Patterns of linkage disequilibrium and haplotype distribution in disease candidate genes

Ji-Rong Long1, Lan-Juan Zhao1,2, Peng-Yuan Liu1, Yan Lu1, Volodymyr Dvornyk1, Hui Shen1,2, Yong-Jun Liu1,2, Yuan-Yuan Zhang1, Dong-Hai Xiong1,2, Peng Xiao1,2 and Hong-Wen Deng*1,2,3

Address: 1Osteoporosis Research Center, Creighton University, Omaha, NE 68131, USA, 2Department of Biomedical Sciences, Creighton University, Omaha, NE 68131, USA and 3Laboratory of Molecular and Statistical Genetics, College of Life Sciences Hunan Normal University, Changsha, Hunan 410081, P. R. China

Email: Ji-Rong Long - jrlongyou@creighton.edu; Lan-Juan Zhao - zhaolanjuan@creighton.edu; Peng-Yuan Liu - pyliu@creighton.edu; Yan Lu - yanlu@creighton.edu; Volodymyr Dvornyk - Dvornyk@creighton.edu; Hui Shen - huishen@creighton.edu; Yong-Jun Liu - jun@creighton.edu; Yuan-Yuan Zhang - zhangy2@creighton.edu; Dong-Hai Xiong - donghai@creighton.edu; Peng Xiao - peng@creighton.edu; Hong-Wen Deng* - deng@creighton.edu

* Corresponding author

Abstract

Background: The adequacy of association studies for complex diseases depends critically on the existence of linkage disequilibrium (LD) between functional alleles and surrounding SNP markers.

Results: We examined the patterns of LD and haplotype distribution in eight candidate genes for osteoporosis and/or obesity using 31 SNPs in 1,873 subjects. These eight genes are apolipoprotein E (APOE), type I collagen α1 (COL1A1), estrogen receptor-α (ER-α), leptin receptor (LEPR), parathyroid hormone (PTH)/PTH-related peptide receptor type 1 (PTHR1), transforming growth factor-β1 (TGF-β1), uncoupling protein 3 (UCP3), and vitamin D (1,25-dihydroxyvitamin D3) receptor (VDR). Yin yang haplotypes, two high-frequency haplotypes composed of completely mismatching SNP alleles, were examined. To quantify LD patterns, two common measures of LD, D’ and r², were calculated for the SNPs within the genes. The haplotype distribution varied in the different genes. Yin yang haplotypes were observed only in PTHR1 and UCP3. D’ ranged from 0.020 to 1.000 with the average of 0.475, whereas the average r² was 0.158 (ranging from 0.000 to 0.883). A decay of LD was observed as the intermarker distance increased, however, there was a great difference in LD characteristics of different genes or even in different regions within gene.

Conclusion: The differences in haplotype distributions and LD patterns among the genes underscore the importance of characterizing genomic regions of interest prior to association studies.

Published: 24 May 2004
Received: 29 January 2004
Accepted: 24 May 2004

This article is available from: http://www.biomedcentral.com/1471-2156/5/11

© 2004 Long et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.
Background
Considerable attention is currently being focused on the genotyping of single nucleotide polymorphisms (SNPs) to search for complex disease susceptibility alleles. The success to detect association between marker alleles and disease critically depends on the extent of linkage disequilibrium (LD) between functional alleles and surrounding markers. Studies on some chromosomes and genome regions have shown that LD is highly variable across human genome and is structured into discrete blocks of sequences separated by hotspots of recombination and/or LD breakdown [1-6]. The LD heterogeneity across the genome is a crucial aspect for genome-wide association studies, which has been emphasized as a powerful approach for complex disease study [7-9]. Due to practical limitations of the expenses, the sample size used on the whole genome scale may not be sufficiently large. For example, the ongoing HapMap represents one of the most comprehensive and largest efforts ever attempted in biomedical research, in which a total of 270 subjects from African, Caucasian and Japanese/Chinese were recruited. However, sample size is an important factor influencing the accuracy of LD evaluation. Small sample size will bias the accuracy of LD estimation [2,10].

A limited but hypothesis-driven approach for association studies is to focus on the polymorphisms located in or around the candidate genes known to be of potentially functional importance for complex traits (based on their cellular and molecular studies). LD patterns in specific candidate genes need be empirically assessed to facilitate efficient experimental designs and execution of association studies [1,11]. So far, only limited information is available about the LD patterns and haplotype characteristics in candidate genes for complex diseases [12,13].

