ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

GENOMICS

Genomics 82 (2003) 194–217

www.elsevier.com/locate/ygeno

Nucleotide variation, haplotype structure, and association with end-stage renal disease of the human interleukin-1 gene cluster

Jeannette T. Bensen,[a,*] Carl D. Langefeld,[b] Gregory A. Hawkins,[a] Linda E. Green,[b] Josyf C. Mychaleckyj,[a] Catherine S. Brewer,[c] Deborah S. Kiger,[a] Scott M. Binford,[a] Carla J.[a] Coleicigno,[c] Dax C. Allred,[c] Barry I. Freedman,[d] and Donald W. Bowden[a,c,d]

[a] *The Center for Human Genomics, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA*
[b] *Department of Public Health Sciences, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA*
[c] *Department of Biochemistry, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA*
[d] *Department of Internal Medicine, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA*

Received 11 December 2002; accepted 14 April 2003

Abstract

A dense gene-based SNP map was constructed across a 360-kb region containing the interleukin-1 gene cluster (*IL1A*, *IL1B*, and *IL1RN*), focusing on *IL1RN*. In total, 95 polymorphisms were confirmed or identified primarily by direct sequencing. Polymorphisms were precisely mapped to completed BAC and genomic sequences spanning this region. The polymorphisms were typed in 443 case–control subjects from Caucasian and African American groups. Consecutive pair-wise marker linkage disequilibrium was not strictly correlated with distance and ranged from $D' = 0.0079$ to 1.000 and $D' = 0.0521$ to 1.0000 in Caucasians and African Americans, respectively. Single markers and haplotypes in *IL1* cluster genes were evaluated for association with end-stage renal disease (ESRD). Eleven SNPs show some evidence of association with ESRD, with the strongest associations in two *IL1A* variants, one SNP, rs1516792-3, in intron 5 ($p = 0.0015$) and a 4-bp insertion/deletion within the 3′UTR, rs16347-2 ($p = 0.0024$), among African Americans with non-T2DM-associated ESRD.

Keywords: Interleukin-1; IL-1; End-stage renal disease; ESRD; Single nucleotide polymorphism; SNP; Linkage disequilibrium; Haplotype

End-stage renal disease (ESRD) with underlying glomerulosclerosis and fibrosis is a major cause of morbidity and mortality among individuals with diabetes, hypertension, and glomerulonephritis [1]. The incidence of ESRD has increased approximately 50% among all racial groups from 1990 to 2000 [1]. While monogenic forms of renal disease exist, ESRD in the general population has a heterogeneous etiology likely resulting from both genetic and environmental factors. African Americans are at 4.5-fold higher risk than Caucasians [1] to develop ESRD. Familial clustering of ESRD with its associated comorbid conditions suggests that possible genetic factors are key in its development.

Multiple lines of evidence support an inflammatory component characterized by glomerular basement membrane thickening and mesangial expansion in the development of nephropathy [2]. Neutral proteases known as matrix metalloproteinases (MMPs) and their inhibitors are essential in maintaining the normal extracellular matrix (ECM) turnover of connective tissue surrounding the mesangial cells. Specifically, changes in the expression of MMP-9 lead to dysregulation of normal ECM turnover [3]. Inflammatory mediators such as IL-1β, TNF-α, and nitric oxide synthetase have been shown to affect the expression of MMP-9 in rat mesangial cells [4].

In humans, polymorphisms in the *IL1* gene cluster on chromosome 2, especially interleukin receptor antagonist (*IL1RN*), have been reported to be associated with diabetic nephropathy [5–8]. The *IL1* gene cluster located on chro-

* Corresponding author. Fax: +1-336-713-7566.
E-mail address: jbensen@wfubmc.edu (J.T. Bensen).

0888-7543/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0888-7543(03)00123-X

mosome 2q12–q14 contains several pro- and anti-inflammatory cytokine genes that are expressed in both physiological and pathological conditions. These cytokines are regulated at many levels and play a key role in the host's response to microbial infection, as well as extracellular matrix turnover and bone resorption. At the cellular level, membrane-bound receptors, soluble receptors, decoy receptors, IL-1 receptor antagonists, and natural anti-IL-1 antibodies exist to maintain the balance between pro- and anti-inflammatory responses within the cell. The association of *IL1* gene cluster polymorphisms with inflammatory diseases and the link between IL-1 and MMP in renal matrix maintenance have made these genes attractive candidates for evaluation in patients with ESRD. We have utilized an experimentally verified gene-based SNP map spanning the 360-kb *IL1* gene cluster region between *IL1A* and *IL1RN* to systematically assess the role the these genes may play in ESRD. This effort incorporated SNP discovery, evaluation of linkage disequilibrium (LD), and association studies with ESRD.

Results

Polymorphisms within the IL1 gene cluster

A dense SNP map was constructed across the 360-kb region containing the interleukin-1 gene cluster [*IL1A*, *IL1B*, and *IL1RN* (receptor antagonist)], focusing on *IL1RN*. In total, 95 polymorphisms, 6 of which were non-SNP variants (Table 1), were confirmed (NCBI dbSNPs) or identified (primarily by direct sequencing): in *IL1A*, 13 SNPs and 1 insertion/deletion (4 bp); in *IL1B*, 5 SNPs, 1 intragenic tetranucleotide repeat; in *IL1F10*, 6 SNPs; and within *IL1RN*, 1 VNTR (86 bp), 3 single-base insertion/deletions, and 65 SNPs. Approximately 27 segments of 600 bp were sequenced directly in 443 chromosomes yielding a total of 162,000 bp. Within this 162 kb we identified 82 high-quality polymorphisms (PolyPhred Rank 1, SNP frequency approximately 1 per 300 bp) consisting of 79 SNPs and 3 single-nucleotide insertion/deletion variants. Overall, 26 SNPs distributed across the *IL1* genes had minor allele frequencies greater than 0.20 (Appendixes 2A–2E). Transition substitutions [66/90 (73%)] were almost three times more common than transversion substitutions [24/90 (27%)]. The vast majority of variants assessed were located in noncoding regions, with 10/95 (10%) within coding regions. Of those in coding regions, three conferred an amino acid change (Table 2).

