

A susceptibility locus for myopia in the normal population is linked to the PAX6 gene region: a genome-wide scan of dizygotic twins.

Running title: PAX6 and myopia in the normal population

Christopher J Hammond^{1,2}, Toby Andrew¹, Ying Tat Mak¹, Tim D Spector¹

¹Twin Research and Genetic Epidemiology Unit, St Thomas' Hospital, London, SE1 7EH, United Kingdom

²West Kent Eye Center, Princess Royal University Hospital, Orpington, BR6 8ND, United Kingdom

Correspondence to:

Christopher J Hammond MD FRCOphth

Twin Research and Genetic Epidemiology Unit, St Thomas' Hospital, London, SE1 7EH, United Kingdom

Email: chammond@btopenworld.com

Tel: +442071886730

Fax: +442071886761

Summary

Myopia is a common, complex trait with considerable economic and social impact and, in highly-affected individuals, ocular morbidity. We performed a classic twin study of 506 unselected twin pairs, and inferred the heritability of refractive error to be 0.89 (95% CI .86-.91). A genome-wide scan of 221 dizygotic twin pairs, analysed using optimal Haseman-Elston regression methods implemented using generalised linear modeling, showed significant linkage (LOD>3.2) to refractive error at four loci, with the maximum LOD score of 6.1 at 40cM on chromosome 11p13. Evidence of linkage for this locus, and the other linkage peaks at chromosomes 3q26 (LOD 3.7), 8p23 (LOD 4.1) and 4q12 (LOD 3.3), remained the same or became stronger after checking model fit and downweighting of outliers. Examination of potential candidate genes showed the PAX6 gene directly below the highest peak at the 11p13 locus. PAX6 is fundamental to identity and growth of the eye, but reported mutations usually result in catastrophic congenital phenotypes such as aniridia. Haplotype tagging of 17 common allele frequency SNPs covering the PAX6 gene identified 5 SNPs which explained 0.999 of the haplotype diversity. Linkage and association analysis of the tagging SNPs showed strong evidence of linkage for all markers with a minimum χ_1^2 of 7.5 ($p = 0.006$), but no association. This suggests that PAX6 may play a role in myopia development, possibly due to genetic variation in an upstream promoter or regulator, although no definite association between PAX6 common variants and myopia was demonstrated in this study.

Introduction

Refractive error is a common, complex trait measured on a continuous scale, with myopia affecting 25-61% of the population (Saw et al. 1996; Wang et al. 1994). Its incidence may be increasing rapidly (Tay et al. 1992), with not only economic implications, but also a significant risk of permanent visual loss in those individuals who have high myopia (“pathological myopia”) which is also becoming more common (Dandona & Dandona 2001; Rosenberg & Klie 1996). Susceptibility loci have been identified in familial (AD) high myopia (Naiglin et al. 2002; Paluru et al. 2003; Young et al. 1998b; Young et al. 1998a), although to date these loci have not been associated with lower degrees of juvenile-onset myopia from population-based samples (“simple myopia”) (Mutti et al. 2002b). High myopia may be influenced by certain candidate genes, such as those involved in Marfan [MIM #154700] or Stickler [MIM #108300] syndromes, but high myopia only accounts for 2-3% of myopia (Dandona & Dandona 2001). To our knowledge no susceptibility loci for simple myopia have been published previously.

The mechanism of myopia development is not fully understood. Simple myopia results from a failure of emmetropisation; excessive eye growth results in images from distant objects being focused in front of the retina. Debate about relative contributions of genes and environment (in particular close work) has continued (Mutti et al. 1996). Cross-sectional studies suggest that the majority of individual variation about the population mean is attributable to genetic variance (Hammond et al. 2001; Mutti et al. 2002a). The rapid growth in the incidence of myopia in SE Asia (Tay et al. 1992), however, suggests that recent generational differences are more

likely to be due to environmental mechanisms such as increased literacy rather than changes in allele frequencies. There are many potential candidate genes controlling both the fundamental development of the eye and variations in eye growth.

The PAX6 gene [MIM 607108] is a member of the paired domain Pax family and encodes transcriptional regulators involved in oculogenesis and body development, aspects of which are conserved throughout phylogeny (Simpson & Price 2002). The gene is widely expressed in the developing eye, both in neuroectoderm and surface ectoderm, as well as in the adult eye (Ashery-Padan & Gruss 2001). Most described PAX6 mutations have been deletions resulting in premature protein truncations, haploinsufficiency and catastrophic congenital phenotypes rather than minor normal variations. For example, PAX6 has been the sole gene implicated in aniridia [MIM #106210], a semi dominantly inherited condition (Glaser et al. 1992), which is characterized by severe iris hypoplasia with other associated abnormalities, and usually results in marked visual impairment (van Heyningen & Williamson 2002). Aniridia is a progressive condition, which suggests the possibility that PAX6 may also play a maintenance role in the adult eye, and is not only involved in oculogenesis. To our knowledge, PAX6 has not been implicated in non-pathogenic myopia development in previous studies.

A classical twin study was undertaken using monozygotic (MZ) and dizygotic (DZ) twins to initially establish the heritability of refractive error. This was followed up with a genome-wide linkage study for the continuous trait using unselected DZ sib-pairs, conducted to ascertain potential susceptibility loci for myopia. Based on the linkage analysis results, an association study for the PAX6 gene was then performed.

Materials and Methods

Refractive error measurement

Refractive error was measured in 506 volunteer twin pairs using a Humphrey-670 Autorefractor, as part of the Twin Eye Study (Hammond et al. 2001). Refractive error was determined by the mean spherical equivalent for the two eyes of each individual measured in diopters. The 280 dizygotic (DZ) and 226 monozygotic (MZ) twin pairs were part of the Twins UK Registry (Spector & MacGregor 2002), and were unselected as to eye traits, having volunteered for the registry unaware of any future potential eye studies or hypotheses. Twins were excluded if they had previous interventions that might have affected refraction, such as cataract or refractive surgery. Informed written consent was provided after approval of the study by the hospital research ethics committee, and whole blood taken for DNA extraction. Zygosity was determined by standardized questionnaire (Martin & Martin 1975) and confirmed by DNA short tandem repeat fingerprinting in cases of doubt. Individuals in the study were considered to be randomly ascertained, because sampling was not based on a subject's eye data.

Heritability analysis

We have previously reported heritability of refractive error in this cohort (Hammond et al. 2001), but have repeated the analysis for the measure of mean spherical equivalent used in the genome-wide scan presented in this study. In short, quantitative model fitting to twin data is based on comparison of the covariances (or correlations) between monozygotic (MZ) and dizygotic (DZ) twin pairs and allows partitioning of

observed phenotypic variance into additive (A) or dominant (D) genetic components and common (C) or unique (E) environmental components. MZ twins are assumed to share the same A and D genetic variance, and DZ pairs share on average half A and a quarter D variance, while C is assumed to be the same for both MZ and DZ twin pairs (the “equal environment” assumption) (Neale & Cardon 1992; Snieder et al. 1997).