The purpose of this study was to determine the LD and haplotype characteristics in eight candidate genes for the two common complex disorders, osteoporosis and obesity. These eight genes are apolipoprotein E (APOE), type I collagen α1 (COL1A1), estrogen receptor-α (ER-α), leptin receptor (LEPR), parathyroid hormone (PTH)/PTH-related peptide receptor type 1 (PTHR1), transforming growth factor-β1 (TGF-β1), uncoupling protein 3 (UCP3), and vitamin D (1,25-dihydroxyvitamin D₃) receptor (VDR).

Results
Haplotype analysis
Allele frequencies of the 31 SNPs are shown in Table 1. The information for the main haplotypes whose frequencies were higher than 1% [14], is presented in Table 2. It is clear that the number of observed haplotypes was smaller than that of haplotypes expected. The number of observed haplotypes varied at different loci. In both the PTHRI gene and the VDR gene, four SNPs were studied. However, seven and twelve haplotypes, respectively, were observed as compared to the expected 16 haplotypes in each gene.

The yin yang haplotypes are defined as two high-frequency haplotypes composed of completely mismatching SNP alleles, i.e., nucleotides differ at every SNP in the haplotype pair [15]. We detected such haplotype phenomenon in two of the eight studied genes (Table 2). In the PTHRI gene, one pair of yin yang haplotypes, GTAA and ACGG (in the order of SNP20-21-22-23), was observed with frequencies of 34.2% and 52.6%, respectively. A pair of yin yang haplotypes, TCT and CTC (in the order of SNP28-30-31), was also observed in the UCP3 gene with frequencies of 23.5% and 40.5%, respectively.

LD analysis
In this study, we used two most common LD measures, D' and r², to determine the extent of LD in the genes (Figure 1). The average of D' values within the eight genes was 0.475 ± 0.325 (±SD) with the range of 0.020–1.000, whereas the r² was 0.158 ± 0.255 (±SD), which ranged from 0.000 to 0.883. Until now, there was no unified standard for meaningful LD between a marker and a functional mutation. Although we do not know whether any of our studied SNPs is functional for osteoporosis and/or obesity, the knowledge of LD patterns revealed through them is useful in inferring effective intermarker distance. When a criterion of D' > 0.33 [4,16] was used as a meaningful LD, 58.8% (30/51) of the SNP pairs met this criterion with the intermarker physical distance ranging from 138 bp (SNP3-SNP4) to ~122 kb (SNP12-SNP14). If we considered D' = 0.7 as a limit for useful LD [17], only 33.3% (17/51 SNP pairs) were "useful". If a value of r² > 0.1 was applied as a criterion for meaningful LD [16,18], 31.4% (16/51 SNP pairs) would be considered in "useful" LD. When we used both D' > 0.33 and r² > 0.1 as a criterion for meaningful LD [4,16,18], these 16 pairs still fell below this threshold. Such "useful" LD was relatively common in a pair-wise physical distance smaller than 8.6 kb, though it was detectable at the distance of 83 kb (SNP13-SNP14).

From Figure 1, the distribution of LD versus physical distance was irregular; however, a decay of LD with increasing distance was still evident. The Spearman correlation coefficient between D' and the distance was -0.70 (P < 0.01). The correlation coefficient was -0.49 (P < 0.01) when LD was quantified by r². Similar relationship was observed between LD measures and ln(distance), in which the correlation coefficient was -0.68 and -0.47 for D' and r², respectively. The D'/ln(distance) was calculated for pair-wise SNPs within each gene. The average of D'/ln(distance) varied a lot among these eight genes with the
Table 1: Information about the 37 SNPs in the eight genes