A total of 170 Caucasian and 273 African American case–control subjects were genotyped for 88 variants with seven SNPs (Table 1) typed in a subset of the study population subjects (total variants typed: 95). Twenty-six of the polymorphisms in African American groups were not polymorphic in either Caucasian sample group (Appendixes 2A–2E), and one (rs315935-2) was polymorphic (<1%) in Caucasian cases and not in African Americans. Of the 67 polymorphic loci, 64 were usefully polymorphic in both ethnic groups. Minor allele frequencies varied greatly between racial groups for any given SNP. For example, the SNP named exon 5 at 4336 SNP-1, actually located in the fourth intron of *IL1B*, has minor allele frequencies of 39 and 16.3% in Caucasian and African American controls, respectively ($p \leq 0.001$). All polymorphisms were in Hardy–Weinberg equilibrium (data not shown).

Linkage disequilibrium

Consecutive marker pair (1–2, 2–3, 3–4, 4–5, etc.) D' for all markers ranged from 0.0079 to 1.0000 (data not shown). In general, marker pairs that were not in complete LD ($D' = 1.00$) exhibited D' values between 0.40 and 0.95. Across this region LD was not strictly correlated with distance; however some trends were evident. Across several large intergenic physical distances of approximately 47 and 143 kb (*IL1A* SNP −889 to *IL1B* exon 5 at 4336 SNP-2 and *IL1B* SNP −511 to *IL1RN* Gaatp33330), LD predictably decreased, with the larger intergenic distance having the lowest D'. A closer examination of LD across shorter distances ranging from 5 bp to 3.2 kb, among markers of varying heterozygosity, failed to identify a consistent fall in D' with distance. If, rather than evaluating consecutive marker pairs, D' statistics are calculated for cumulative marker pairs (1–2, 1–3, 1–4, 1–5, 1–6, . . ., 1–*n;* 2–1, 2–3, 2–4, . . ., 2–*n;* etc.) for all markers, D' again clearly falls at very predictable landmarks between genes, and this observation is consistent across all sample groups.

Ethnic variation

Comparing ethnic groups, approximately twice the number of marker pairs with an arbitrary criterion of $D' < 0.95$ are observed among African Americans ($N = 34$) than among Caucasians ($N = 15$), suggesting more limited LD in the African American sample groups (data not shown). The majority of marker-to-marker pairs with reduced D' are in noncoding regions such as the promoter, 5′UTR, intron, 3′UTR, and flanking genomic sequence. Of those marker pairs common to both ethnic groups, several displayed large differences in D' (data not shown). The contrast of ethnic patterns of LD by distance using two measures of LD, D and D', is shown in Fig. 1A through 1D. LD values are calculated for all possible pair-wise marker combinations (1–2, 1–3, 1–4, . . ., 1–96; 2–1, 2–3, 2–4, . . ., 2–96) and plotted against the known distance between the marker pair. A comparison of these measures shows a poor correlation between D' and distance across the *IL1* gene cluster, while D clearly decreases with increasing distance. In African Americans D decays quite rapidly at distances under 50 kb within the *IL1* gene cluster, while LD decays more slowly in

Table 1
Location (bp) of variants within BAC and *IL1* gene cluster

SNP Name	Primer set range	BAC 1[a]	BAC 2[a]	BAC 3[a]	Absolute position
rs16347-3	X03833 11309. . .11490	11423			548
rs16347-2		11419			552
rs16347-1		11409			562
rs1516792-3	X03833 7208. . .7920	7635			4336
rs1516792-2[b]		7572			4399
rs1516792-1		7496			4475
rs17561-3	X03833 5949. . .6642	6282			5689
rs17561-2		6166			5805
rs17561-1		6153			5818
rs20540	X03833 3996. . .4655	4282			7689
rs1609682-4	X03833 3074. . .3719	3534			8437
rs1609682-3		3423			8548
rs1609682-2		3330			8641
rs1609682-1		3302			8669
SNP-889	X03833 421. . .690 and AC07953 651. . .921	549	793		11422
Exon 5 at 4336 SNP-2	AC079753 47797. . .48450		48223		58852
Exon 5 at 4336 SNP-1			48300		58929
rs1133558-2	AC079753 51783. . .52453		52141		62770
rs1133558-1			52220		62849
SNP-511	AC079753 52587. . .52891		52700		63329
Gaatp33330	AC016724 28980. . .29166			29069	205878
IL1F10-1 Ex2 5′UTR	AC016724 121811. . .122072			121961	298770
IL1F10-2 Ex3 Not Code	AC016724 123459. . .123712			123598	300407
IL1F10-3 Ex4 cSNP-1	AC016724 123867. . .124161			123965	300774
IL1F10-4 Ex4 cSNP-2				123986	300795
IL1F10-5 Ex5 3′UTR-1	AC016724 124592. . .124862			124634	301443
IL1F10-6 Ex5 3′UTR-2				124681	301490
rs315929-1	AC016724 160301. . .160916			160412	337221
rs315929-2				160475	337284
rs315929-3				160526	337335
rs315929-4				160586	337395
rs315929-5				160733	337542
rs315929-6				160745	337554
rs315929-7				160748	337557
rs315931-1	AC016724 160895. . .161545			160982	337791
rs315931-2				161022	337831
rs315932				161156	337965
rs315933-1				161178	337987
rs315933-2				161280	338089
rs315921-1	AC016724 162964. . .163622			163169	339978
rs315921-2				163229	340038
rs315921-3				163242	340051
rs315921-4				163416	340225
1731-1934-1	AC016724 166377. . .167045			166606	343415
1731-1934-2				166611	343420
1731-1934-3				166687	343496
1731-1934-4				166743	343552
1731-1934-5				166762	343571
1731-1934-6				166809	343618
1731-1934-7				166931	343740
rs315919-1	AC016724 167064. . .167710			167245	344054
rs315919-2				167247	344056
rs315919-3				167257	344066
rs315919-4				167359	344168
rs315919-5				167392	344201
rs315935-1	AC016724 172248. . .172893			172538	349347
rs315935-2				172651	349460
rs315935-3				172775	349584
Ex2 +8006-1	AC016724 178071. . .178705			178380	355189
Ex2 +8006-2				178435	355244