The variance due to A, C, D, E and age were estimated using normal theory and maximum likelihood methods implemented in Mx (Neale 1997). The significance of variance components A, C, D and age was assessed by removing each sequentially and testing the deterioration in model fit. Submodels were contrasted to full models using nested comparisons and an asymptotically chi-squared fit statistic. If there was no significant change in model fit, then the parameter could be removed, leading to a parsimonious model explaining the variance and covariances. Data handling and preliminary analyses were conducted using STATA (StataCorp 1997).

Genome-wide linkage study

A genome wide linkage analysis was performed on 221 of the 280 dizygotic twin pairs from the heritability study for which there were genomic marker data. The genomic scan was based on DNA extracted from venous blood samples of the study subjects. Scans involved using standard fluorescence-based genotyping methodologies (Pritchard et al. 1995; Reed et al. 1994) for the analysis of 737 highly polymorphic microsatellite markers from the ABI Prism linkage mapping set (Applied Biosystems) and Genethon Genetic Linkage Map (Dib et al. 1996), as has been described previously (Wilson et al. 2003). The estimated genotyping error rate was < 1%.

Linkage analysis

Multipoint genome-wide linkage analyses were performed using unadjusted mean spherical equivalent of both eyes (in diopters) and optimal Haseman and Elston regression methods, implemented using a Generalised Linear Model (GLM) (Barber et al. 2004). This method, implemented in Stata (Statacorp, College Station, TX), is algebraically equivalent to other likelihood techniques (Almasy & Blangero 1998; Kruglyak & Lander 1995), but has the advantage of being robust to deviations in multivariate normality by freely estimating the coefficient of variation (i.e. the mean and variance-corrected residual error) and by utilizing a robust Huber-estimate of variance (Barber et al. 2004). Regression diagnostics were used to check the reliability of model fit, including an iterative robust regression routine applied to GLM using Huber followed by biweight iterations (StataCorp 1997). On a genome-wide basis, any observed divergence between GLM and iterative robust GLM (IR) results were used to indicate potentially poor fitting models for specific regions.

The map positions and ordering of all marker loci were determined from the Genethon Genetic Linkage Maps. Approximate support intervals were generated using a 1 LOD drop interval approach. Twin pairs formed independent families and no additional sibs were considered in the analysis.

Association analysis

Out of the known genes observed under the linkage peak on chr.11p13 (see Table 1), only PAX6 is directly implicated in the identity and growth of the eye (Dominguez et al. 2004) and has a well documented role in oculogenesis. We therefore undertook a

three-step candidate gene association study (Weale et al. 2003) by 1) characterising the haplotypic structure of the gene 2) identification of a subset of highly informative single-nucleotide polymorphism (SNP) markers in the region, referred to as tagging SNPs (tSNPs), 3) typing the tSNPs for the clinical sample and testing these for linkage and association, using QTDT (Abecasis et al. 2000a).

PAX6 SNP genotyping

SNPs were selected from public SNP databases (SNP Consortium and National Centre of Biotechnology Information) by criteria including population frequency validation, multiple submitters, high profile submitters and absence of repeats masking in the sequence. 17 SNPs in a 34 kb region on chromosome 11p13 were genotyped for 22 individuals to assess linkage disequilibrium in the region, covering 8 kb upstream and 4 kb downstream of the PAX6 gene. Details of these 17 SNPs are given in Table 2.

The PCR method is based on the accumulation of fluorescence during the PCR process by degradation of an internally quenched allele-specific probe bound to the same template. (Heid et al. 1996) Allelic discrimination is achieved by the use of two oligonucleotide probes, each complementary to one of the two alleles and labelled with a different fluorescent reporter dye. Applied Biosystems (Warrington, Cheshire, UK) performed the design of PCR primers and allelic specific Minor Groove Binder (MGB) probes and optimisation of reaction conditions for all the 17 SNPs. The company also supplied reagents and consumables used for TaqMan® PCR in our study.

The PCR reaction was performed with the GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems) in a volume of 25 µL containing 12.5 µL of TaqMan 2x Universal PCR Master Mix, 20-50 ng of genomic DNA, and the optimised primer/probe solution as supplied (Applied Biosystems). The PCR program included an initial step at 95 °C for 10 min followed by 40 cycles at 92 °C for 15 s and 60 °C for 1 min. Endpoint reading of fluorescence and analysis of the genotypes were performed with the ABI Prism® 7700 Sequence Detection System (Applied Biosystems).

Assessment of Pax6 haplotypic structure

Linkage disequilibria measures r^2 and D' (Hedrick 1987) were used to assess pairwise LD between markers, both within the immediate vicinity of Pax6 using 17 public domain SNPs with common minor allele frequencies ($MAF \geq 7\%$) and also a 250Kb region either side of Pax6 using HapMap data. tSNPs were originally identified using haplotype diversity criteria (Johnson et al. 2001). However, due to the high levels of disequilibria found in the gene, other more recent and potentially more efficient marker based selection criteria were used, such as association r^2 (Weale et al. 2003), which in this case also produced the same tagging solutions. Haplotype tagging in the 34Kb region identified 5 tSNPs, which explained 0.999 of the haplotype diversity arising from 17 common allele frequency SNPs (see Figure 1). The five tSNPs were subsequently genotyped in the original sample of 221 DZ twin pairs.

Results

The 506 twin pairs included in the heritability study were aged 49-79 years, with a mean age of 62.2 (S.D. 5.7) years. Their mean refractive error (spherical equivalent: SE) was +0.39 diopters (S.D. 2.38), with a range of -12.12D to +7.25D. The distribution was leptokurtic (5.87), with a left skew (-0.91), in keeping with other population studies of refractive error. There were no significant differences between MZ and DZ twins for age (62.4 vs 62.0 years) and refractive error (mean 0.36D vs 0.40D). Age accounted for only ~1% of the variance, and so was excluded from further analysis.

Refractive errors for twin 1 versus twin 2 for mean spherical equivalent (SE) are plotted in Figure 2. The pairwise intraclass correlation coefficient was 0.89 for MZ twins, and 0.49 for DZ twins, suggesting a significant genetic contribution to the variance. Univariate model fitting results for mean SE are displayed in Table 3, and show the best-fitting model to be the one that explained the variation of SE due to additive genetic effects (A) and unique environmental effects (E). The effects of shared family environment (C) and dominant genes (D) could be eliminated without significant deterioration in model fit. The parameter estimates for mean SE obtained by this model suggested a very high heritability of 0.89 (95% CI .86-.91), with the rest of variance explained by individual environmental effects (0.11, 95% CI .09-.14). These results do not differ significantly from our previous reported heritability analysis of right and left eyes separately (Hammond et al. 2001).

