HAPLOTYPIC ANALYSES OF THE *IGF2-INS-TH* GENE CLUSTER IN RELATION TO CARDIOVASCULAR RISK TRAITS

Santiago Rodríguez[1] *; Tom R. Gaunt[1] *; Sandra D. O'Dell[1]; Xiao-he Chen[1]; Dongfeng Gu[1]; Emma Hawe[2]; George J. Miller[2]; Stephen E. Humphries[2]; Ian N. M. Day[1]

* Equal contributorship

[1] *Human Genetics Division, School of Medicine, University of Southampton, Southampton General Hospital, UK. http://www.sgel.humgen.soton.ac.uk*

[2] *Centre for the Genetics of Cardiovascular Disease, Royal Free and University College London Medical School, London, UK.*

ABSTRACT

Three important genes are clustered on human chromosome 11p15: insulin-like growth factor 2 (*IGF2*), insulin (*INS*) and tyrosine hydroxylase (*TH*). We undertook haplotypic studies of the whole cluster with regard to nine traits related to body composition, plasma triglyceride (TG) levels and blood pressure (BP) in 2743 middle aged males. Six out of the ten common haplotypes found showed significant association with at least one trait. Two specific haplotypes protect against obesity. Raised plasma TG levels and elevated BP and fat mass are associated with two different haplotypes containing the same specific alleles for *TH* and *INS* genes. A retrospective power analysis based on the observed haplotype frequencies indicates that our study has a 90% power to detect small differences between the mean values for each haplotype and those observed for all individuals, for the cardiovascular traits analysed. Our haplotype studies also integrate a complex literature mostly based on only one of the genes. The simplest hypothesis is that haplotypes containing a long repeat insertion in the *INS* gene promoter (*INS* "class III" alleles) may be predisposing for metabolic syndrome traits (including hypertriglyceridemia, high BP, type II diabetes and PCOS), whereas low body mass, and type I diabetes reflect the effects of at least one different haplotype, which is marked by the shortest alleles of the "class I" group in the *INS* promoter.

Figure 1

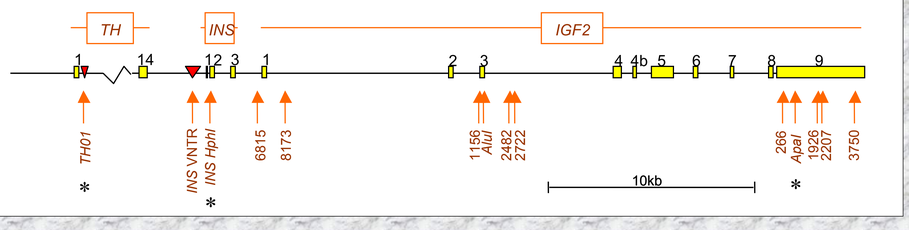


Figure 1. Genetic map of the *IGF2-INS-TH* region spanning around 40 kb of 11p15. Exons of *TH*, *INS* and *IGF2* are numbered and represented by boxes. Indicated by an asterisk are three markers that tag the information given by the 12 markers analysed (see table 1).

		TH01						*INS* VNTR		
IGF2 (11 SNP)	Code	6	7	8	9	9.3	Total	Class I	Class III	Total
12112111111	A	**63**	0	6	5	**28**	102 (18.1%)	**119**	0	119 (19.8%)
11221211111	B	0	6	2	12	**50**	70 (12.4%)	14	**50**	64 (10.6%)
21121112211	C	1	0	8	**42**	5	56 (9.9%)	**49**	2	51 (8.5%)
11212111111	D	0	**41**	4	2	2	49 (8.7%)	**33**	5	38 (6.3%)
11221212211	E	0	0	1	1	**27**	29 (5.1%)	0	**24**	24 (4.0%)
Total		64	47	21	62	112	306 (54.3%)	215	81	296 (49.2%)

Table 1. Interallelic associations found between the five commonest *IGF2* haplotypes and *TH01* and *INS* VNTR alleles in NPHSII men. The information given by *TH01* and the 11 *IGF2* SNPs can be reduced to a combination of the microsatellite and one single *IGF2* SNP (*ApaI* in red). Addition of the *INS HphI* site permits distinction between classI/classIII alleles at the *INS* VNTR locus (1).