| SNP ID | Gene | dbSNP accession | polymorphism | Domain | Distance (bp) | Frequency (%) | Forward Primer | Reverse Primer |
|--------|------|-----------------|--------------|--------|---------------|---------------|----------------|----------------|
| 1 | APOE | ss12568587 | G-C | 5' UTR | 357 | 35.7 | TCCCCAGGAGCCGCTGGA | CCCCCAGGACCGACCC |
| 2 | APOE | ss12568609 | G-A | Intron 2 | 1277 | 39.9 | CCCCCAGGAGCCGCTGGA | AATCTCTGGGCTCAAGTG |
| 3 | APOE | ss12568607 | T-C | Exon 4 | 1497 | 14.9 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 4 | APOE | ss12568612 | C-T | Exon 4 | 138 | 8.7 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 5 | COLIA1 | ss12568606 | G-T | 5' UTR | 357 | 15.4 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 6 | COLIA1 | ss12568597 | G-T | Intron 2 | 1277 | 18.8 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 7 | COLIA1 | ss12568598 | G-T | Exon 8 | 2365 | 0.0 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 8 | COLIA1 | ss12568584 | A-G | Exon 4 | 9862 | 19.8 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 9 | ER-α | ss12568579 | G-A | Exon 1 | 156 | 48.4 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 10 | ER-α | ss12568596 | T-C | Intron 1 | 34258 | 44.9 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 11 | ER-α | ss12568619 | G-C | Intron 3 | 66110 | 25.7 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 12 | ER-α | ss12568618 | G-C | Exon 4 | 39074 | 9.8 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 13 | ER-α | ss12568585 | T-C | Intron 4 | 39074 | 21.8 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 14 | ER-α | ss12568605 | G-A | Exon 6 | 83068 | 11.9 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 15 | ER-α | ss12568617 | A-G | Exon 8 | 32431 | 20.3 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 16 | LEPR | ss12568615 | A-G | Exon 4 | 290 | 15.4 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 17 | LEPR | ss12568604 | G-C | Exon 14 | 3013 | 18.2 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 18 | LEPR | ss12568614 | A-G | Exon 20 | 62803 | 35.3 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 19 | PTH1R | ss12568589 | G-A | Intron 1 | 1 | 0.8 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 20 | PTH1R | ss12568592 | A-G | Intron 2 | 4250 | 39.8 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 21 | PTH1R | ss12568591 | C-T | Intron 8 | 5436 | 37.4 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 22 | PTH1R | ss12568588 | G-A | Intron 10 | 1912 | 39.7 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 23 | PTH1R | ss12568590 | G-A | Exon 13 | 1246 | 38.0 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 24 | TGF-β1 | ss12568613 | C-T | 5' UTR | 2132 | 31.1 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 25 | TGF-β1 | ss12568603 | A-C | Intron 4 | 12354 | 2.1 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 26 | TGF-β1 | ss12568593 | C-T | Exon 5 | 83 | 0.8 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 27 | TGF-β1 | ss12568602 | C-T | Intron 5 | 9644 | 27.4 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 28 | UCP3 | ss12568601 | C-T | 5' UTR | 2122 | 26.5 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
| 29 | UCP3 | ss12568600 | G-A | Exon 2 | 2122 | 0.0 | CCCCCGACGCGCTTCA | AATCTCTGGGCTCAAGTG |
The highest value of 0.085 in PTHR1 and the lowest value of 0.001 in TGF-β1. The Spearman correlation coefficient between the two LD measures was 0.81 ($P < 0.01$).

**Discussion**

Determining the haplotype and LD characteristics of candidate genes and/or genomic regions has important implications for strategies of complex diseases gene mapping [19]. In the present study, we illustrated the haplotype distribution and LD patterns in eight candidate genes for two complex diseases, obesity and osteoporosis, in 1873 subjects from 405 Caucasian nuclear families.

In this study, the yin yang haplotype pairs, mismatching at every SNP, were observed in two genes, PTHR1 and UCP3, but not in the other six. There are several possible reasons for such differences among genes. Other things being the same, lower LD levels result in smaller yin yang haplotype size, the genomic span per yin yang haplotype region [15]. The different SNP density and LD patterns among genes may be partly responsible for the discrepancies about yin yang haplotypes among genes. Lower minor allele frequencies in some genes may be another reason since minor allele frequency $\geq 10\%$ is a threshold for yin yang haplotype analysis [15]. Moreover, yin yang haplotypes were affected by the factors that might be different among genes, such as the mutation and recombination rates [15].

So far, LD was inferred either by using unrelated random sample or family data. The LD values evaluated from these two different data sets were correlated, but in some instances they were quite different. Family subjects contain more information than unrelated random individuals, and the inference of LD using family data is more accurate [4]. Sample size is another important factor.
influencing the accuracy of LD evaluation. Generally, larger sample size can minimize sampling error and produce more accurate evaluation of LD [20]. In this study, haplotypes were reconstructed based on the nuclear families in a larger sample rather than unrelated individuals as the other studies [12,18,21,22]. This increased the precision of LD evaluation.