The 221 DZ twin pairs included in the genome-wide linkage analysis did not significantly differ from the whole cohort (mean age 61.4 years, mean SE +0.31D). Maximum evidence of linkage in the twins was observed at chromosome 11p13 (LOD

6.1), with further loci at chromosomes 3q26 (LOD 3.7) and 4q12 (LOD 3.3) and 8p23 (LOD 4.1) and 11q23-24 (LOD 2.9), as demonstrated in Figure 3. Evidence for these loci became stronger after checking model fit residuals and the downweighting of unduly influential outliers using iterative regression GLM(IR), as displayed in Table 4. The linkage results for the four peaks with LOD scores greater than 3.2 are displayed in Figure 4, with marker information. To examine whether this linkage pertained to myopia as opposed to hyperopia (long sight), twin pairs were divided into myopes (mean SE ≤ 0 , for 85 pairs) and hyperopes (mean SE > 0 , for 136 pairs) using the arbitrary threshold of zero. Figure 4 demonstrates that the linkage signal was maintained in the myopic group despite the smaller numbers, albeit with a lower LOD score, as well as the hyperopic group. This suggests that the locus is likely to confer myopia susceptibility across the whole spectrum of refractive errors.

The nearest markers to the maximum LOD score of 6.1 at 40cM from the p' terminal on chromosome 11p13 were D11S904 (37cM, two-point LOD score 4.28) and D11S935 (49cM, two-point LOD score 2.2). Support intervals, using the maximum LOD-1 approach, equated to 35-45cM, in which there are 44 described genes or potential genes. Some possible candidate genes are listed in Table 1. The most likely candidate gene in this area is PAX6, at 39-42cM, directly beneath the highest linkage peak at 40cM. We therefore attempted to determine whether PAX6 was associated with myopia in this population.

Analysis of the tSNP genotypes and haplotypes using QTDT showed strong evidence of linkage for all markers (minimum of $\chi_1^2 = 7.5$, $p = 0.006$), as detailed in Table 5.

There was marginal evidence for population stratification for SNPs 1 & 3 (and

resulting haplotypes). All five SNP polymorphisms were in Hardy Weinberg equilibrium and there was no evidence for phenotypic association with any of the SNPs or haplotypes, using both tests of total association and tests of association robust to population stratification (Abecasis et al. 2000b). Tests of linkage controlling for association also did not significantly reduce the observed linkage signal.

Discussion

These results show a high heritability for refractive error of 0.89. This is the first study of myopia, to our knowledge, demonstrating significant linkage to susceptibility loci using a genome-wide screen. In this population-based sample we have found four loci with LOD scores greater than 3.2. A QTL in the vicinity of PAX6 that predisposes individuals to myopia has not previously been reported; only major developmental eye abnormalities have to date been associated with mutations in this gene. This means that PAX6 may play an important role post-natally in continued eye growth.

The strong evidence of linkage to PAX6 suggests that genetic variant(s) are located in the vicinity of the gene, perhaps within regulatory or promoter regions of PAX6 whose important role (Kellis et al. 2003) and greater numbers are increasingly recognised (Dermitzakis et al. 2003). There have been advances in the understanding of the control of eye growth. The dosage of PAX6 may also be important; overexpression of PAX6 in transgenic mice led to reduced eye size (Schedl et al. 1996). Recent work on *Drosophila* eyes have shown that two distinct Pax genes

control eye identity and growth separately, and that these equate to the PAX6 and PAX6(5a) isoforms of the gene in humans (Dominguez et al. 2004). This raises the exciting possibility that PAX6, or one of its isoforms, may be involved in eye growth in childhood, and therefore in the etiology of myopia.

Animal studies have suggested that eye growth is regulated by the quality of retinal image (“emmetropisation”) (Flitcroft 1998). If humans have the same mechanism, human myopia may occur if a child inherits a dysfunctional emmetropisation mechanism (Norton 1994), with the trigger being close work. The exact mechanisms are unclear. Humoral interactions occur between scleral and choroidal layers in response to negative and positive fitted lenses in chicks (Marzani & Wallman 1997), and this effect is also seen in primates (Hung et al. 2000). Retinoic acid may have a mediating role between vision and growth (Mertz & Wallman 2000). Further examination of the four loci significantly linked to myopia in this study may allow further elucidation of the mechanisms of myopia development.

The high heritability of myopia in this twin cohort of 0.89 (95% CI .86-.91) is almost identical to that reported in the Danish twin study (0.89-0.94) (Lyhne et al. 2001), and consistent with other family studies (The Framingham Offspring Eye Study Group 1996; Zadnik et al. 1994). The twins in this study were unselected for refractive error, having volunteered for osteoporosis and other studies, reducing the possibility of ascertainment bias. The twins in this study are representative of the general population for many common traits and exposures such as those in the cardiovascular and musculoskeletal systems (Andrew et al. 2001), supporting the generalisability of these results. Although the LOD scores of these loci are impressive for a common

complex trait, we need to confirm them in a replication study, to prioritise potential regions of interest for further research.

The failure to find a phenotypic association with a common SNP in the PAX6 gene could have occurred for one of several reasons – either the variant(s) lies within the gene and has been missed by the tagging SNPs, or the variant lies outside the PAX6 gene or the linkage signal is a false positive result.

For the first scenario, if the potential causal variant is common, it is unlikely one or a combination of the tagging SNPs we selected are not in LD with the polymorphism, since the common haplotype diversity captured was over 99% (Johnson et al. 2001). In addition, in a sampling experiment in which each of the original 17 SNPs was sequentially dropped (Weale et al. 2003), the 3 or 4 tagging SNP solution captured all the information in the dropped SNP (r^2 of 1 in all cases). The only exception to this was rs2239789 (SNP 8) where the r^2 was only 21%. This is because SNP8 lies in a region of the gene with low LD between exons 8 and 9 (see Figure 1). It is possible that a (combination of) rare variant(s), not in LD with the common tagging SNPs, may be associated with myopia; sequencing would be required for their identification.

However, if common and rare functional polymorphisms within the PAX6 gene are ruled out, then the alternative explanations for the linkage signal are either a different gene nearby (such as ELP4, see Table 1), or a variant in a nearby regulator of PAX6. Further detailed examination of PAX6 and its isoforms and regulatory regions is required.

Conclusion

Myopia is a common trait that is increasing in incidence and is strongly heritable.

Four loci have been identified from a genome-wide linkage analysis, with the highest peak on chromosome 11p13 (LOD 6.1). We have found strong linkage to refractive error in the PAX6 vicinity, a candidate gene at this locus. Further mapping is required to confirm whether the linkage to PAX6 is due regulatory loci or, in fact, due to linkage to another unrecognised nearby gene.