Haplotype	*IGF2 ApaI*	*INS* VNTR	*TH01*	Frequency	*IGF2* (11 snps)
*IGF2-INS-TH**1	1 (G)	Class I	6	825 (20.0%)	A
*IGF2-INS-TH**2	1 (G)	Class I	7	716 (17.4%)	D
*IGF2-INS-TH**3	1 (G)	Class I	9.3	435 (10.6%)	A
*IGF2-INS-TH**4	1 (G)	Class III	9.3	419 (10.2%)	B
*IGF2-INS-TH**5	2 (A)	Class I	9	372 (9.0%)	C
*IGF2-INS-TH**6	2 (A)	Class III	9.3	323 (7.8%)	E
*IGF2-INS-TH**7	1 (G)	Class I	8	154 (3.7%)	A or D
*IGF2-INS-TH**8	1 (G)	Class III	8	145 (3.5%)	B
*IGF2-INS-TH**9	1 (G)	Class I	9	145 (3.5%)	A or D
*IGF2-INS-TH**10	2 (A)	Class I	9.3	142 (3.4%)	E

Table 2. Characteristics of the *IGF2-INS-TH* haplotypes considered for association analysis with cardiovascular traits. Consistent with approaches to reduce risk of false positive findings related to low-frequency haplotypes (2), we considered only haplotypes with a frequency higher than 3%, representing 89% of all haplotypes.

Figure 2

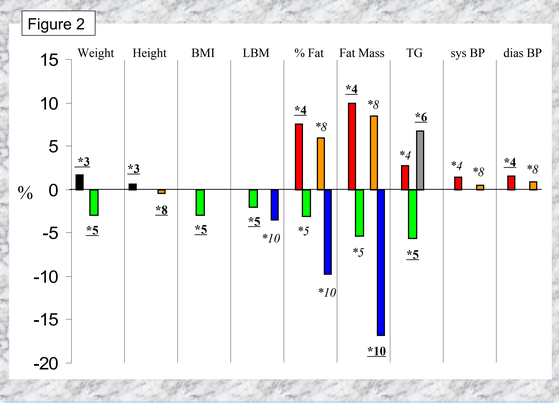


Figure 2. Significant associations (in bold and underlined) between *IGF2-INS-TH* haplotypes and cardiovascular risk traits. The bars represent the percentage difference observed between the mean value of a given haplotype and the overall sample mean for each trait.

Figure 3

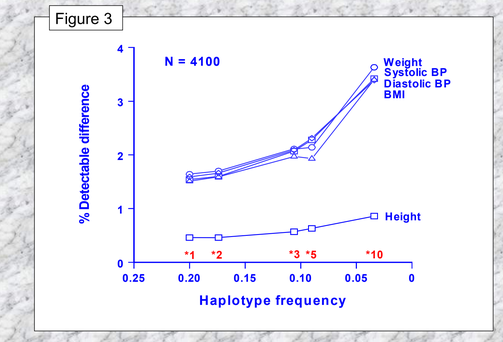
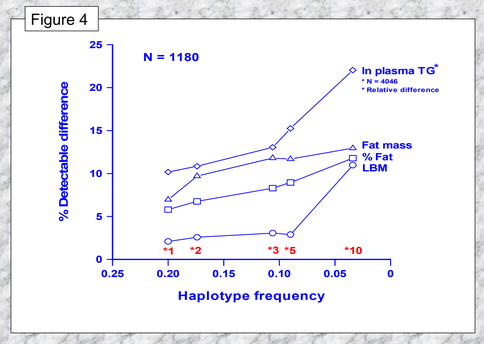


Figure 4



Figures 3 and 4. Percentage of detectable differences between the mean values for given haplotypes and the mean values for all individuals for each trait. There is a 90% power to detect very small differences for weight, systolic BP, diastolic BP, BMI and height (less than 4% for the rarest haplotype and around or less than 1.5% for the commonest one) (Figure 3). Detectable differences for fat mass, percentage fat and LBM were around 10% for the rarest haplotype, and approximately 20% for plasma TG (Figure 4) whereas considerably smaller differences could be detected with a 90% power for the more frequent haplotypes.

