Ann-Christine Syvänen^{e1}, Elena Tremoli^{f1,g1}, Fabrizio Veglia^{g1}, Björn Zethelius^{h1}, Hanna M. Björck^a, Johan G Eriksson^{i1,j1,k1}, Albert Hofman^{j1,l1}, Oscar H Francoⁱ, Hugh Watkins^{q,r}, J Wouter Jukema^{m1,n1}, Jose C Florez^{o1,p1,q1}, Nicholas J Wareham^e, James B Meigs^{p1,q1,r1}, Erik Ingelsson^{d,s1}, Damiano Baldassarre^{f1,g1} and Anders Hamsten^a on behalf of the IMPROVE study group*.

Affiliations

^a Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden

^b Mental Health and Wellbeing, Institute of Mental Health and Wellbeing, University of Glasgow, Glasgow, UK

^c Department of Immunology, Genetics and Pathology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

^d Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

^e MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK

^f Boston University Data Coordinating Center, Boston, MA, USA

^g Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

^h National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA

ⁱ Department of Epidemiology, Leiden University Medical Center, Leiden, the Netherlands

^j Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, The Netherlands

^k MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK

^l Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland

^m Folkhälsan Research Centre, Helsinki, Finland

ⁿ Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

^o Department of Cardiology, Karolinska University Hospital, Stockholm, Sweden

^p Assistance Publique–Hôpitaux de Paris, Service Endocrinologie-Métabolisme, Groupe Hospitalier Pitié-Salpêtrière, Unités de Prévention Cardiovasculaire, Paris, France

^q Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK

^r Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK

^s Centre for Cardiovascular Genetics, Institute Cardiovascular Science, University College London, UK.

^t Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio Campus, Kuopio, Finland

^u Department of Public Health and Caring Sciences, Geriatrics; Molecular Geriatrics, Uppsala University, Uppsala, Sweden

^v Department of Medical Sciences, Uppsala University, Uppsala, Sweden

^w Department of Clinical and Experimental Medicine, Internal Medicine, Angiology and Arteriosclerosis Diseases, University of Perugia, Perugia, Italy

^x Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands

^y Department of Biostatistics, University of Liverpool, Liverpool, UK

^z Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland

^{a1} Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland

^{b1} Dept of Cell and Molecular Biology, National Bioinformatics Infrastructure Sweden,
Science for Life Laboratory, Uppsala University

^{c1} Department of Medicine, University Medical Center Groningen, Groningen, the
Netherlands

^{d1} Department of Medicine, University Medical Center Groningen, University of
Groningen, Groningen, the Netherlands

^{e1} Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory,
Uppsala University, Uppsala, Sweden

^{f1} Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano, Milan,
Italy

^{g1} Centro Cardiologico Monzino, IRCCS, Milan, Italy

^{h1} Department of Public Health and Caring Sciences/Geriatrics, University Hospital,
Uppsala

ⁱ¹ Department of General Practice and Primary Health Care, University of Helsinki and
Helsinki University Hospital, Helsinki, Finland

^{j1} National Institute for Health and Welfare, Finland

^{k1} Folkhälsan Research Center, Helsinki, Finland

^{l1} Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA,
USA

^{m1} Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands

ⁿ¹ Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands

^{o1} Diabetes Unit and Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA

^{p1} Programs in Metabolism and Medical & Population Genetics, Broad Institute, Cambridge, MA, USA

^{q1} Department of Medicine, Harvard Medical School, Cambridge, MA, USA

^{r1} General Medicine Division, Massachusetts General Hospital, Boston, MA, USA

^{s1} Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA

*Additional members of the IMPROVE study group are listed in the Online Data Supplement

Corresponding author:

Dr Rona J. Strawbridge

L8:03, Centre for Molecular Medicine

Karolinska Universitetssjukhuset Solna

Stockholm, Sweden, 171 76

Email: Rona.Strawbridge@ki.se

¹Present address:

Department of Health and Wellbeing

University of Glasgow

1 Lilybank Gardens, G12 8RZ

Telephone: 00 44 (0)141 3303482

Email: Rona.Strawbridge@glasgow.ac.uk

Word count: 5913

Keywords: Proinsulin, atherosclerosis, intima-media-thickness, single nucleotide polymorphisms, genetic variants, Mendelian Randomisation

Abstract (248/250)

Background and aims: Increased proinsulin relative to insulin levels have been associated with subclinical atherosclerosis (measured by carotid intima-media thickness (cIMT)) and are predictive of future cardiovascular disease (CVD), independently of established risk factors. The mechanisms linking proinsulin to atherosclerosis and CVD are unclear. A genome-wide meta-analysis has identified nine loci associated with circulating proinsulin levels. Using proinsulin-associated SNPs, we set out to use a Mendelian randomisation approach to test the hypothesis that proinsulin plays a causal role in subclinical vascular remodelling.

Methods: We studied the high CVD-risk IMPROVE cohort (n=3,345), which has detailed biochemical phenotyping and repeated, state-of-the-art, high-resolution carotid ultrasound examinations. Genotyping was performed using Illumina Cardio-Metabo and Immuno arrays, which include reported proinsulin-associated loci. Participants with type 2 diabetes (n=904) were omitted from the analysis. Linear regression was used to identify proinsulin-associated genetic variants.

Results: We identified a proinsulin locus on chromosome 15 (rs8029765) and replicated it in data from 20,003 additional individuals. An 11-SNP score, including the previously identified and the chromosome 15 proinsulin-associated loci, was significantly and negatively associated with baseline IMT_{mean} and IMT_{max} (the primary cIMT phenotypes) but not with progression measures. However, MR-Eggers refuted any significant effect of the proinsulin-associated 11-SNP score, and a non-pleiotropic SNP score of three variants (including rs8029765) demonstrated no effect on baseline or progression cIMT measures.