Table 1 *(continued)*

SNP Name	Primer set range	BAC 1[a]	BAC 2[a]	BAC 3[a]	Absolute position
Ex2 +8006-3				178439	355248
Ex2 +8006-4				178446	355255
Ex2 +8006-5				178467	355276
Ex2 +8006-6				178508	355317
Ex2 +8006-7				178511	355320
Ex2 +8006-8				178549	355358
Ex2 +8006-9				178572	355381
Ex2 +8006-10				178631	355440
Ex2 +8006-11				178858	355465
Int2 VNTR	AC016724 179245. . .179656			179281	356090
rs315955-1	AC016724 180293. . .180912			180603	357412
rs315955-2				180642	357451
rs315955-3				180804	357613
rs315954-1	AC016724 180744. . .181424			180961	357770
rs315954-2				180984	357793
rs314953-1				181176	357985
rs314953-2				181177	357986
rs314953-3				181198	358007
rs315952-1	AC016724 181431. . .181993			181477	358286
rs315952-2				181546	358355
rs315952-3				181552	358361
rs315951-1				181759	358568
rs315951-2				181783	358592
rs315951-3				181932	358741
rs9005-1	AC016724 182331. . .182984			182585	359394
rs9005-2				182804	359613
rs315950-1	AC016724 182803. . .183464			182948	359757
rs315950-2				182978	359787
rs315950-3				182994	359803
rs315950-4				183015	359824
rs315950-5				183043	359852
rs315950-6				183314	360123
rs315949-1	AC016724 183481. . .184156			183697	360506
rs315949-2				183725	360534
rs315949-3				183736	360545
rs315949-4				183947	360756

[a] BAC 1, Accession No. X03833.1; BAC 2, Accession No. AC079753.7; BAC 3, Accession No. AC016724.11.
[b] Not polymorphic in any study group.

Caucasians reaching comparable LD levels at approximately 350 kb.

To identify patterns of LD across the *IL1* gene cluster, rather than strictly relating D' calculated for *all pair-wise combinations* to distance as described above, *consecutive* (1–2, 2–3, 3–4, 4–5, . . ., 95–96) *high-frequency* (minor allele frequency $\geq$0.20) marker-to-marker D' values were calculated for each ethnic control group (Fig. 2A). Overall, regions between 0.5 and 11 kb were in LD with D' values above 0.80. Three intergenic D' valleys ranging in size from 36 to 237 kb were evident in both ethnic groups. In addition, a D' valley was observed between the promoter region and the coding portion of *IL1RN* in African Americans that was not as pronounced in Caucasians. To characterize LD further within this genomic region, reference markers [two markers were selected, one in *IL1A* (Fig. 2B) and one in *IL1RN* (Fig. 2C)] were selected and D' values were calculated between high-frequency markers and the reference marker as the distance from the reference marker increased. In general, patterns of LD are quite similar between the two ethnic groups (Fig. 2C). In Caucasians, the telomeric region of *IL1RN* appears in greater LD with the reference marker than in African Americans (Fig. 2C). There is no evident pattern of increase or decrease in D' for a marker pair comparing cases with controls in either ethnic group (data not shown).

Single-marker association analysis with ESRD

The *single-marker* association analysis with ESRD revealed significant associations with multiple markers (Fig. 3). SNP rs1516792-3 in intron 5 of *IL1A* was significantly associated with ESRD in both African American ESRD populations. The minor allele for this SNP is seen only in controls, in which the minor allele frequency is quite small (1.9%).