In the present study, two most common measurements of LD, D' and r², were used, both measurements varying between 0 and 1. According to their calculating formula [23], they are positively correlated, as observed in this study, but sometimes they are quite different. For example, D' value for SNP6-SNP8 was 1.000, while the value for r² was only 0.004. Such a discrepancy between the two measures was also observed by Tiret et al [17]. One possible explanation for the discrepancy is that the rare allele frequency of SNP8 was very low, only 1.9%, and r² is more sensitive to allele frequencies than D' [23].

Although a trend of decay of LD with increasing physical distance was observed in the present study, there was a great variation in LD with distance. For example, the physical distances for SNP33-SNP34 in VDR and SNP20-SNP21 in PTHR1 were both about 5 kb, however, the D' values were 0.062 and 0.964, and the r² were 0.004 and 0.826, respectively. In the ER-α gene, the physical distance of SNP14-SNP15 was about 2.6 times that of SNP13-SNP14, while the D' values were 0.675 and 0.028, and r² were 0.599 and 0.019, respectively. Such irregularity has been observed in the other studies [6,17,18,24] and is expected, since LD is also affected by other factors, such as natural selection, the rates of recombination and mutation, and gene conversion [25].

Several limitations of this study should be acknowledged. Compared with some recent analyses [4,14,26], we only selected a limited number of SNPs in each gene, which may bring some difficulty to define the LD characteristics. Though 15-kb chromosome-wide resolution for LD pattern and haplotype structure study is acceptable but not ideal [27]. In addition, the yin yang haplotype analysis might be affected by the limited SNPs and genes examined. The SNPs were chosen based on the information from available literature and public databases rather than evenly distributed within gene. Another limitation is that this study was restricted to Caucasians of European descent, which limited the ability to perform cross-population comparisons for yin yang haplotype coverage and haplotype diversities etc. Since genetic diversity and LD extent differ between populations, studying contrasted populations is informative. However, if the purpose is to provide guidelines for the design of association studies in a given population, information drawn from that population should be preferable [5], especially because Caucasians are more susceptible to osteoporosis and obesity compared with other populations [28,29].

In conclusion, the haplotype distribution was different among the eight studied candidate genes for osteoporosis and/or obesity. The LD characteristics were different between genes or even at different regions within gene, though LD decays with the increase of physical distance. The differences in haplotype distributions and LD patterns among genes underscore the importance of characterizing each locus of interest prior to association studies.
Methods

Subjects

The study subjects came from the ongoing genetic studies of osteoporosis and obesity that have been approved by the Creighton University Institutional Review Board. All subjects were Caucasians of European origin. We have recruited 405 nuclear families composed of 1,873 healthy individuals, including 840 parents, 744 daughters and 389 sons. All individuals signed informed-consent documents before entering the project.

Genotyping

After searching public SNP databases, such as dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) and available literature, 37 SNPs in the eight candidate genes for obesity and/or osteoporosis were chosen. The selection criteria were as follows: 1) functional relevance and importance; 2) position in or around the gene; and 3) their use in previous genetic studies. Information about the 37 SNPs is presented in Table 1. The genotyping procedure for all SNPs was similar, involving polymerase chain reaction (PCR) and invader assay reaction (Third Wave Technology, Madison, WI). PCR was performed in a 10 µl reaction mixture. The sequences of the PCR primers for all 37 SNPs are presented in Table 1. After amplification, an invader reaction (Third Wave Technology, Madison, WI). PCR was performed in a 10 µl reaction mixture. The sequences of the PCR primers for all 37 SNPs are presented in Table 1. After amplification, an invader reaction was performed in a 7.5 µl reaction volume. The genotype for every sample was called according to the ratio of the fluorescence intensity of the two dyes, which was read using a CytoFluor 4000 multi-well plate reader (Applied Biosystems, Foster City, CA). Program PedCheck (http://watson.hgen.pitt.edu/register/soft_doc.html) was employed to verify Mendelian inheritance of all marker alleles within each family [30].

We initially genotyped all 37 SNPs in a random sample of 190 to 380 subjects. The minor allele frequency of six SNPs was found less than 1.0%. Because the precision and variance of LD estimates suffer from low allele frequencies [6], these six SNPs were excluded from further genotyping and data analysis. Finally, 31 SNPs were analyzed in this study. Among them, 14 were located in exons, 12 in introns, and 5 in untranslated regions. SNP23 in the TGF-β1 gene is an insertion/deletion polymorphism of a cytosine (+/- C) and the other 30 SNPs are nucleotide substitutions, including 25 transitions and 5 transversions. These 31 SNPs totally spanned ~451 kb. The average inter-marker distance was 63 kb (ranging from 138 bp to 291 kb).