Conclusions: We identified a novel proinsulin-associated locus and demonstrated that whilst proinsulin levels are associated with cIMT measures, proinsulin *per se* is unlikely to have a causative effect on cIMT.

Introduction

Hyperinsulinemia, most often a result of systemic insulin resistance, is associated with increased risk of cardiovascular disease (CVD) and severity of coronary artery disease (CAD) ¹. Under normo-glycemic conditions, proinsulin is fully processed to insulin and C-peptide before secretion, with very low concentrations of proinsulin being detectable in the blood. During hyperinsulinemia, increased demand for insulin production results in pancreatic β -cells failing to correctly process proinsulin to mature insulin ². Thus, circulating levels of proinsulin are increased (relative to insulin levels) and this might be a link between insulin resistance and CVD. **In support of this hypothesis, anti-diabetic agents which modulate glucose levels without increasing insulin secretion (for example SGLT2-inhibitors) have shown cardiovascular benefits, whilst those which stimulate insulin (and potentially proinsulin) secretion, such as sulfonylureas, have shown no CVD benefits and in some studies, increased CVD risk ³.** Furthermore, subclinical changes in the wall of the carotid artery, measured by carotid intima-media thickness (cIMT), have been shown to correlate positively with proinsulin levels in healthy individuals ⁴ and increased proinsulin levels have been demonstrated to predict future CVD, independently of established risk factors ⁵⁻⁹. However, it remains to be determined whether increased proinsulin levels merely reflect CVD risk-related processes, or whether proinsulin has a direct adverse functional effect on the vasculature.

~~Identification of genetic variants (single nucleotide polymorphisms (SNPs)) convincingly associated with circulating proinsulin levels ¹⁰ has paved the way for Mendelian randomisation (MR) approaches to assess causality of proinsulin on CVD phenotypes.~~

~~The basis of the MR approaches is that if proinsulin has a causal effect on a cardiovascular phenotype, then one would expect that proinsulin-associated genetic variants would demonstrate consistent associations with cIMT measures. A gene score of proinsulin-associated loci~~ Genetic predisposition to increased proinsulin levels ¹⁰

previously demonstrated no effect on CVD events, however we proposed that increased proinsulin levels could be important in earlier stages of CVD development, such as those reflected by cIMT.

The aims of this project were firstly to explore the existence of additional proinsulin-associated variants using a large panel of candidate loci for metabolic or inflammatory processes and secondly to use MR to assess causality of proinsulin on early-stage atherosclerosis, represented by cIMT. The IMPROVE cohort is uniquely positioned to address the second aim, having extensive genotyping, circulating proinsulin and insulin levels and detailed cIMT measurements in 2,441 participants without type 2 diabetes (T2D). ~~Accordingly, we set out to test the hypothesis that proinsulin exerts a causal detrimental effect on subclinical vascular processes reflected by cIMT.~~ Demonstration that proinsulin has a causal role in cIMT would highlight insulin resistance as a target for CVD prevention. As insulin sensitisation by dietary, lifestyle and pharmaceutical interventions is possible, the role of proinsulin in CVD warrants further investigation.

Materials and Methods

Cohort description: IMPROVE

The IMPROVE study **recruitment and ultrasound protocol have** been described in detail ^{11, 12}. In short, IMPROVE consists of participants without symptoms or history of CVD at enrolment, who presented with at least three classical risk factors for CVD (**including family history of CVD, diabetes or impaired fasting glucose, hypoalphalipoproteinemia, hypertension, hypertriglyceridemia, hypercholesterolemia, current smoking** ^{11, 12}). Between March 2004 and April 2005, 3,711 participants were recruited from seven centres in five European countries (Finland, Sweden, the Netherlands, France and Italy). Extensive ultrasound examinations of cIMT were carried out at baseline and after 15 and 30 months, with linear regression of cIMT over time being used to calculate cIMT progression variables ^{11, 12}. Primary cIMT phenotypes used in the present study were the baseline mean and maximum of the common carotid artery intima-media thickness (IMT_{mean} and IMT_{max}, respectively) ^{11, 12}, as only these baseline measures would be available in replication cohorts. Secondary analyses were conducted on progression of IMT_{mean} and IMT_{max} as well as baseline and progression measures of additional segments (Online Data Supplement) ^{11, 12}. At baseline, a structured medical history and lifestyle questionnaire was completed and blood was sampled. Standard clinical and biochemical phenotyping was performed. Participants with T2D (n=904, defined as having been diagnosed or treated for diabetes, or fasting glucose ≥ 7 mmol/L) were excluded from all analyses. Fasting intact plasma proinsulin was measured with a sandwich enzyme-linked immunosorbent assay (DRG Instruments GmbH, Marburg,

Germany, Online Data Supplement). Fasting plasma insulin was analysed by electrochemiluminescence immunoassay (Meso Scale Discovery, Gaithersburg, MD, USA, Online Data Supplement). Cohort characteristics are described in Supplemental Table 1.

Genotyping in IMPROVE

IMPROVE has been genotyped using both the Illumina CardioMetabo 200k ¹³ and Immunochip 200k ¹⁴ bead array platforms. **Single nucleotide polymorphisms (SNPs)** were excluded for low call rate (<95%), low minor allele frequency (<0.01) and for deviation from Hardy-Weinberg equilibrium ($p < 1 \times 10^{-6}$). Participants were excluded due to low call rate (<95%), aberrant sex assignment or cryptic relatedness (IBD >0.2). Multi-dimensional scaling (MDS) components 1-3 were calculated in PLINK ¹⁵ to enable adjustment for population structure. After quality control 2,441 participants without T2D and 254,756 SNPs with minor allele frequency >1% were included in the analysis.