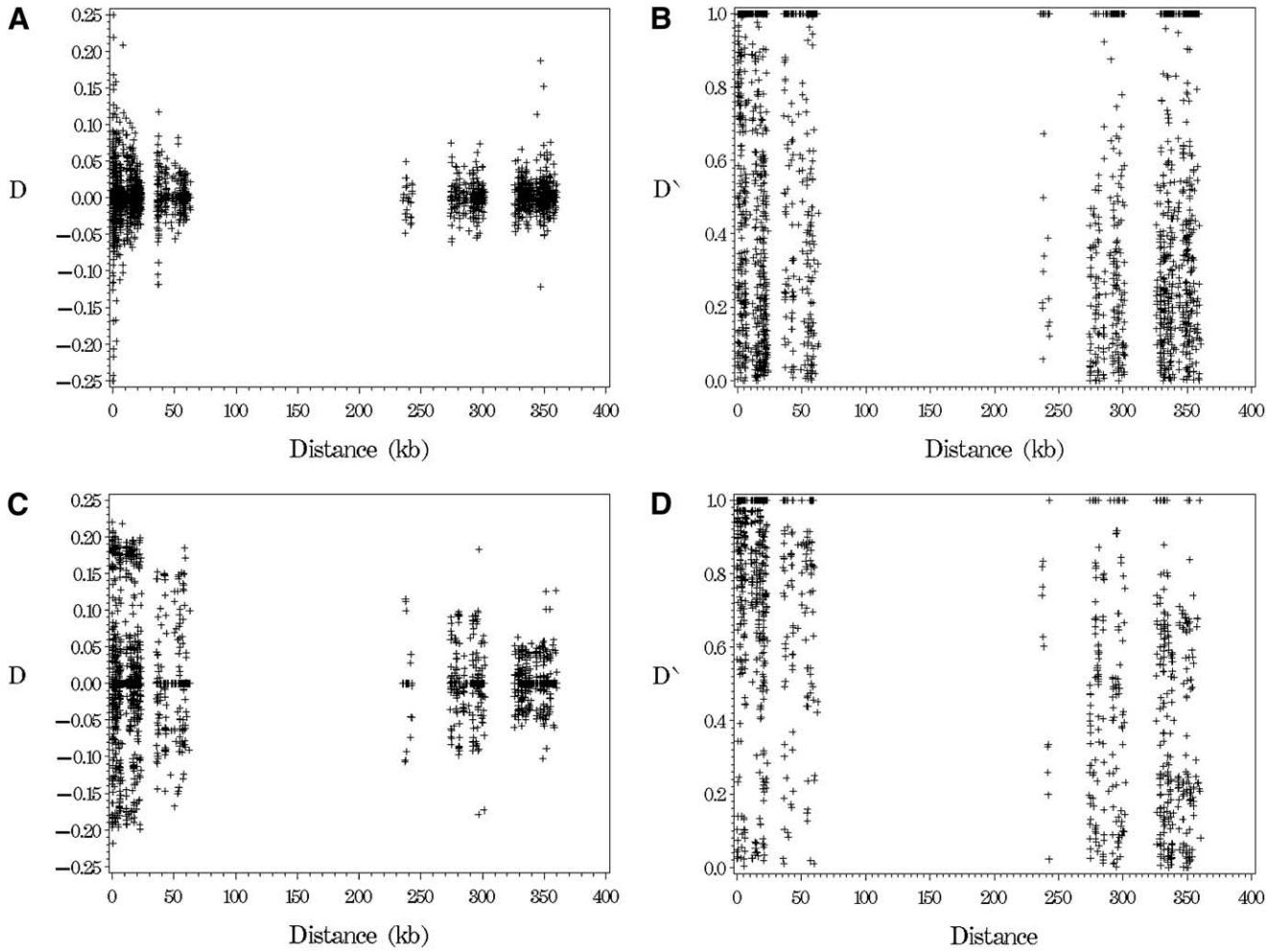

Fig. 1. (A) Pairwise LD (*D*) plotted by physical distance (kb) in the Caucasian control sample. (B) Pairwise LD (*D′*) plotted by physical distance (kb) in the Caucasian control sample. (C) Pairwise LD (*D*) plotted by physical distance (kb) in the African American control sample. (D) Pairwise LD (*D′*) plotted by physical distance (kb) in the African American control sample.

Additional regions of significant association identified only in the African American non-T2DM/ESRD group include a 3′UTR 4-bp insertion/deletion polymorphism in *IL1A* (rs16347-2, $p = 0.0024$) along with two intron 3 SNPs (rs1609682-3, $p = 0.0250$, and rs1609682-1, $p = 0.0033$), in *IL1F10* a 3′UTR SNP (IL1F10-5 Ex5 3′UTR-1, $p = 0.0474$), in *IL1RN* two promoter SNPs (rs315929-3, $p = 0.0345$, and rs315929-7, $p = 0.0448$) and an intron 2 SNP

Table 2
IL1 gene cluster coding SNPs

Gene	Location[a]	SNP	Codon change	Amino acid change
IL1A[b]	Exon 4	rs20540	C276[c]T	I92I
IL1A	Exon 5	rs17561-3	T340G	S114A
IL1B	Exon 5	Ex5 at 4336 SNP-2	C315T	F105F
IL1F10	Exon 3	IL1F10-2 Ex3	T72C	D24D
IL1F10	Exon 4	IL1F10-3 Ex4 cSNP-1	T131C	I44T
IL1F10	Exon 4	IL1F10-4 Ex4 cSNP-2	C152A	A51D
IL1RN	Exon 2	Ex2 +8006-1	T120C	A40A
IL1RN	Exon 4	rs315952-1	C399T	S133S
IL1RN	Exon 4	rs315952-2	C468T	D156D
IL1RN	Exon 4	rs315952-3	C474T	P158P

[a] Precise BAC location provided in Table 1.
[b] *IL1A* is transcribed in reverse orientation.
[c] Position relative to first nucleotide in initiator methionine codon.