Figure 5

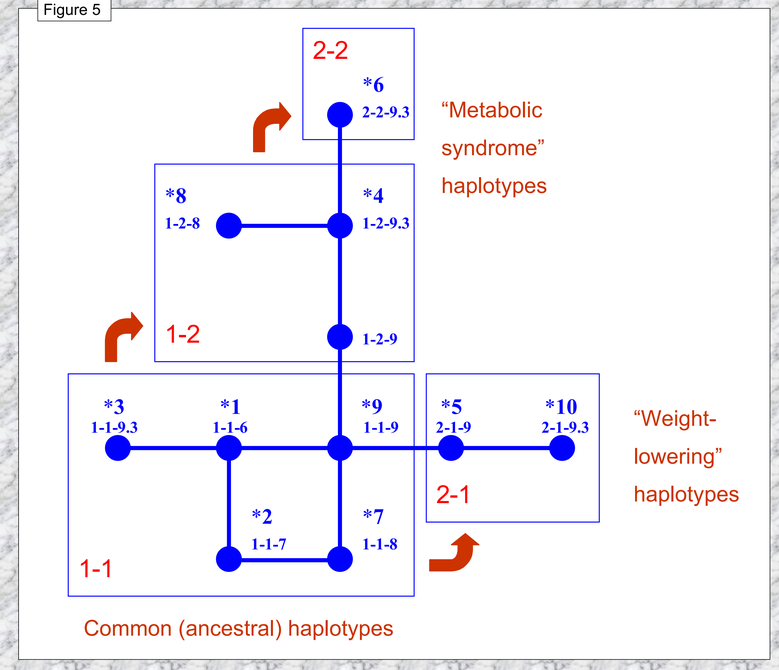


Figure 5. Cladogram relating the 10 haplotypes considered at the *IGF2-INS-TH* region. The commonest haplotypes, which are likely the oldest (3), marked by the presence of the common alleles at *IGF2 ApaI* and *INS HphI*, do not show association with the traits analysed. From this clade, two new clades with effect on cardiovascular traits emerged and these are tagged by alleles either of *IGF2 ApaI* or *INS HphI*. Mutation at *IGF2 ApaI* marks weight-lowering haplotypes, whereas mutation at *INS HphI* (equivalent to class III alleles at the *INS* VNTR), marks haplotypes with increased risk for metabolic syndrome traits. Finally, a more recent clade marked by the presence of the rarer alleles at both SNPs has effects on plasma TG variation.

Table 3. Summary of published apparent traits or disease effects associated with particular alleles and haplotypes at the *IGF2-INS-TH* region. The weight-lowering haplotype *5 contains the *INS VNTR* class I alleles most prominently linked and associated with type 1 diabetes, and it also contains the *TH01* allele 9 overrepresented in normotensive controls in a hypertension case study (4). *INS VNTR* class III associations with type 2 diabetes, PCOS and high TG levels in diabetics must reflect some or all of haplotypes *4, *6 and *8, of which *4 and *8 appear to associate with raised fat mass and diastolic BP and *6 with raised TG level. However, observations in hypertension and longevity relating to *TH01* allele 9.3 are difficult to interpret because this allele is found in association with a variety of *IGF2* and *INS* haplotypes.

IGF2 ApaI	*INS VNTR*	*TH01*	Apparent trait or disease effect	*IGF2-INS-TH* haplotype
2	class I	9	weight lowering	*5
2	class I	9.3	weight lowering	*10
1	class III	9.3	increased metabolic syndome risk	*4
2	classIII	9.3	high TG	*6
	class I		type 1 diabetes	apparently *5
	class I		childhood obesity	
	class I		increased prostate cancer risk	
	class III		type 2 diabetes	likely one or more of *3, *4, *6, *8
	class III		polycystic ovary syndrome	
		9	normotension	apparently *5
		9.3	hypertension	likely one or more of *3, *4, *6, *8
		9.3	increased longevity	

Conclusions

The analysis of the haplotypic effects of the *IGF2-INS-TH* region on a spectrum of cardiovascular risk traits revealed two weight-lowering haplotypes (*5 and 10) and two haplotypes confering increased metabolic syndrome risk (*4 and *6). The weight-lowering effect of *5 (1 BMI unit) represents about a 10% increase in the rate of coronary events (5). The potential significance of haplotype *4 to cardiovascular risk is not trivial. A 1.29mmHg higher diastolic BP equates with about a 6-8% increment in stroke risk and a 4-5% increment in CHD risk (6). Comparable estimations can be made for the 10% fat mass increment and for the 10% higher plasma TG of haplotype *6 in relation to the metabolic syndrome. This systematic description identifies tagging markers and specific haplotypes (particularly *4 for risk and *5 for protection), permitting deeper epidemiological and functional investigation of *IGF2-INS-TH* as a cardiovascular risk polygene.