Acknowledgements

The authors would like to thank the Wellcome Trust who funded the original phenotyping, and the Chronic Diseases Research Foundation for continuing support. We acknowledge Sequenom Inc (San Diego, CA) for genotyping information, and would like to thank Mathew Barber for his help with data analysis early in the project and Kouros Ahmadi for his expert advice on candidate gene SNP tagging. Finally, we are grateful to all the twin volunteers for their ongoing help and support.

Electronic Database Information

HapMap data: <http://www.hapmap.org>; February 2004

PAX6, Marfan, Stickler and Aniridia MIM refs: Online Mendelian Inheritance in Man (OMIM): <http://www.ncbi.nlm.nih.gov/Omim/>

SNPs in figure 5: UCSC Genome Browser Gateway version July 2003 (<http://www.genome.ucsc.edu/cgi-bin/hgGateway?db=hg8>).

References

- Abecasis GR, Cardon LR, Cookson WO (2000a) A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.* 66:279-292
- Abecasis GR, Cookson WO, Cardon LR (2000b) Pedigree tests of transmission disequilibrium. *Eur. J. Hum. Genet.* 8:545-551
- Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* 62:1198-1211
- Andrew T, Hart DJ, Snieder H, de Lange M, Spector TD, MacGregor AJ (2001) Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin. Res.* 4:464-477
- Ashery-Padan R, Gruss P (2001) Pax6 lights-up the way for eye development. *Curr. Opin. Cell Biol.* 13:706-714
- Barber MJ, Cordell HJ, MacGregor AJ, Andrew T (2004) Gamma regression improves Haseman-Elston and variance components linkage analysis for sib-pairs. *Genet. Epidemiol.* 26:97-107
- Dandona R, Dandona L (2001) Refractive error blindness. *Bulletin of the World Health Organization* 79:237-243
- Dermitzakis ET, Reymond A, Scamuffa N, Ucla C, Kirkness E, Rossier C, Antonarakis SE (2003) Evolutionary discrimination of mammalian conserved non-genic sequences (CNGs). *Science* 302:1033-1035

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380:152-154

Dominguez M, Ferres-Marco D, Gutierrez-Avino FJ, Speicher SA, Beneyto M (2004) Growth and specification of the eye are controlled independently by Eyegone and Eyeless in *Drosophila melanogaster*. *Nat. Genet.* 36:31-39

Flitcroft DI (1998) Ophthalmologists should consider the causes of myopia and not simply treat its consequences. *British Journal of Ophthalmology* 82:210-211

Glaser T, Walton DS, Maas RL (1992) Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. *Nat. Genet.* 2:232-239

Hammond CJ, Snieder H, Gilbert CE, Spector TD (2001) Genes and environment in refractive error: the twin eye study. *Invest Ophthalmol. Vis. Sci.* 42:1232-1236

Hedrick PW (1987) Gametic disequilibrium measures: proceed with caution. *Genetics* 117:331-341

Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR. *Genome Res.* 6:986-994

Hung LF, Wallman J, Smith EL, III (2000) Vision-dependent changes in the choroidal thickness of macaque monkeys. *Invest Ophthalmol. Vis. Sci.* 41:1259-1269

Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd

JA (2001) Haplotype tagging for the identification of common disease genes. *Nat. Genet.* 29:233-237

Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* 423:241-254

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.* 57:439-454

Lyhne N, Sjolie AK, Kyvik KO, Green A (2001) The importance of genes and environment for ocular refraction and its determiners: a population based study among 20-45 year old twins. *Br. J. Ophthalmol.* 85:1470-1476

Martin NG, Martin PG (1975) The inheritance of scholastic abilities in a sample of twins. I. Ascertainments of the sample and diagnosis of zygosity. *Annals of Human Genetics* 39:213-218

Marzani D, Wallman J (1997) Growth of the two layers of the chick sclera is modulated reciprocally by visual conditions. *Invest Ophthalmol. Vis. Sci.* 38:1726-1739

Mertz JR, Wallman J (2000) Choroidal retinoic acid synthesis: a possible mediator between refractive error and compensatory eye growth. *Exp. Eye Res.* 70:519-527

Mutti DO, Mitchell GL, Moeschberger ML, Jones LA, Zadnik K (2002a) Parental myopia, near work, school achievement, and children's refractive error. *Invest Ophthalmol. Vis. Sci.* 43:3633-3640

Mutti DO, Semina E, Marazita M, Cooper M, Murray JC, Zadnik K (2002b) Genetic loci for pathological myopia are not associated with juvenile myopia. *Am. J. Med. Genet.* 112:355-360

Mutti DO, Zadnik K, Adams AJ (1996) Myopia. The nature versus nurture debate goes on. *Investigative Ophthalmology & Visual Science* 37:952-957

Naiglin L, Gazagne C, Dallongeville F, Thalamas C, Idder A, Rascol O, Malecaze F, Calvas P (2002) A genome wide scan for familial high myopia suggests a novel locus on chromosome 7q36. *J. Med. Genet.* 39:118-124

Neale MC (1997) *Mx: statistical modeling*. Dept of Psychiatry, Medical College of Virginia, Box 126 MCV, Richmond, VA 23298

Neale MC, Cardon LR (1992) *Methodology for genetic studies of twins and families*. Kluwer Academic Publishers, Dordrecht

Norton TT (1994) A new focus on myopia. *JAMA* 271:1362-1363

Paluru P, Ronan SM, Heon E, Devoto M, Wildenberg SC, Scavello G, Holleschau A, Makitie O, Cole WG, King RA, Young TL (2003) New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. *Invest Ophthalmol. Vis. Sci.* 44:1830-1836

Pritchard LE, Kawaguchi Y, Reed PW, Copeman JB, Davies JL, Barnett AH, Bain SC, Todd JA (1995) Analysis of the CD3 gene region and type 1 diabetes: application of fluorescence-based technology to linkage disequilibrium mapping. *Hum. Mol. Genet.* 4:197-202

Reed PW, Davies JL, Copeman JB, Bennett ST, Palmer SM, Pritchard LE, Gough SC, Kawaguchi Y, Cordell HJ, Balfour KM (1994) Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. *Nat. Genet.* 7:390-395

Rosenberg T, Klie F (1996) Current trends in newly registered blindness in Denmark. *Acta Ophthalmologica Scandinavica* 74:395-398

Saw SM, Katz J, Schein OD, Chew SJ, Chan TK (1996) Epidemiology of myopia. *Epidemiologic Reviews* 18:175-187

Schedl A, Ross A, Lee M, Engelkamp D, Rashbass P, van H, V, Hastie ND (1996) Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell* 86:71-82

Simpson TI, Price DJ (2002) Pax6; a pleiotropic player in development. *Bioessays* 24:1041-1051

Snieder H, Boomsma DI, van Doornen LJP (1997) Heritability of respiratory sinus arrhythmia: dependency on task and respiration rate. *Psychophysiology* 34:317-328

Spector TD, MacGregor AJ (2002) The St. Thomas' UK Adult Twin Registry. *Twin. Res.* 5:440-443

StataCorp (1997) *Intercooled Stata for Windows 95*. [5.0] College Station, StataCorp.