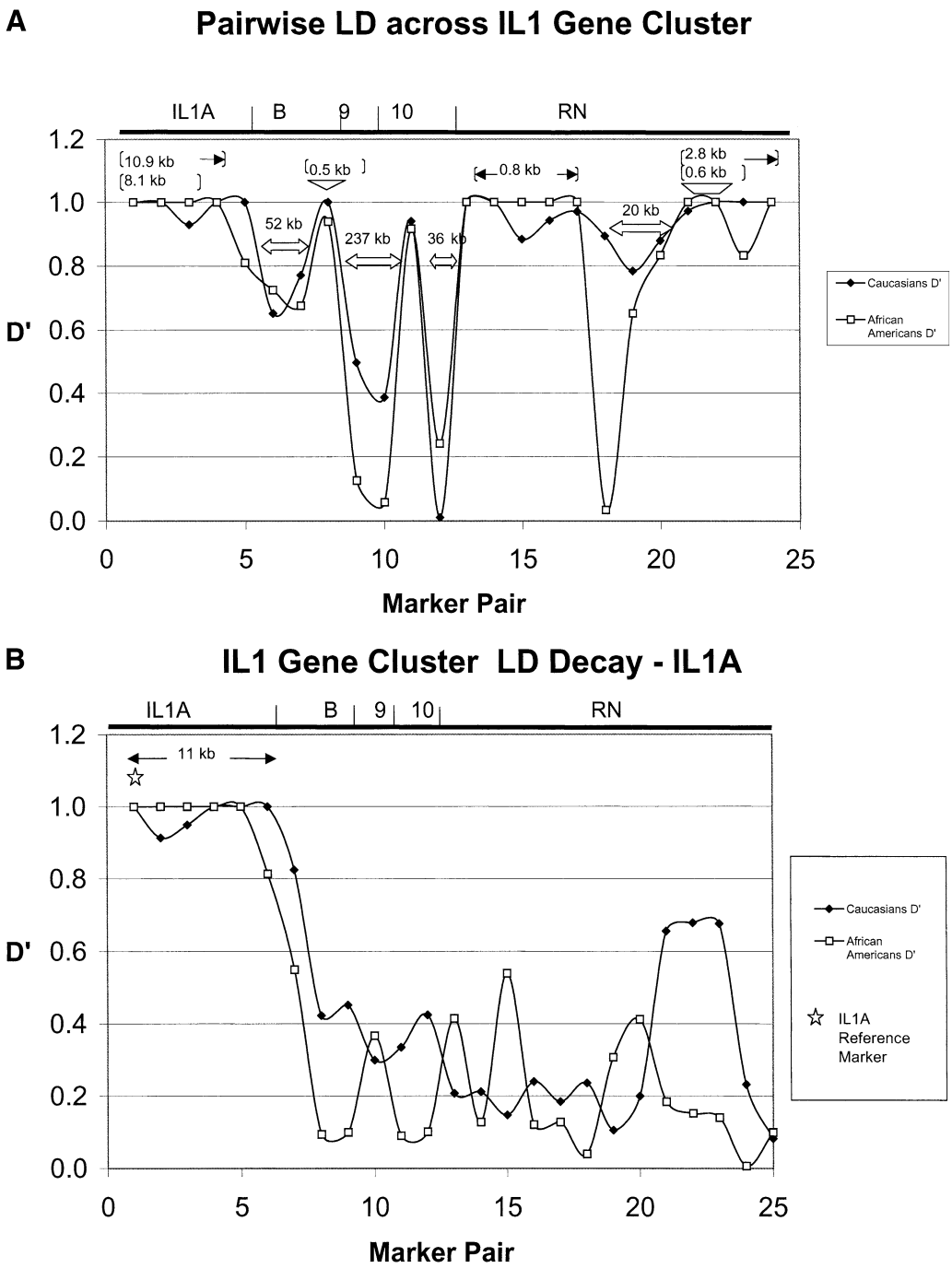

Fig. 2. (A) LD (D') between consecutive high-frequency (minor allele ≥0.20) marker pairs across *IL1* gene cluster in Caucasian and African American control groups. While physical distance is not drawn to scale, marker location can be identified using the gene boundary map (solid dark line) provided above the graph. Approximate distances are provided within parentheses and represented by arrows. (B) Pairwise LD (D') decay of high-frequency markers (minor allele ≥0.20) from *IL1A* reference marker (rs16347-2) in Caucasian and African American control groups. For example, if marker 2 is the reference marker, D' values were calculated from marker 2–7, 2–9, 2–11, 2–13, 2–14, and so on given that markers 7, 9,1 1, 13, and 14 are high-frequency markers. (C) Pairwise LD (D') decay of high-frequency markers (minor allele ≥0.20) from *IL1RN* reference marker (rs315931-1) in Caucasian and African American control groups. See (B) description.

(Ex2 +8006-6, $p = 0.0344$). Of the eight disease-associated markers identified in African American non-T2DM/ESRD cases four had minor allele frequencies ≥17%, while the other four had minor allele frequencies ≤8%. Two markers were significantly associated only in the African American T2DM/ESRD cases (rs315933-2 and rs315951-2), with rs315933-2 being the most highly significant ($p = 0.0148$).