Statistical analysis

Allele frequencies of each SNP were estimated in 1,873 subjects using a maximum likelihood method implemented in the program SOLAR which is available at http://www.sfrb.org/sfrb/public/software/solar. The haplotypes of each gene for all subjects were reconstructed using the program Genehunter version 2.1 (http://www.hgmp.mrc.ac.uk/Registered/Option/genehunter.html). Genehunter extracts complete multipoint inheritance information to infer maximum likelihood haplotypes for all individuals in nuclear families [31]. Haplotype frequencies were estimated from the unrelated subjects (parents of the nuclear families).

Two normalized measures, $D'$ and $r^2$, were used to characterize the LD patterns within the studied candidate gene. Pair-wise $D'$ values are computed as: $D' = D/D_{\text{max}}$. $D_{\text{max}}$ is the maximum $D$ value occurring within the given set of markers. $r^2$ is defined as $r^2 = (f_{11} - f_{22})^2 / (f_{11} + f_{12} + f_{21} + f_{22})$, where $f$s are the sample estimates of haplotype frequencies [23]. An exponential decay function was employed to fit the relationship between $D'$ and $r^2$. $D'$ and $r^2$ both as index of LD, correlation coefficient between them was calculated using SPSS software.

Author’s contribution

JRL performed SNP genotyping, combined the results of analysis and drafted the manuscript. LJZ performed SNP genotyping and data analysis. PYL and YL performed statistical analysis. LD contributed to the manuscript preparation. HS, YJL, YZZ, DHX and PX performed SNP genotyping. HWD conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

Acknowledgements

The investigators were partially supported by grants from Health Future Foundation of the USA, the National Institute of Health, the State of Nebraska Cancer and Smoking Related Disease Research Program, the State of Nebraska Tobacco Settlement Fund, US Department of Energy, Chinese National Science Foundation, and the Ministry of Education of P. R. China.