Tay MT, Au Eong KG, Ng CY, Lim MK (1992) Myopia and educational attainment in 421,116 young Singaporean males. *Ann Acad Med Singapore* 21:785-791

The Framingham Offspring Eye Study Group (1996) Familial aggregation and prevalence of myopia in the Framingham Eye Study. *Archives of Ophthalmology* 114:326-332

van Heyningen V, Williamson KA (2002) PAX6 in sensory development. *Hum. Mol. Genet.* 11:1161-1167

Wang Q, Klein BE, Klein R (1994) Refractive status in the Beaver Dam Eye Study. *Investigative Ophthalmology & Visual Science* 35:4344-4347

Weale ME, Depondt C, Macdonald SJ, Smith A, Lai PS, Shorvon SD, Wood NW, Goldstein DB (2003) Selection and evaluation of tagging SNPs in the neuronal-sodium-channel gene SCN1A: implications for linkage-disequilibrium gene mapping. *Am. J. Hum. Genet.* 73:551-565

Wilson SG, Reed PW, Bansal A, Chiano M, Lindersson M, Langdown M, Prince RL, Thompson D, Thompson E, Bailey M, Kleyn PW, Sambrook P, Shi MM, Spector TD (2003) Comparison of genome screens for two independent cohorts provides replication of suggestive linkage of bone mineral density to 3p21 and 1p36. *Am. J. Hum. Genet.* 72:144-155

Young TL, Ronan SM, Alvear AB, Wildenberg SC, Oetting WS, Atwood LD, Wilkin DJ, King RA (1998a) A second locus for familial high myopia maps to chromosome 12q. *Am. J. Hum. Genet.* 63:1419-1424

Young TL, Ronan SM, Drahozal LA, Wildenberg SC, Alvear AB, Oetting WS, Atwood LD, Wilkin DJ, King RA (1998b) Evidence that a locus for familial high myopia maps to chromosome 18p. *Am. J. Hum. Genet.* 63:109-119

Zadnik K, Satariano WA, Mutti DO, Sholtz RI, Adams AJ (1994) The effect of parental history of myopia on childrens eye size. JAMA 271:1323-1327

Table 1

Potential candidate genes under peak area (maxLOD – 1, equating to a LOD>5.1) on chromosome 11 (35cM-45cM).

Symbol	Locus	Name	Known associated eye phenotypes
BBOX1	11p14.2	butyrobetaine (gamma), 2-oxoglutarate dioxygenase	
KCNA4	11p14	Potassium voltage-gated channel	
C11orf8	11p14-p13	Chromosome 11 open reading frame 8, fetal brain protein 239	
BDNF	11p13	Brain-derived neurotrophic factor	Neuroprotectant in vertebrate eyes
TCL2	11p13	Leukemia, acute T-cell	
ELP4	11p13	PAX6 neighbor gene	?aniridia
PAX6	11p13	PAX6	Aniridia, Peters anomaly, cataract, keratitis
EVR3	11p13-p12	Familial exudative vitreoretinopathy	AD phenotype with retinal traction and detachment
WT1	11p13	Wilms tumour	Distributed through neural retina, disruption=small eyes (mice)
GA17	11p13	Dendritic cell protein	
CSTF3	11p13	Cleavage stimulation factor	
HIPK3	11p13	Homeodomain interacting protein kinase 3	

Table 2. Details of the 17 SNPs selected for PAX6 gene from public domains.

No.	Public ID	Location	Position, bp	Major allele	Minor allele	Minor allele frequency	Target sequence
1	<u>rs3026401</u>	3' UTR	31771833	A	G	0.219	TGTAAGAAATGA[A/G]TACCATATAGGG
2	rs608293	3' UTR	31772589	A	G	0.091	CCCTTAAATGGT[A/G]AACAACTGGTTT
3	<u>rs662702</u>	3' UTR	31773379	G	A	0.065	GGAAGTGCAGA[G/A]AAGGGCTATGTG
4	<u>rs1506</u>	3' UTR	31774607	T	A	0.199	TCCACTTACAGC[T/A]GGGTGTAGATCT
5	rs2071754	intron 12	31776891	A	G	0.152	CTGTGGCCAGTG[A/G]AAGGACTAGCTC
6	rs3026389	intron 12	31777838	G	C	0.152	TCCACATACAGA[G/C]CGGTATGGGGAA
7	rs667773	intron 10	31779671	C	T	0.065	TAACCTGTCCCA[C/T]CTGATTTCCAGG
8	<u>rs2239789</u>	intron 9	31780205	A	T	0.465	TTATAAGGAAAA[A/T]TGATGATTTGGC
9	<u>rs628224</u>	intron 8	31783483	C	T	0.152	CTTCTAAAGTAA[C/T]GAACTGTCTCCG
10	rs592859	intron 8	31783644	G	C	0.091	CTTAGGTTTATC[G/C]TGGGGGTGGGGG
11	rs694617	intron 4	31789827	C	A	0.909	GGGAGGGCTGCT[C/A]TCTAAGTCGGGG
12	rs1894620	5'UTR	31798825	G	C	0.095	AGGCAGGAACGA[G/C]AGGGTGAGGCCC
13	rs1806159	5'UTR	31800936	G	A	0.087	CCACTGGACAAT[G/A]TTATTTTAAAGG
14	rs1540320	5'UTR	31801752	G	A	0.087	GTCAGCCCCTCT[G/A]TCCCCCGGGCAG
15	rs1806158	5'UTR	31803097	G	C	0.087	AAGAATGCGGCC[G/C]ACAGAGCTGGGC
16	rs1806180	5'UTR	31804574	C	T	0.087	ATCTGGGATTTT[C/T]CTGTTTTCTCC
17	rs1806155	5'UTR	31805434	C	A	0.091	TCTAGGCCCAGA[C/A]TAGAGTGCCAG

Minor allele frequencies are results from this study.

Tagging SNPs are underlined.

Table 3: Model-fitting results for univariate analysis of mean spherical equivalent

Model	χ^2	$\Delta\chi^2$	df	p
ACE	13.390			
ADE	14.119	0.729		
AE	14.119	0.729	1	0.4
CE	135.425	121.306	1	<0.001

Abbreviations: χ^2 = Chi-square goodness of fit statistic; $\Delta\chi^2$ = change in χ^2 comparing submodel with full model, df = change in degrees of freedom between submodel and full model, p = probability that $\Delta\chi^2$ is zero

Table 4

Significant linkage loci (LOD>2.0) in genome-wide scan, using Genethon map chromosomal positions. Abbreviations: GLM generalised linear model, GLM_{IR} generalised linear model with iterative regression (see text)

Chromosome	Position (cM)	LOD (GLM)	p-value (GLM)	LOD (GLM _{IR})	p-value (GLM _{IR})
11p13	40	6.1	.0000001	8.0	.000000001
3q26	185	3.7	.00004	5.1	.000001
8p23	0	4.1	.00002	4.8	.000003
4q12	65	3.3	.0001	4.6	.000004
11q23-24	125	2.9	.0003	4.4	.000005

Table 5. Test for linkage, association and combined linkage and association using QTDT.