In summary, three SNPs (rs1516792-3, rs315933-2, and rs315951-2) in the African American T2DM/ESRD group and eight SNPs (rs16347-2, rs1516792-3, rs1609682-3,

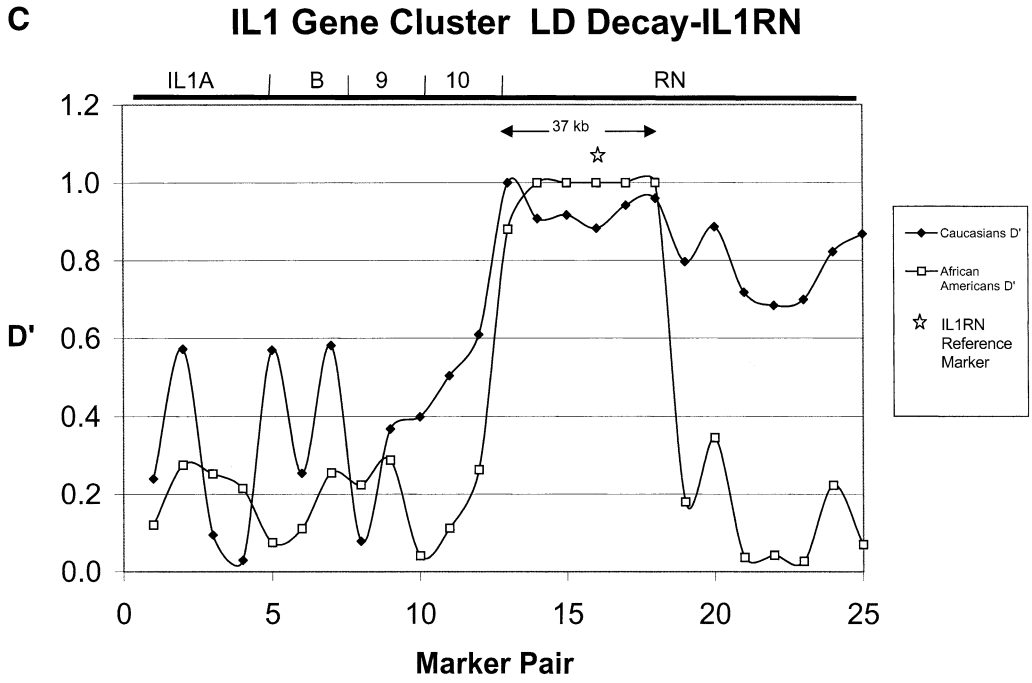

Fig. 2 (*continued*)

rs1609682-1, IL1F10-5 Ex5 3′UTR-1, rs315929-3, rs315929-7, and Ex2 +8006-6) in the non-T2DM/ESRD group were associated with disease. No markers were associated with ESRD in Caucasians.

Haplotype association analysis

The identification of multiple individual SNPs in the *IL1* gene cluster associated with ESRD suggests the possibility that combinations of these SNPs may contribute increased risk. Initially two- and three-marker moving-window haplotype analysis was performed in each study group. This analysis, however, provided little additional information compared to the single marker analysis described above. To explore the possibility that a combination of associated genes may function together to predispose to disease, we constructed haplotypes of associated sequence variants.

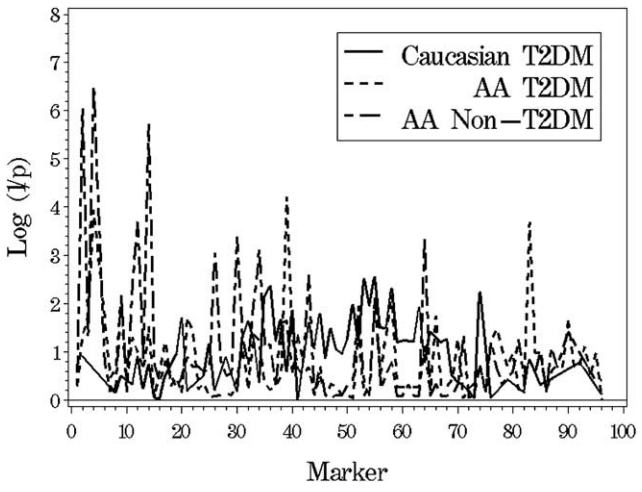

Fig. 3. Single-marker association analysis in three case–control comparison groups: Caucasian T2DM, Caucasian T2DM/ESRD vs Caucasian controls; AA T2DM, African American T2DM/ESRD vs African American controls; AA Non-T2DM, African American non-T2DM/ESRD vs African American controls. A significant association ($p = 0.05$) is indicated by y-axis values $[\ln(1/p)] \geq 3.0$.

African American T2DM/ESRD

Using the three associated markers (Table 3) observed in African American T2DM/ESRD, haplotypes were constructed in African American controls and T2DM/ESRD cases, respectively (Table 4). Two different haplotypes were inferred in cases and four in controls. Significant differences ($p = 0.005$) were evident between haplotype frequency distributions, with haplotype 1 occurring more frequently in cases than in controls. Two and a half percent (7/276) and 9.5% (25/258) of marker sites within the haplotype data were unknown in cases and controls, respectively. In this case PHASE assigns a probability for the most likely marker allele within the haplotype. The vast majority of phase probability estimates at these sites were over 95%. LD analysis between the three associated markers in the African American controls indicated complete LD between SNPs rs315933-2 and rs315951-2 ($D' = 1.0000$) but not between rs1516792-3 and rs315933-2 ($D' = 0.5906$).