Identification of proinsulin-associated SNPs in IMPROVE

To identify proinsulin-associated variants ¹⁰, 254,756 SNPs were analysed for influence on proinsulin levels. This also enabled assessment of the reported ¹⁰ proinsulin-associated loci. The reported lead ¹⁰, the analysed proxy or the strongest SNP for each locus are presented in Supplemental Table 2.

Levels of proinsulin, insulin and glucose were natural log normalised prior to analysis. Linear regression, assuming an additive genetic model, was used to analyse the effect of **single nucleotide polymorphisms (SNPs)** on proinsulin levels. To enable comparison

of SNP effects with the proinsulin **genome-wide association study** (GWAS) ¹⁰, the same three models explored in that paper were tested: insulin model (age, sex, population structure and insulin), glucose model (age, sex, population structure, insulin and glucose) and **body-mass index (BMI)** model (age, sex, population structure, insulin and BMI). Loci with $p < 1 \times 10^{-5}$ (suggestive evidence of association) in any model were selected for replication. Of note, IMPROVE was not part of the prior proinsulin GWAS ¹⁰.

Replication of proinsulin-associated SNPs

Replication of proinsulin-associated loci was attempted in seven independent cohorts (Fenland (<http://www.mrc-epid.cam.ac.uk/Research/Studies/Fenland>), FHS ¹⁶⁻¹⁸, HBCS ¹⁹, PIVUS ^{20, 21}, ULSAM ^{22, 23}, PROCARDIS cases ²⁴ and PROCARDIS controls ²⁴, $n=18,773-20,003$ participants). Replication cohorts are described in the Online Data Supplement and cohort characteristics are presented in Supplemental Table 3. The same analyses and models as the discovery analysis were applied. To improve power and examine consistency of effects, replication results were combined in an inverse variance meta-analysis, using GWAMA²⁵.

Expression quantitative trait locus analysis

To further characterize the proinsulin-associated locus on chromosome 15 (rs8029765), genotype-specific gene expression patterns were assessed. This was conducted in liver biopsies from participants undergoing elective surgery for aortic valve and/or ascending aortic disease in ASAP ²⁶ (Online Data Supplement) and biopsies from liver or heart and peripheral blood mononuclear cells from organ donation or autopsy specimens in the GTEx ²⁷ database (Online Data Supplement).

In silico analysis of effect on T2D

Proinsulin-increasing alleles of previously identified SNPs rather surprisingly demonstrated a mixture of positive, negative and null effects on T2D risk¹⁰. Thus to assess the impact of the novel proinsulin-associated locus on chromosome 15 (rs8029765) on risk of T2D, an *in silico* lookup was performed in the DIAGRAM consortium data. This study of primarily Europeans included 34,840 T2D cases and 114,981 controls with GWAS or metabochip data and the results were downloaded from the DIAGRAM consortium website (<http://diagram-consortium.org/downloads.html>).

Effects of proinsulin on cIMT measures in IMPROVE

All phenotypes were assessed for normality and natural logarithmic transformation was applied where necessary, prior to further statistical analysis. The analysis of proinsulin levels for effect on cIMT measures were adjusted in basic (age, sex and population structure) and extended models (age, sex, population structure, BMI, **systolic blood pressure (SBP)**, **high density lipoprotein (HDL)**, **triglycerides (TGs)** and current smoking)²⁸. As the mean cIMT measure encompasses the maximum cIMT measure, these variables were not considered independent and therefore related, multiple testing was not corrected for and $p < 0.05$ was considered significant. Secondary analyses explored the consistency of effects of proinsulin on other cIMT segments and on analysis of progression measures, where the baseline measure of the same segment was also included as a covariate. Because the secondary analyses were exploratory only, the size and direction of effects, rather than significance, is the focus of these

analyses. Analyses were performed in STATA (STATAcorp LP, College Station, Texas, USA).

Mendelian Randomization (MR) to assess causality of proinsulin on cIMT in IMPROVE

In order to be valid instruments for the MR analyses, the proinsulin-associated loci should demonstrate the expected associations with proinsulin levels. Accordingly, the effect size and direction were examined for all proinsulin-associated loci ¹⁰. Previously published results were considered the gold standard (given the much larger sample size in the GWAS (n~20,000) compared with this study) and were thus referred to in further analyses. Only subjects with complete genotyping for the 11 proinsulin-associated SNPs (10 SNPs in 9 reported loci and the chromosome 15 SNP/locus identified here) were included in this analysis (N=35 excluded).

As individual SNPs had only small effects on proinsulin levels, we calculated proinsulin-associated SNP scores by summing the number of alleles associated with higher proinsulin levels. SNP scores were calculated unweighted (sum of number of proinsulin-increasing alleles) and weighted (sum of the number of alleles multiplied by the effect size of the allele reported in the GWAS ¹⁰). As there was little difference between the variance of proinsulin explained by the unweighted and weighted SNP scores, for simplicity the unweighted scores were used for further analysis. The majority of proinsulin-associated SNPs have also been associated with other metabolic traits of relevance to CVD. This pleiotropy may introduce confounding or explain some of the proinsulin associations with cIMT. To determine whether this was the case, a further

SNP score was constructed, including only the SNPs which had no reported metabolic effects (GWAS catalogue, <https://www.ebi.ac.uk/gwas/>, Online Data Supplement). Analyses were adjusted using the basic model (age, sex and population structure). Instrument variable regression, with the proinsulin SNP scores as the instruments, was used to assess causality of proinsulin on cIMT variables and was conducted in STATA (STATAcorp LP, College Station, Texas, USA). Instrument variable analysis assumes that SNPs have no pleiotropic effects. Therefore, Egger's regression (which can be used to assess associations which include pleiotropic SNPs) was also conducted in R (Mendelian Randomization package).