Markers:	Multipoint linkage		Association			Combined L & A
	χ^2	p-value	PS p-value	Total p-value	Permute (Robust to PS) p-value	p-value
SNP1	7.5	0.006	0.09	-	-	0.006
SNP3	7.5	0.006	0.05	-	-	0.006
SNP4	7.5	0.006	-	-	-	0.006
SNP8	7.5	0.006	-	-	-	0.006
SNP9	7.5	0.006	-	-	-	0.006
Haplotypes	6.8	0.009	0.008	-	-	0.009

Abbreviations: PS Population Stratification, L Linkage, A Association, χ^2 = Chi-square goodness of fit statistic

Figure Legends

Figure 1

Linkage disequilibria map for the 17 PAX6 SNPs genotyped (n=22) over a 32Kb region. The LD statistic plotted is r^2 . The plot orientation is 3' to 5' with SNPs 1, 9 and 16 (rs3026401, rs628224 and rs1806180) also genotyped by the Hapmap Project.

Figure 2

Scatter plots of refractive error measured as mean spherical equivalent for each twin (in diopters, D) plotted for twin 1 versus twin 2 in 226 monozygotic (MZ) and 280 dizygotic (DZ) twin pairs.

Figure 3

Genome-wide scan linkage analysis peaks using mean spherical equivalent data (n=221 pairs).

Figure 4

Linkage analysis using GLM for chromosomes 3, 4, 8 and 11 for all sib-pairs, “myopic” pairs (mean spherical equivalent ≤ 0) and “hyperopic” pairs (mean SE > 0), demonstrating significant linkage peaks (LOD > 3.2 , (Kruglyak and Lander 1995)). Marker positions are identified, as well as the position of the PAX6 gene on chromosome 11.

Figure 5

Schematic representation of PAX6 gene and position of SNPs (presented 3' to 5').

Exons are represented as vertical lines and introns as a horizontal line. Tagging SNPs used in this study are underlined. 5'UTR: 5' un-translated region, 3'UTR: 3' un-translated region.