African American non-T2DM/ESRD

Using the seven markers that showed evidence of association with non-T2DM/ESRD (Table 3), 7 and 12 different

Table 3
Risk haplotypes marker order and gene location

(A) Three-marker haplotype markers—T2DM/ESRD cases						
Haplotype position	SNP name	Gene	Allele		Location	*p* value
			1	2		
1	rs1516792-3	*IL1A*	A	G	Intron 5	0.0194
2	rs315933-2	*IL1RN*	C	T	Promoter	0.0148
3	rs315951-2	*IL1RN*	C	T	Exon 4ic 3′UTR	0.0251
(B) Seven-marker haplotype markers—non-T2DM/ESRD cases						
Haplotype position	SNP name	Gene	Allele		Location	*p* value
			1	2		
1	rs16347-2	*IL1A*	+	−	3′UTR	0.0024
2	rs1516792-3	*IL1A*	A	G	Intron 5	0.0015
3	rs1609682-3	*IL1A*	T	C	Intron 3	0.0250
4	rs1609682-1	*IL1A*	A	C	Intron 3	0.0033
5	rs315929-3	*IL1RN*	G	A	Promoter	0.0345
6	rs315929-7	*IL1RN*	G	A	Promoter	0.0448
7	Ex2 +8006-6	*IL1RN*	G	T	Intron 2	0.0344

haplotypes were constructed and frequencies determined for both African American non-T2DM/ESRD and control groups, respectively (Table 4). Highly significant differences existed in the 12 seven-marker haplotype frequencies between cases and controls ($p \leq 0.00005$), with haplotypes 1 and 4 occurring more frequently in cases. In addition, the haplotype pair (1,4) occurred significantly more frequently in cases (17/96) than in controls (6/86) (using χ^2 statistic, $p \leq 0.05$). Approximately 17% (117/672) and 24% (146/602) of marker sites within the haplotype data had unknown phase in African American cases and controls. Typically, only a single marker within a haplotype was missing phase. For these sites PHASE computed the probability that the estimated phase assignment was correct. Ninety percent of phase probabilities were greater than 80% at unknown sites. Linkage disequilibrium (D') was measured between each

Table 4
Reconstructed risk haplotype (HT) frequency

(A) Three markers						
HT	African American controls			African American T2DM/ESRD cases[a]		
	3-Marker HT	HT number	HT frequency (%)	3-Marker HT	HT number	HT frequency (%)
1	111	153	89	111	179	97
2	121	13	7.5	121	5	3
3	112	3	1.5			
4	211	3	1.5			
(B) Seven Markers						
HT	African American Controls			African American non–T2DM/ESRD cases[b]		
	7-Marker HT	HT number	HT frequency (%)	7-Marker HT	HT number	HT frequency (%)
1	1111212	84	49	1111212	107	56
2	2122112	6	4		0	0
3	2122212	12	7	2122212	11	6
4	1111222	14	8	1111222	36	19
5	2112212	24	14	2112212	21	11
6	1111112	12	7	1111112	14	7
7	2112112	4	2		0	0
8	1111211	8	4	1111211	2	1
9	2112222	1	1		0	0
10	1211112	3	2		0	0
11	2122222	4	2		0	0
12	1111122	0	0	1111122	1	<1

[a] $p \leq 0.005$ (SE = 0.0022).
[b] $p \leq 0.00005$ (SE = 0.0000).

Table 5
Linkage disequilibrium comparisons—African American seven-marker haplotype

Marker pair	Location[a]	Controls D'
rs16347-2–rs1516792-3	IL1A 3′UTR to Intron 5	0.9998
rs1516792-3–rs1609682-3	IL1A Intron 5 to Intron 3	0.1372
rs1609682-3–rs1609682-1	IL1A Intron 3 to Intron 3	1.0000
rs1609682-1–rs315929-3	IL1A Intron 3 to IL1RN promoter	0.1567
rs315929-3–rs315929-7	IL1RN promoter to promoter	1.0000
rs315929-7–Ex2 +8006-6	IL1RN promoter to Intron 2	1.0000

[a] IL1A transcribed in reverse orientation

consecutive marker pair of the seven-marker haplotype (Table 5). Noteworthy drops in LD were identified within *IL1A* between intron 3 (rs1516792-3) and intron 5 (rs1609682-3), spanning a distance of 4.2 kb, as well as between a large 329-kb intergenic region between *IL1A* (rs1609682-1) and *IL1RN* (rs315929-3). Haplotypes associated with disease are evident in both African American groups, haplotype 1 in African American T2DM/ESRD (odds ratio = 4.3) and both haplotypes 1 (odds ratio = 1.4) and 4 (odds ratio = 2.9) among African American non-T2DM/ESRD cases (Table 6). Risk haplotypes 1 and 4 differ only at SNP rs315929-7, with allele 1 in haplotype 1 and allele 2 in haplotype 4. Both haplotypes are associated with non-T2DM/ESRD in African American cases. SNP rs315929-7 is an *IL1RN* promoter polymorphism. While less frequent, the odds ratio associated with carrying haplotype 4 (at least one copy) is more than twice that for haplotype 1. Of note, no homozygous (4,4) individuals were identified in the control group, while three were present in the African American non-T2DM/ESRD cases.

Discussion

Previous reports on the association between *IL1* gene polymorphisms and diabetic nephropathy have been inconsistent [5–7,9]. To address systematically whether there is a link between genes in this region and ESRD we constructed a dense, gene-based SNP map of the *IL1* gene cluster to evaluate the association with ESRD. To facilitate this analysis a detailed characterization of LD across this genomic region is described. The association analysis included the evaluation of both single markers and haplotypes for association with ESRD.