References

1. Ardlie KG, Kruglyak L, Seielstad M: Patterns of linkage disequilibrium in the human genome. Nat Rev Genet. 2002, 3:299-309.
2. Reich DE, Cargill M, Bull S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kyouymjian R, Farhadian SF, Ward R, Lander ES: Linkage disequilibrium in the human genome. Nature 2001, 411:99-204.
3. Wall JD, Pritchard JK: Haplotype blocks and linkage disequilibrium in the human genome. Nat Rev Genet 2003, 5:8-97.
4. Moffatt MF, Traherne JA, Abecasis GR, Cookson WO: Single nucleotide polymorphism and linkage disequilibrium within the TCR alpha/delta locus. Hum Mol Genet 2000, 9:1011-1019.
5. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer C], Chew A, Han JH, Duan J, Carr JL, Lee MS, Koshiyama B, Kumar AM, Zhang G, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaber SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Ruano G, Vovis GF: Haplotype variation and linkage disequilibrium in 313 human genes. Science 2001, 293:489-493.
6. Abecasis GR, Noguchi E, Heinzmann A, Traherne JA, Bhattacharyya S, Leaves N, Anderson GG, Zhang Y, Lench NJ, Carey A, Cardon LR, Moffatt MF, Cookson WO: Extent and distribution of linkage
disequilibrium in three genomic regions. Am J Hum Genet 2001, 68:191-197.
7. Tusie N, Mangahas K: The future of genetic studies of complex human diseases. Science 1996, 273:1516-1517.
8. Weiss KM, Clark AG: Linkage disequilibrium and the mapping of complex human traits. Trends Genet 2002, 18:19-24.
9. Jorde LB: Linkage disequilibrium and the search for complex disease genes. Genome Res 2000, 10:1435-1444.
10. Teare MD, Dunning AM, Durocher F, Rennart G, Easton DF: Sampling distribution of summary linkage disequilibrium measures. Am Hum Genet 2002, 66:223-233.
11. Stephens M, Smith NJ, Donnelly P: A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001, 68:978-989.
12. Bonnen PE, Wang PJ, Kimmel M, Chakraborty R, Nelson DL: Haplotype and linkage disequilibrium architecture for human cancer-associated genes. Genome Res 2002, 12:1846-1853.
13. Zhu X, Yan D, Cooper RS, Luke A, Ikeda MA, Chang YP, Weder A, Chakravarti A: Linkage disequilibrium and haplotype diversity in the genes of the renin-angiotensin system: findings from the family blood pressure program. Genome Res 2003, 13:173-181.
14. Groucher Pj, Maschereti S, Hampe J, Huse K, Frenzel H, Stoll M, Wu TK, Nikolau S, Yang SK, Kimmich J, Schreiber S: Haplotype structure and association to Crohn's disease of CARD15 mutations in two ethnically divergent populations. Eur J Hum Genet 2003, 11:6-16.
15. Zhang J, Rowe WL, Clark AG, Buettow KH: Genomewide Distribution of High-Frequency, Completely Mismatching SNP Haplotype Pairs Observed To Be Common across Human Populations. Am J Hum Genet 2003, 73:1073-1081.
16. Kruglyak L: Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 1999, 22:139-144.
17. Tintle N, Poirier O, Niclous V, Herrmann SM, Perret C, Aupperle RJ, Mooney BL, Kouchoukos N, Nowell PC, Frenzel H, Stoll M, Wu TK, Nikolau S, Yang SK, Kimmich J, Schreiber S: Haplotype structure and association to Crohn's disease of CARD15 mutations in two ethnically divergent populations. Eur J Hum Genet 2003, 11:6-16.
18. Nakajima T, Jorde LB, Ishigami T, Umemura S, Emi M, Lalouel JM, Inoue I: Nucleotide diversity and haplotype structure of the human angiotensinogen gene in two populations. Am J Hum Genet 2002, 70:108-123.
19. Weale ME, Depondt C, Macdonald SJ, Smith A, Lai PS, Shorvon SD, Wood NW, Goldstein DB: Selection and evaluation of tagging SNPs in the neuronal-sodium-channel gene SCN1A: implications for linkage-disequilibrium gene mapping. Am J Hum Genet 2003, 73:551-565.
20. Guo SW: Linkage disequilibrium measures for fine-scale mapping: a comparison. Hum Hered 1997, 47:301-314.
21. Peterson RJ, Goldman D, Long JC: Effects of worldwide population subdivision on ALDH2 linkage disequilibrium. Genom Res 1999, 9:844-852.
22. Albagha OM, Tasker PN, McGuigan FE, Reid DM, Ralston SH: Linkage disequilibrium between polymorphisms in the human TNFRSF1B gene and their association with bone mass in perimenopausal women. Hum Mol Genet 2002, 11:2289-2295.
23. Devlin B, Risch N: A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 1995, 29:311-322.
24. Taillon-Miller P, Baur-Sardina I, Saccone NL, Putzel J, Latinen T, Cao A, Kere J, Pilia G, Rice JP, Kwock PY: Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq26 and Xq28. Nat Genet 2000, 25:324-328.
25. Ardile K, Liu-Cordero SN, Eberle MA, Daly M, Barrett J, Winchester E, Landers ES, Kruglyak L: Lower-than-expected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion. Am J Hum Genet 2001, 69:582-589.
26. Zabetian CP, Buchbaum SG, Elston RC, Kohnke MD, Anderson GM, Gelernter J, Cubells JF: The structure of linkage disequilibrium at the DBH locus strongly influences the magnitude of association between diallelic markers and plasma dopamine beta-hydroxylase activity. Am J Hum Genet 2003, 72:1389-1400.
27. Dawson E, Abecasis GR, Bumpstead S, Chen Y, Hunt S, Beare DM, Pabail J, Dibling T, Tinsley E, Kirby S, Carter D, Papaspyridonos M, Livingstone S, Ganske R, Lohmusar A, Eberle MA, Daly M, Barrett J, Winchester E, Lander ES, Kruglyak L: Lower-than-expected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion. Am J Hum Genet 2001, 69:582-589.
28. Liu YZ, Liu YJ, Recker RR, Deng HW: Molecular studies of identification of genes for osteoporosis: the 2002 update. J Endocrinol 2003, 17:147-196.
29. Chagnon YC, Rankinen T, Snyder EE, Weisnagel SJ, Perusse L, Bougu L, Coadracter C: The human obesity gene map: the 2002 update. Obes Res 2003, 11:313-367.
30. O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype inaccuracies in linkage analysis. Am J Hum Genet 1998, 63:259-266.
31. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996, 58:1347-1363.
32. Lewontin RC: The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 1964, 49:49-67.