Fig 1

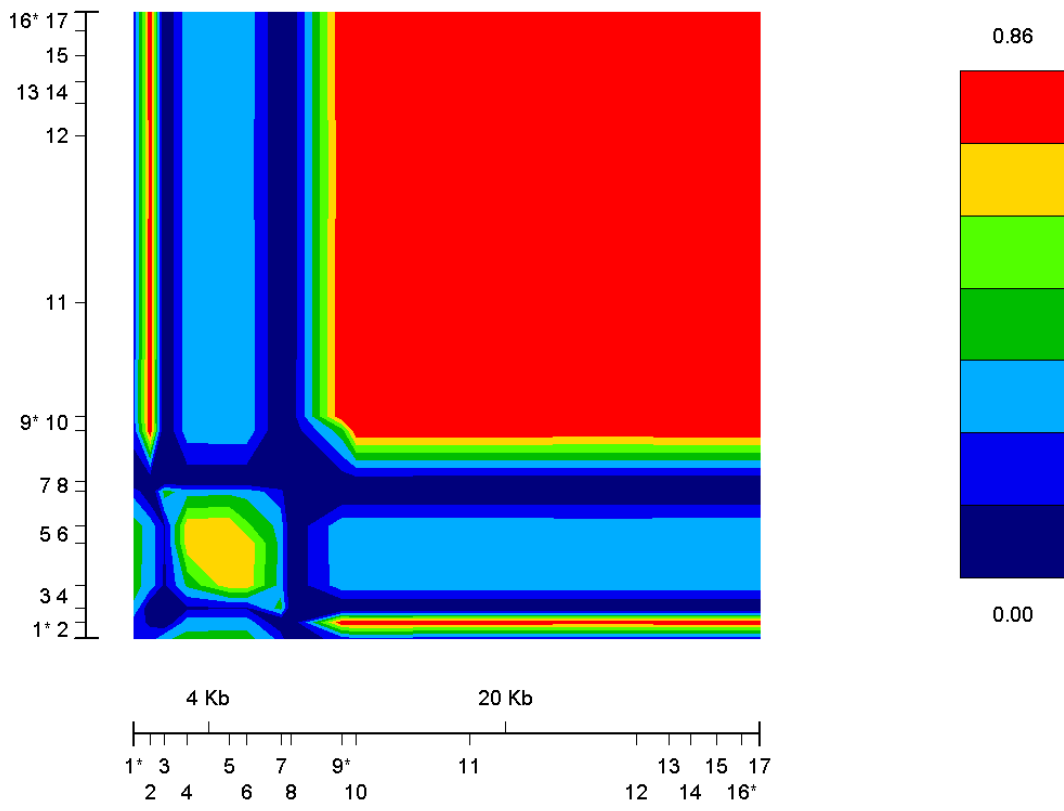


Figure 2

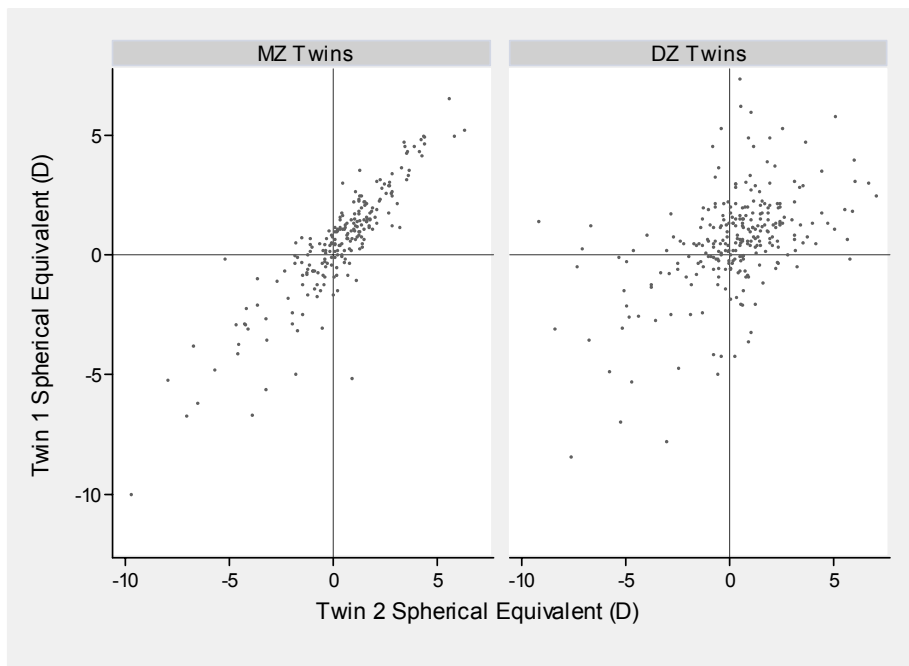


Figure 3

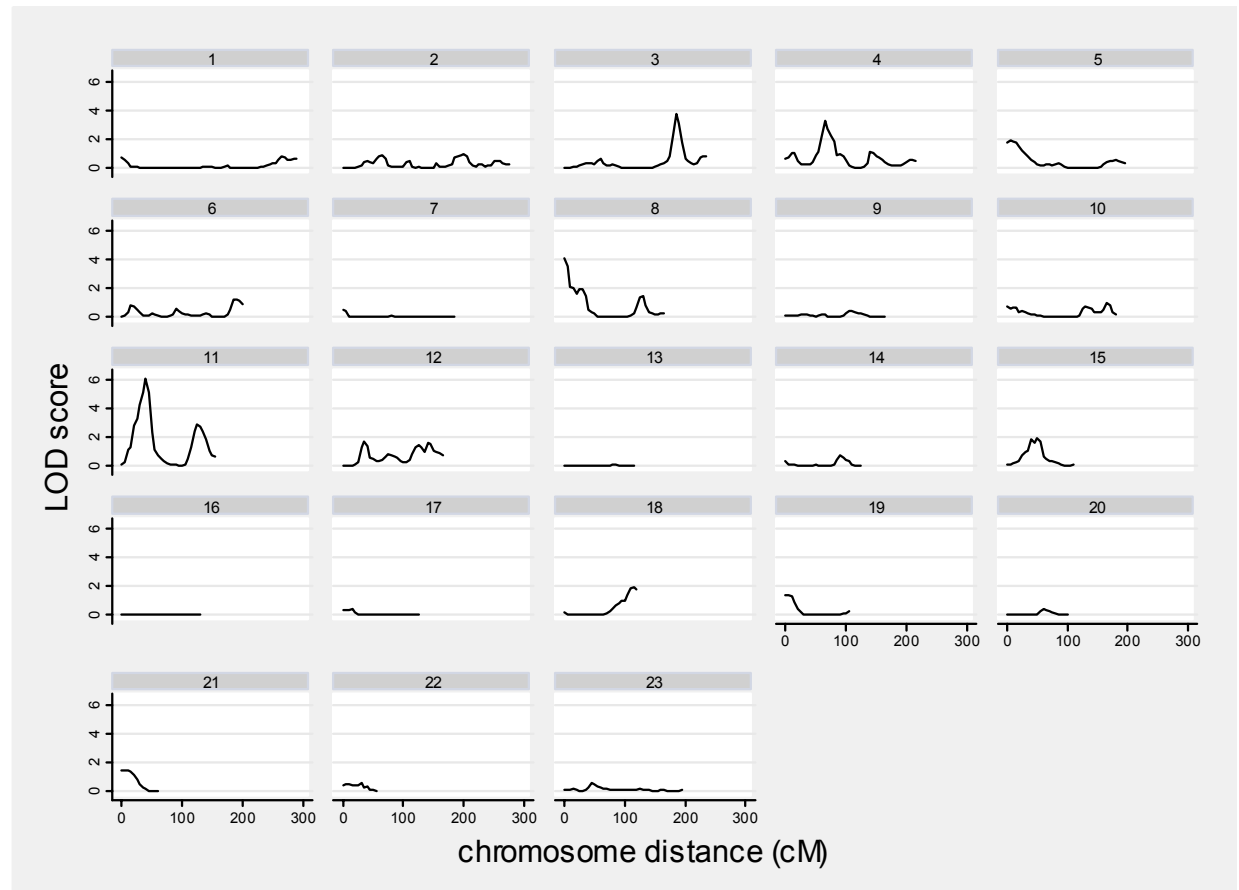


Figure 4a

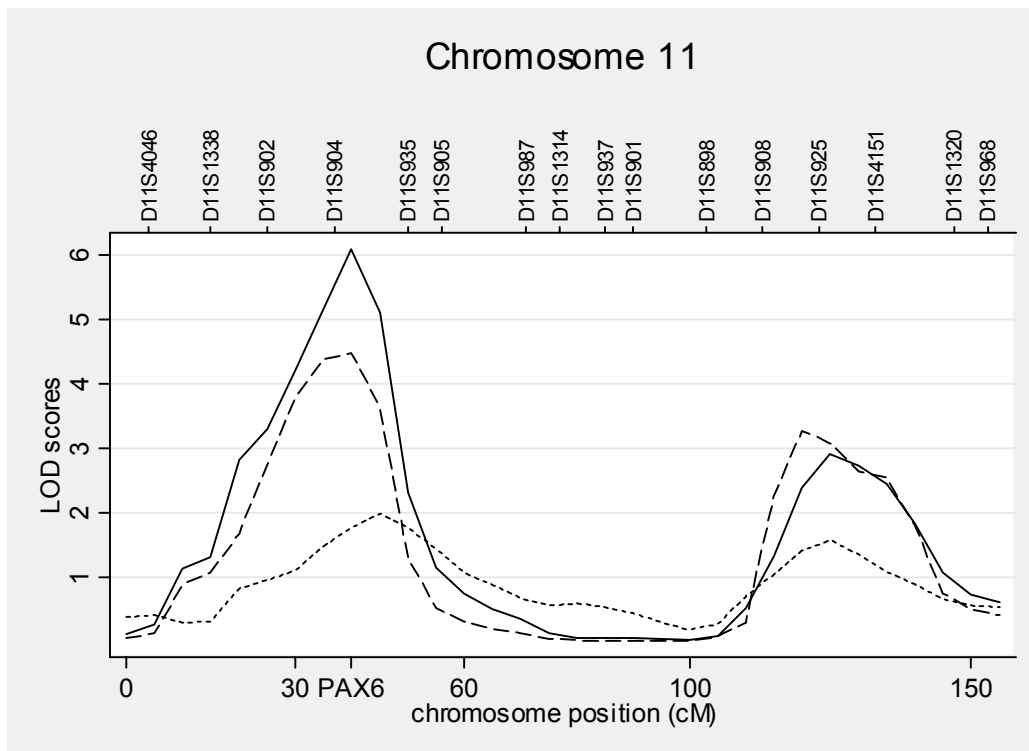