Material and Methods

Experimental conditions for the genotyping of polymorphisms are as previously described (7). In brief, the eleven SNPs at *IGF2* and *HphI* were typed by PCR adopting the Amplification Refractory Mutation System (ARMS) (8) or by digesting each amplification product by the corresponding restriction enzyme. The *TH* microsatellite *TH01* and the *INS* minisatellite were amplified with primers and PCR conditions previously described (9-11). The amplification products (VNTR, microsatellite and SNP assays) in all cases were resolved by Microplate Array Diagonal Gel Electrophoresis (MADGE) as previously described (12-15). Three statistically different approaches were adopted, Haplotype Trend Regression (HTR) (16); haplo.score (17); and Phase (18) (combined with SPSS) to test for association of indirectly-deduced haplotypes with weight, height, BMI, LBM, percentage of fat, fat mass, plasma TG levels, systolic BP and diastolic BP. Plasma TG values were log normalised prior to analysis. These approaches were compared to determine the statistical robustness of their conclusions, given that haplotype uncertainty is a major issue complicating regression analyses in samples of unrelated individuals (19). Retrospective power analyses were also conducted for the detected associations and computed the percentage of detectable differences for each trait for a given power of 90% assuming a significant value of 0.05. Power calculations were performed as previously described (20) using a Normal Power Calculator (http://calculators.stat.ucla.edu/powercalc/) assuming normal distributions for the traits and unequal variances between the mean trait values for each haplotype and those observed for all individuals. A cladogram was constructed using the principle of parsimony from the haplotypes considered in the analyses. This was guided using cladistic analyses constructed in the program Evolutionary-Based Haplotype Analysis Package (EHAP) (21).

References

1. - Bennett,S.T., Todd,J.A. (1996) Annu.Rev.Genet., 30, 343-370.
2. - Comeron,J.M. et al (2003) Pac.Symp.Biocomput., 478-489.
3. - Zapata,C. et al (2001) Genetics, 158, 1235-1251.
4. - Sharma,P. et al (1998) Hypertension, 32, 676-682.
5. - Shaper,A.G. et al (1997) BMJ, 314, 1311-1317.
6. - MacMahon,S. et al (1990) Lancet, 335, 765-774.
7. - Gaunt,T.R. et al (2001) Hum.Mol.Genet., 10, 1491-1501.
8. - Newton,C.R. et al (1989) Nucleic Acids Res., 17, 2503-2516.
9. - Weber,J.L., May,P.E. (1989) Am.J.Hum.Genet., 44, 388-396.
10. - Polymeropoulos,M.H. et al (1991) Nucleic Acids Res., 19, 3753.
11. - Bennett,S.T. et al (1995) Nat.Genet., 9, 284-292.
12. - Day,I.N., Humphries,S.E. (1994) Anal.Biochem., 222, 389-395.
13. - O'Dell,S.D. et al (2000) Hum.Mutat., 15, 565-576.
14. - O'Dell,S.D. et al (2000) Biotechniques, 29, 500-506.
15. - Chen,X.H. et al (2002) Biotechniques, 32, 1080-2, 1084, 1086.
16. - Zaykin,D.V. et al (2002) Hum.Hered., 53, 79-91.
17. - Schaid,D.J. et al (2002) Am.J.Hum.Genet., 70, 425-434.
18. - Stephens,M. et al (2001) Am.J.Hum.Genet., 68, 978-989.
19. - Zhao,H. et al (2003) Pharmacogenomics., 4, 171-178.
20. - Thompson,S.K. (2002) Sampling. John Wiley & Sons, Inc., NY.
21. - Seltman,H. et al (2003) Genet.Epidemiol., 25, 48-58.

Acknowledgements

S.R. was an EC Marie Curie Research Fellow. I.N.M.D. was a Lister Institute Professor. The UK Medical Research Council and the British Heart Foundation are thanked for support.