Replication of MR

Replication of the MR experiment was attempted in 12,113 participants from 5 cohorts (2 of which contributed to the SNP-proinsulin replication analysis) with IMT_{mean} and IMT_{max} phenotypes. The replication cohorts are described in the Online Data Supplement and characteristics are presented in Supplemental Table 4. The cIMT measures were broadly comparable with IMPROVE, with multiple measurements of the carotid artery enabling assessment of both the maximum and mean values, both left and right carotid arteries being measured (except in ULSAM), and the 1st cm proximal to the bifurcation being omitted from common carotid artery measurements (except in PIVUS). Meta-analysis of SNP score associations with IMT measures was conducted in R using the `rma` function in the METAFOR package (<https://cran.r-project.org/web/packages/metafor/index.html>)

Results

Identification of proinsulin-associated SNPs

~~In order to identify proinsulin-associated loci, The combined chip was analysed for association with proinsulin levels and adjusted for relevant confounders (consistent with the proinsulin GWAS, namely population structure, age, sex, insulin as well as BMI or glucose¹⁰).~~

Analysis of the combined chip (adjusting for population structure, age, sex, insulin as well as BMI or glucose¹⁰) identified ~~Here,~~ a total of 16 SNPs ~~distributed across~~ in 3 loci reaching the suggestive level of significance for association with proinsulin ($p < 1 \times 10^{-5}$) (Figure 1 and Supplemental Tables 5 and 6). ~~Consistent with the published proinsulin GWAS¹⁰, more proinsulin-associated SNPs were identified using the glucose-adjusted model.~~ Most (n=13) of these SNPs were located in the *ARAP1* locus. This locus was identified previously¹⁰, so this finding was taken to be a positive control and gave us confidence in the other loci identified. Conditional analysis for the two known signals (rs77756118 and rs11603334) in the *ARAP1* locus rendered these SNPs non-significant (Figure 2). ~~Two signals were observed:~~ On chromosome 6, rs2323000 and rs6910151 represent the same signal. On chromosome 15 (rs8029765), the signal is located 29Mb from the *LARP6* and 20Mb from the *VPS13C* loci. Conditioning on *LARP6* (strongest SNP) and *VPS13C* (reported lead SNP) signals had negligible effects on the chromosome 15 signal (Figure 2 and Supplemental Table 7), which thus appeared to be a distinct signal. The reason that this SNP was not identified in the previous GWAS meta-analysis was most likely its relatively large effect size in IMPROVE (Beta=0.09),

compared with that in the GWAS discovery analysis ($\text{Beta}=0.01$)¹⁰. A stronger effect size in the IMPROVE study compared with the GWAS is plausible: We believe proinsulin levels are important in CVD, and the IMPROVE cohort is enriched for CVD risk factors. In contrast, the GWAS included a much wider spectrum of CVD burden. Hence two proinsulin-associated SNPs, rs2323000 and rs8029765, were taken forward to replication.

Replication of proinsulin-associated SNPs

Replication ~~results~~ of the two proinsulin-associated SNPs ~~was attempted in~~ from seven independent cohorts ~~and results~~ were meta-analysed ~~across all studies~~ ($n=18,773$ - $20,004$). The meta-analysis refuted the chromosome 6 locus association with proinsulin levels (Table 1). However, nominal significance was observed for the minor allele of rs8029765 (chromosome 15) being associated with higher proinsulin levels in most models (insulin model: $\text{Beta}\pm\text{Se } 0.013\pm0.006$, $p=0.04$ (combined discovery and replication 0.020 ± 0.006 , $p=0.001$); BMI model: 0.014 ± 0.006 , $p=0.03$ (combined discovery and replication 0.020 ± 0.006 , $p=0.002$)). Additionally adjusting for glucose did not materially alter the effect size (0.012 ± 0.006 , $p=0.0709$ (combined discovery and replication 0.020 ± 0.006 , $p=0.003$)). The replication meta-analyses indicated that the effect sizes in women were not significantly different from those in men ($p=0.0659$), with no evidence of heterogeneity between the strata ($I^2=0$, sex heterogeneity $p=0.80$). These data provide evidence for a novel proinsulin-associated locus on chromosome 15, which did not influence risk of T2D (OR 1.00, confidence interval (CI) 0.95-1.06, $p=0.89$, $n=51,411$)²⁹. rs8029765 (and the SNP in high LD, rs4450375, $R^2>0.8$,

Supplemental Figure 1) has not previously been reported to be associated with any of the metabolic or cardiovascular traits we studied.

Expression QTL for the chromosome 15 locus

The relevance of the chromosome 15 locus was assessed by examining nearby genes for genotype-specific expression. In the GTEx heart (n=190), liver (n=97) or PBMC (n=338) data, generated from relatively healthy individuals, no significant eQTLs were identified. However, in the ASAP cohort, which is significantly burdened with cardiovascular disease, levels of expression were detectable for eight genes in the chromosome 15 locus: *MAN2A2*, *HDDC3*, *UNC45A*, *RCCD1*, *PRC1*, *VPS33B*, *SV2B*, *SLCO3A1*. In liver samples (n=209), rs8029765 was significantly associated with expression of *UNC45A* (Supplemental Figure 2, $p=5.6 \times 10^{-3}$) with the proinsulin-increasing allele demonstrating lower expression of *UNC45A*. The unc-45 myosin chaperone A encoded by the *UNC45A* gene is a widely expressed chaperone protein ³⁰.