Figure 4b

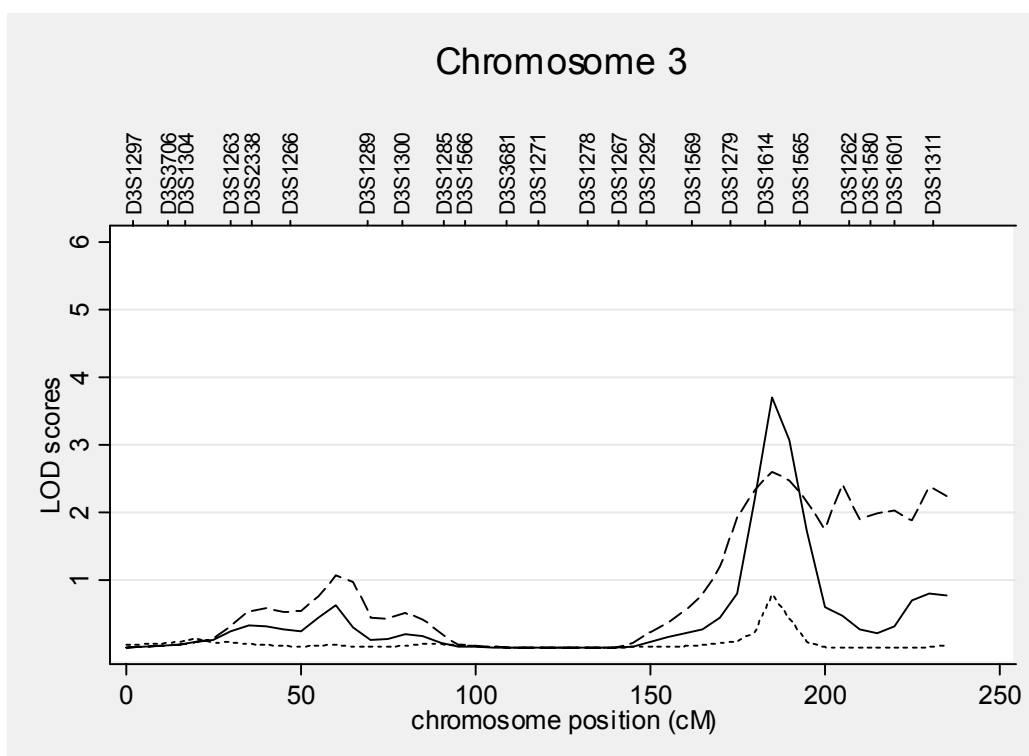


Figure 4c

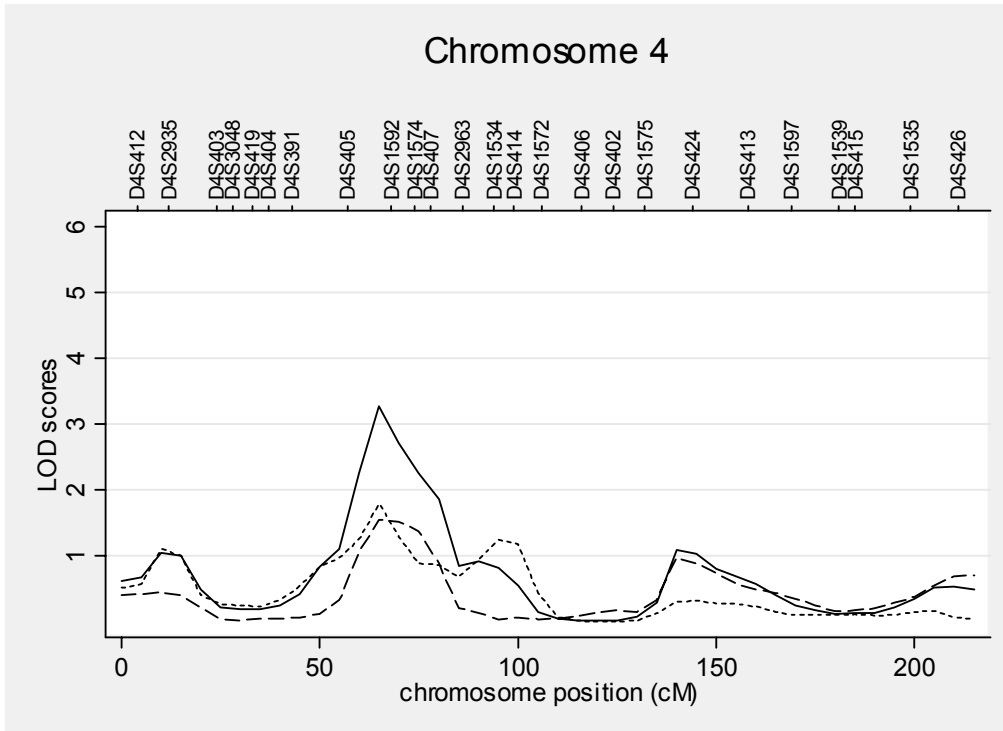


Figure 4d

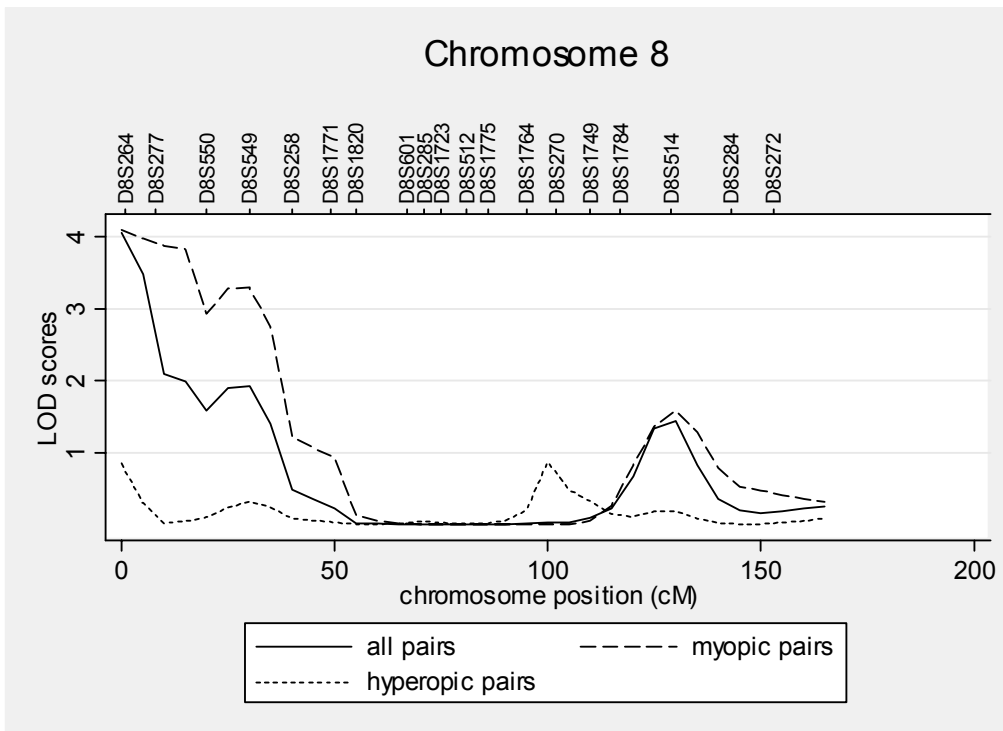


Figure 5