Effects of Proinsulin on IMT_{mean} and IMT_{max} .

As anticipated, proinsulin levels demonstrated correlations with almost all established risk variables for CVD in the IMPROVE cohort (Supplemental Table 8). ~~The effect of proinsulin levels on measures of cIMT was examined.~~ For further analysis of IMT measures, models were adjusted for relevant confounders ²⁸. No statistically significant associations were observed between proinsulin levels and ~~with~~ the primary cIMT phenotypes IMT_{mean} and IMT_{max} . Nominal associations ($p < 0.05$) were observed with the IMT_{max} progression, but not baseline measures, independently of established CVD risk factors (Table 2). Secondary analyses indicated that the common carotid (excluding the

1st cm proximal to the bifurcation) demonstrated (non-significant) associations between higher proinsulin levels and larger cIMT measures (Supplemental Table 9). All other segments demonstrated inverse associations.

Effect of proinsulin-increasing SNP scores on IMT_{mean} and IMT_{max}

Reported proinsulin-associated SNPs demonstrated effects sizes generally comparable with those previously reported ¹⁰ (Supplemental Tables 2 and 10), thus were appropriate for use in the MR analysis. A total of 11 proinsulin-associated SNPs were included in the SNP score used for the MR analysis (10 previously reported SNPs in 9 loci and the chromosome 15 SNP, rs8029765). Instrument variable analysis demonstrated that the 11-SNP score was significantly and inversely associated with baseline IMT_{mean} and IMT_{max} but not with progression measures (Table 3). However, MR-Eggers (which is more robust when pleiotropy is present ³¹) refuted a significant association between the 11-SNP score and IMT_{mean} (-0.020 ± 0.033 $p=0.5721$) or IMT_{max} (-0.060 ± 0.054 $p=0.2738$). As pleiotropic SNPs can bias estimates ³¹, a further score was constructed, using only non-pleiotropic SNPs (*SNX7* rs1571500, *LARP6* rs7163439 and *UNC45A* rs8029765). This SNP score was not associated with baseline or progression of cIMT measures (Table 3). Replication of the non-pleiotropic SNP score was attempted in 12,113 participants with baseline mean and maximum IMT measures comparable with those in IMPROVE. Meta-analysis of the effect of the 3-SNP score on baseline IMT_{mean} and IMT_{max} demonstrated no association (Supplemental Figure 4 and Supplemental Table 11). Analyses of secondary IMT phenotypes demonstrated that no

single segment differed from the primary phenotypes analysed (Supplemental Table 12). Proinsulin levels according to SNP score are presented in Supplemental Table 13.

Mendelian Randomization refutes causality of proinsulin on cIMT

Proinsulin levels vary over time ~~with and thus a single~~ measurement of levels reflecting the exposure of the vasculature to proinsulin at a single point in time. In contrast, the proinsulin-increasing SNP score reflects a lifetime of exposure to proinsulin. For causality to be demonstrated, the direction of effect of the proinsulin-increasing SNP score on cIMT needs to be consistent with the effect of higher levels of proinsulin on cIMT. A further requirement is that the SNPs only act on cIMT through proinsulin levels. Here, the direction of effect was consistent between proinsulin levels and proinsulin-increasing SNP scores: higher levels of proinsulin were (non-significantly) associated with smaller cIMT measures in IMPROVE and proinsulin-increasing SNP scores were associated with smaller cIMT measures. However, this finding does not imply causality, because when only the non-pleiotropic SNPs were included in the SNP score, no association with cIMT was observed. Taken together, proinsulin *per se* is unlikely to have a causal effect on cIMT.

Discussion

It has previously been reported that circulating levels of proinsulin are predictive of CVD⁵⁻⁹. To the best of our knowledge, this is the first study to ~~examine address the question~~ ~~of~~ whether proinsulin has a causal effect on early vascular processes, such as increased thickness of the carotid artery wall measured by cIMT. In this study we identified a proinsulin-associated locus and used this, and other previously reported proinsulin-associated SNPs¹⁰, in a MR experiment, which demonstrated that proinsulin is unlikely to be a causal factor in determining cIMT.

The robust proinsulin-associated loci generally had sizes and directions of effect in IMPROVE that were comparable with those previously reported¹⁰, which added support to the validity of the proinsulin-associated SNP, and enabled the use of proinsulin SNP scores in an MR analysis. The locus (lead SNP rs8029765) demonstrated a genotype-specific effect on *UNC45A* expression, with the allele associated with higher levels of proinsulin being associated with lower expression of *UNC45A*. Low levels of *UNC45A* have been shown to inhibit cell proliferation and differentiation whereas high levels are associated with a number of cancers³². The chaperone protein encoded by *UNC45A* is also involved in intra-cellular trafficking, via its association with myosin^{32, 33}. It is possible that genetically low *UNC45A* levels could influence proinsulin levels, either via aberrant trafficking of vesicles (leading to release of incompletely processed proinsulin) or as a result of a reduced number of β -cells in the pancreas (leaving the pancreas more vulnerable to damaging stress during insulin resistance and hyperinsulinemia). Five genetic variants in the *UNC45A* genomic region have been reported in association

with metabolic traits (GWAS catalogue, dated 20161017). However, these are all ~500kb from the chromosome 15 locus and are at most in low linkage disequilibrium with rs8029765 ($r^2 < 0.11$, Supplemental Figure 3). In addition, they show no association with proinsulin levels in this study (smallest $p > 0.2411$), suggesting that these signals are unrelated. The lack of effect of rs8029765 on T2D is consistent with other proinsulin-associated loci, for example *PCSK1*, which demonstrated strong effects on proinsulin levels but convincingly null effects on T2D¹⁰.

The MR analyses can be used to demonstrate causality if i) proinsulin levels are associated with cIMT; ii) SNPs demonstrate robust associations with proinsulin; and iii) SNPs influence cIMT only through proinsulin, and not via other mechanisms. Proinsulin levels have been shown to influence cIMT in a previous report from a cross-sectional study of healthy middle-aged men⁴. However, in IMPROVE the association was weak and non-significant. Robust associations between SNPs and proinsulin levels have been reported¹⁰ and were consistent in IMPROVE. Whilst some association was observed between the 11-SNP score and cIMT, adjustment for pleiotropy (using MR-Eggers) or omission of the pleiotropic SNPs (3-SNP score) rendered the associations with cIMT non-significant. These results suggest that associations between proinsulin and cIMT are most likely due to confounding and pleiotropic effects of SNPs. Despite these findings, CVD prevention strategies which aim to reduce insulin resistance are still of value, as insulin resistance promotes other CVD risk factors including obesity and leads to T2D. In addition, proinsulin processing to insulin also releases c-peptide, which

negatively correlate with HDL levels ³⁴, therefore reducing hyperinsulinemia could enable increased HDL levels.

A limitation of this study is lack of power to detect an association, with much larger sample sizes (n>350,000) being required to conclusively address this question. MR analyses are frequently weakened by small effect sizes of genetic variants and inability to collect large enough sample sizes to enable drawing of definitive conclusions. In the absence of sufficient sample sizes with cIMT measures, alternative methodologies would be required to confirm these findings. Metabolic phenotypes are particularly challenging because of the complexity of interlinking mechanisms.

A unique strength of this study is the detailed measurements of the cIMT. The secondary analyses in the present study demonstrated that a positive association between proinsulin and cIMT was limited to only one segment in the carotid tree, which is an interesting and novel observation. However the uniqueness of IMPROVE is also a limiting factor in that there are few available cohorts with comparable cIMT measures, which precludes replication of the cIMT segments considered in the secondary analysis.

IMT measures are often rather imprecise, especially when considering progression over a relatively short time frame. We recognise this limitation, but note that the standard deviations for the studies included in this analysis are not excessive, that there is consistency across the studies and we have focused on the baseline measures and merely comment upon progression data. The weak association between proinsulin and cIMT is a weakness of this study, however there is prior independent evidence for the association between proinsulin and cIMT ⁴. Biomarker measurements reflect a snapshot

Commentator [RS1]: Don't want to just say that its crap... or were leaving ourselves open to the "then why did you bother doing this study?" question!
If there is a better way of rephrasing this, please let me know!

in time, whereas genetic variants reflect the lifetime exposure to biomarker levels. Thus by using genetic variants robustly associated with proinsulin levels, MR analyses give a more stable lifetime risk assessment. For a genetic discovery analysis, the IMPROVE cohort is relatively small. However, the high CVD-risk profile of all subjects renders it a relevant cohort for association studies of CVD-associated biomarkers. In addition, all participants had at least 3 of 9 established CVD risk factors ¹², meaning that the repertoire of potential confounders was very varied across the cohort and thereby limiting the potential for selection bias. Replication of the chromosome 15 proinsulin-associated SNP was conducted in a population of similar size to that utilised in the proinsulin GWAS ¹⁰, thus giving confidence to this finding.

In conclusion, we have identified a proinsulin-associated locus and used MR to demonstrate that proinsulin *per se* is unlikely to have a causal effect on cIMT, a proxy measurement for early atherosclerosis.

Disclosures Conflict of interest

A.D has received consultancy and research support from Metagenics Inc. (outside the scope of submitted work). Metagenics Inc. had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript. **E.I** is an advisor and consultant for Precision Wellness, Inc., and advisor for Cellink for work unrelated to the present project.

Funding Sources-Financial support

IMPROVE was supported by the European Commission (Contract number: QLG1-CT-2002-00896), the Swedish Heart-Lung Foundation, the Swedish Research Council (projects 8691 and 09533), the Kn

ut and Alice Wallenberg Foundation, the Foundation for Strategic Research, the Stockholm County Council (project 592229), the Strategic Cardiovascular and Diabetes Programmes of Karolinska Institutet and Stockholm County Council, the European Union Framework Programme 7 (FP7/2007-2013) for the Innovative Medicine Initiative under grant agreement n° IMI/115006 (the SUMMIT consortium), the Academy of Finland (Grant #110413), the British Heart Foundation (RG2008/08, RG2008/014) and the Italian Ministry of Health (Ricerca Corrente). The **Fenland** Study is funded by the Wellcome Trust and the Medical Research Council (MC_U106179471). We are grateful to all the volunteers, General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams. This research includes data and resources from the **FHS** of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research

was conducted using the Linux Clusters for Genetic Analysis (LinGA) computing resources at Boston University Medical Campus. Also supported by National Institute for Diabetes and Digestive and Kidney Diseases (NIDDK) R01 DK078616 and NIDDK K24 DK080140, and a Career Development Award from the American Diabetes Association. The **HBCS** has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation. The **NEO** study thanks all participants, general practitioners and research nurses involved in this study. We thank the NEO study group, Pat van Beelen, Petra Noordijk and Ingeborg de Jonge for the coordination, lab and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. **PIVUS** acknowledges the support of the Knut och Alice Wallenberg Foundation (Wallenberg Academy Fellow), European Research Council (ERC-2013-StG; grant no. 335395), Swedish Diabetes Foundation, Swedish Heart-Lung Foundation (grant no. 20120197), Swedish Research Council (grant no. 2012-1397), and Wellcome Trust (grant no. WT098017, WT090532, and WT064890). The analyses were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX)

under Project b2011036. **PROCARDIS** was funded by the European Commission Framework 6 (FP6) program (LSHM-CT-2007-037273), the British Heart Foundation, AstaZeneca, the Swedish Research Council (8691), the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation and the Torsten and Ragnar Söderberg Foundation. H.W acknowledges the support of the Wellcome Trust core award (090532/Z/09/Z) and the BHF Centre of Research Excellence. A.G and H.W acknowledge European Union Seventh Framework Programme FP7/2007-2013 under grant agreement no. HEALTH-F2-2013-601456 (CVGenes@Target) & and A.G, the Wellcome Trust Institutional strategic support fund. **RSI** and **RSII**: The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam; the Netherlands Organization for Scientific Research; the Netherlands Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly; The Netherlands Heart Foundation; the Ministry of Education, Culture and Science; the Ministry of Health Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. Support for genotyping was provided by the Netherlands Organization for Scientific Research (NWO) (175.010.2005.011, 911.03.012), the Netherlands Genomics Initiative (NGI)/NWO project nr. 050-060-810 and Research Institute for Diseases in the Elderly (RIDE). **ULSAM** was supported by the Knut och Alice Wallenberg Foundation (Wallenberg Academy Fellow), European Research Council (ERC-2013-StG; grant no. 335395), Swedish Diabetes Foundation, Swedish Heart-Lung Foundation (grant no. 20120197), Swedish Research Council (grant no. 2012-1397), and Wellcome Trust (grant no. WT098017, WT090532, and

WT064890). The analyses were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. Genotyping of Swedish samples was performed by **the SNP&SEQ Technology Platform** in Uppsala (www.genotyping.se) which is part of the National Genomics Infrastructure at Science for Life Laboratory, Uppsala University Sweden. The SNP&SEQ Technology Platform is also supported by the Swedish Research Council for Infrastructures and the Knut and Alice Wallenberg Foundation.

M.K is supported by The Netherlands Organisation for Scientific Research Grant VENI, 91616079. **B.S.** is financially supported by the Knut and Alice Wallenberg Foundation as part of the National Bioinformatics Infrastructure Sweden at SciLifeLab. **M.d.H** is supported by grants from the Swedish Heart-Lung Foundation (20140543), Swedish Research Council (2015-03657), and NIH (R01DK106236). **C.M.L.** is a Wellcome Trust Research Career Development Fellow (086596/Z/08/Z). **O.H.F** works in ErasmusAGE, a center for aging research across the life course funded by Nestlé Nutrition (Nestec Ltd), Metagenics, Inc, and AXA. **E.I.** is supported by grants from the Knut and Alice Wallenberg Foundation, the European Research Council (ERC Starting Grant, 335395), the Swedish Heart-Lung Foundation (20120197), and the Swedish Research Council (2012-1397). **J.B.M** is supported by R01DK078616 and K24 DK080140. **D.O.M-K** is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023). **J.W.J** acknowledges that the research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F2-2009-223004 PHASE. **A.P.M** is a Wellcome Trust Senior

Fellow in Basic Biomedical Science (grant number WT098017). **S.E.H** is funded by the National Institute for Health Research UCL Hospitals Biomedical Research Centre. None of the funding sources were involved in the design, collection, analysis or interpretation of the study.

Acknowledgments

We thank all participants and those involved in the recruitment of the cohorts included in this study.

References

- [1] Bavenholm, P, Proudler, A, Tornvall, P, et al., Insulin, intact and split proinsulin, and coronary artery disease in young men, *Circulation*, 1995;92:1422-1429.
- [2] Roder, ME, Porte, D, Jr., Schwartz, RS, et al., Disproportionately elevated proinsulin levels reflect the degree of impaired B cell secretory capacity in patients with noninsulin-dependent diabetes mellitus, *J Clin Endocrinol Metab*, 1998;83:604-608.
- [3] Paneni, F and Luscher, TF, Cardiovascular Protection in the Treatment of Type 2 Diabetes: A Review of Clinical Trial Results Across Drug Classes, *Am J Cardiol*, 2017;120:S17-S27.
- [4] Boquist, S, Ruotolo, G, Tang, R, et al., Alimentary lipemia, postprandial triglyceride-rich lipoproteins, and common carotid intima-media thickness in healthy, middle-aged men, *Circulation*, 1999;100:723-728.
- [5] Kahn, SE, Leonetti, DL, Prigeon, RL, et al., Relationship of proinsulin and insulin with noninsulin-dependent diabetes mellitus and coronary heart disease in Japanese-American men: impact of obesity--clinical research center study, *J Clin Endocrinol Metab*, 1995;80:1399-1406.
- [6] Yudkin, JS, May, M, Elwood, P, et al., Concentrations of proinsulin like molecules predict coronary heart disease risk independently of insulin: prospective data from the Caerphilly Study, *Diabetologia*, 2002;45:327-336.
- [7] Yudkin, JS, Phillips, DIW and Stanner, S, Proteinuria and progressive renal disease: birth weight and microalbuminuria, *Nephrol Dial Transpl*, 1997;12:10-13.
- [8] Lindahl, B, Dinesen, B, Eliasson, M, et al., High proinsulin concentration precedes acute myocardial infarction in a nondiabetic population, *Metabolism*, 1999;48:1197-1202.
- [9] Zethelius, B, Byberg, L, Hales, CN, et al., Proinsulin is an independent predictor of coronary heart disease: Report from a 27-year follow-up study, *Circulation*, 2002;105:2153-2158.
- [10] Strawbridge, RJ, Dupuis, J, Prokopenko, I, et al., Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes, *Diabetes*, 2011;60:2624-2634.
- [11] Baldassarre, D, Hamsten, A, Veglia, F, et al., Measurements of carotid intima-media thickness and of interadventitia common carotid diameter improve prediction of cardiovascular events: results of the IMPROVE (Carotid Intima Media Thickness [IMT] and IMT-Progression as Predictors of Vascular Events in a High Risk European Population) study, *J Am Coll Cardiol*, 2012;60:1489-1499.
- [12] Baldassarre, D, Nyssonen, K, Rauramaa, R, et al., Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study, *Eur Heart J*, 2010;31:614-622.
- [13] Voight, BF, Kang, HM, Ding, J, et al., The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits, *PLoS Genet*, 2012;8:e1002793.
- [14] 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;43:1193-1201.
- [15] Purcell, S, Neale, B, Todd-Brown, K, et al., PLINK: a tool set for whole-genome association and population-based linkage analyses, *Am J Hum Genet*, 2007;81:559-575.
- [16] Splansky, GL, Corey, D, Yang, Q, et al., The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination, *Am J Epidemiol*, 2007;165:1328-1335.

- [17] Feinleib, M, Kannel, WB, Garrison, RJ, et al., The Framingham Offspring Study. Design and preliminary data, *Prev Med*, 1975;4:518-525.
- [18] Dawber, TR, Meadors, GF and Moore, FE, Jr., Epidemiological approaches to heart disease: the Framingham Study, *Am J Public Health Nations Health*, 1951;41:279-281.
- [19] Eriksson, JG, Early growth and coronary heart disease and type 2 diabetes: findings from the Helsinki Birth Cohort Study (HBCS), *Am J Clin Nutr*, 2011;94:1799S-1802S.
- [20] Andersson, J, Sundstrom, J, Gustavsson, T, et al., Echogenicity of the carotid intima-media complex is related to cardiovascular risk factors, dyslipidemia, oxidative stress and inflammation: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study, *Atherosclerosis*, 2009;204:612-618.
- [21] Ingelsson, E, Sundstrom, J, Lind, L, et al., Low-grade albuminuria and the incidence of heart failure in a community-based cohort of elderly men, *Eur Heart J*, 2007;28:1739-1745.
- [22] Hedstrand, H, A study of middle-aged men with particular reference to risk factors for cardiovascular disease, *Ups J Med Sci Suppl*, 1975;19:1-61.
- [23] Wohlin, M, Sundstrom, J, Lannfelt, L, et al., Apolipoprotein E epsilon4 genotype is independently associated with increased intima-media thickness in a recessive pattern, *Lipids*, 2007;42:451-456.
- [24] Farrall, M, Green, FR, Peden, JF, et al., Genome-wide mapping of susceptibility to coronary artery disease identifies a novel replicated locus on chromosome 17, *PLoS Genet*, 2006;2:e72.
- [25] Magi, R and Morris, AP, GWAMA: software for genome-wide association meta-analysis, *BMC Bioinformatics*, 2010;11:288.
- [26] Folkersen, L, van't Hooft, F, Chernogubova, E, et al., Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease, *Circ Cardiovasc Genet*, 2010;3:365-373.
- [27] Consortium, GT, The Genotype-Tissue Expression (GTEx) project, *Nat Genet*, 2013;45:580-585.
- [28] Persson, J, Strawbridge, RJ, McLeod, O, et al., Sex-Specific Effects of Adiponectin on Carotid Intima-Media Thickness and Incident Cardiovascular Disease, *J Am Heart Assoc*, 2015;4:e001853.
- [29] Morris, AP, Voight, BF, Teslovich, TM, et al., Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes, *Nat Genet*, 2012;44:981-990.
- [30] Fratev, F, Osk Jonsdottir, S and Pajeva, I, Structural insight into the UNC-45-myosin complex, *Proteins*, 2013;81:1212-1221.
- [31] Bowden, J, Davey Smith, G and Burgess, S, Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression, *Int J Epidemiol*, 2015;44:512-525.
- [32] Ni, W and Odunuga, OO, UCS proteins: chaperones for myosin and co-chaperones for Hsp90, *Subcell Biochem*, 2015;78:133-152.
- [33] Jilani, Y, Lu, S, Lei, H, et al., UNC45A localizes to centrosomes and regulates cancer cell proliferation through ChK1 activation, *Cancer Lett*, 2015;357:114-120.
- [34] Li, Y, Li, Y, Meng, L, et al., Association between serum C-peptide as a risk factor for cardiovascular disease and high-density lipoprotein cholesterol levels in nondiabetic individuals, *PLoS One*, 2015;10:e112281.

Figure Legends

Figure 1: Manhattan plots of the association between SNPs and proinsulin levels in participants without T2D, adjusted for the glucose model. SNPs within 250Kb up or downstream of SNPs with $p < 1 \times 10^{-5}$ are highlighted in green.

Figure 2: Regional plots of suggestive loci. (A) Chromosome 11 locus. (B) Chromosome 11 locus after conditioning on the previously reported *ARAP1* loci (rs11603334 and rs77756118). (C) Chromosome 15 locus. (D) Chromosome 15 locus adjusted for the previously reported, nearby loci (*LARP6* rs7163439 and *VPS13C* rs4502156). (E) Chromosome 6 locus.