Ethnic variations in allele frequency were evident between our Caucasian and African American control groups and in some cases were quite striking, with a 24% difference in minor allele frequency. As expected, we identified a greater number of population-specific markers in African Americans ($N = 26$) than in Caucasians ($N = 1$), and of these many were of lower frequency compared to SNPs verified in both ethnic groups.

Consistent with many prior studies, LD in this study was not strictly correlated with distance between markers. LD was measured as both D and D' across a relatively short (360 kb) genomic region (Figs. 1A through 1D). Variability in LD was observed across both short (ex. 76 bp between markers with $D' = 0.10$) and long segments (ex. 142 kb between markers with $D' = 0.50$) of the *IL1* gene cluster (data not shown). One potential explanation for this may be our use of both high-frequency (minor allele frequencies $\geq$0.20) and low-frequency markers. Approximately 30% of markers uniformly distributed across the region had a minor allele frequency $\leq$10%, leading to inflated measures of D'. Figs. 1A through 1D show the association between LD and distance, with D reflecting the predicted decrease in LD with increasing distance. We observed the generally accepted observation that younger populations (Caucasians) had greater LD over the same genomic regions compared to an older population (African Americans) [10–16]. Specifically, at 350 kb Caucasians had levels of LD comparable to those observed in African Americans at 50 kb. Similar to findings by Pritchard and Przeworski [17], this trend was not visible when D' was used, which may be a reflection of the sensitivity of D' to low allele frequency present in a significant proportion of the total data set. When D' across the region was evaluated using only high-frequency markers ($>$0.20), both consecutive marker pair D' (Fig. 2A) and D' between marker and reference SNPs (Figs. 2B and 2C) displayed comparable LD trends between ethnic groups across this gene cluster; however, one notable exception existed (Fig. 2C). Among Caucasians LD is preserved ($D' > 0.70$) from *IL1F10* exon 4 to the *IL1RN* 3′ flanking region, a distance of 60 kb, while among African Americans, following a 37-kb region of LD a dramatic drop occurs between the *IL1RN* promoter (rs315933-1) and intron ic1 long (rs315919-5) (an inter-SNP distance of 6 kb). Interestingly, in our evaluation using a reference SNP in *IL1A*, this 3′ end of *IL1RN* (SNPs rs315953-1, rs315952-1, and rs315951-1) that remains in LD in Caucasians also shares comparable levels of LD with *IL1A* in Caucasians. This

Table 6
Haplotype odds ratios

Group/haplotype pair	No. of cases	No. of controls	Odds ratio (95% CI)
African American T2DM/ESRD			
1,1	87	69	4.3 (1.5, 12.2)
Not 1,1	5	17	
African American non-T2DM/ESRD			
(1,1) and (1,_)	76	63	1.4 (0.7, 2.8)
No haplotype 1	20	23	
(4,4) and (4,_)	33	13	2.9 (1.4, 6.1)
No haplotype 4	63	73	

finding may reflect inherent variability in measures of D' at shorter genomic intervals even in high-frequency markers [18–20]. As hypothesized by Cox and colleagues [21] and Pritchard and Przeworski [17], mutation, genetic drift, gene conversion, and admixture may play a greater role at shorter distance than recombination in larger genomic segments.

Single-marker disease association analyses within the *IL1* gene cluster suggest that this region contains multiple susceptibility loci for ESRD in the African American population. This genomic region, however, seems to contribute little to disease among Caucasian T2DM/ESRD cases. No Bonferroni multiple comparisons adjustments were made due to the exploratory nature of this study. In addition, such a correction for correlated SNPs residing in close physical space would likely provide an overly conservative estimate of type I error. For this reason, the results of the single-marker association should be interpreted with caution. To explore further the possibility that the SNPs associated with disease in both the African American case groups may interact together to predispose to disease, three- and seven-marker haplotypes were constructed using PHASE and frequencies compared between cases and controls. PHASE reconstructs haplotypes given population-based genotype data with a reported up to 50% reduction in error rate over the EM algorithm. While missing data can significantly affect reconstruction algorithms, PHASE has been shown to perform well in typical data sets such as ours, with approximately 5% missing genotype data. In each case group, the associated haplotypes spanned nearly 355 kb of genomic sequence, extending from *IL1A* to *IL1RN*. Fewer haplotypes were reconstructed in both case groups (Tables 4A and 4B), suggesting less genetic variability within this region. These haplotypes had significantly different frequency distributions in cases compared to controls, suggesting a role in the development of ESRD. Due to the exploratory nature of this analysis, however, these results should be interpreted with caution. Evaluation of individual SNPs and risk haplotypes is currently being carried out in a greatly expanded population of cases and controls to evaluate these issues further.

SNP rs315929-7 allele 2 is the single difference between "risk" haplotype 1 [OR = 1.4 (0.7, 2.8)] and haplotype 4 [OR = 2.9 (1.4, 6.1)] in African American non-T2DM/ESRD cases. This minor allele variant within the *IL1RN* promoter (6 kb upstream) may play a significant role in altering a transcription factor binding site or other regulatory element. Significant gene duplication has occurred in this region, thus this variant may even reside in a yet undescribed "*IL1RN*-like" gene. Given the inflammatory component of ESRD, one might expect this variant to down-regulate *IL1RN;* however, no functional data are currently available for this SNP. Further studies evaluating the functional relevance of markers within these associated haplotypes, as well as an assessment of their interaction, will contribute to our understanding of the molecular mechanisms involved in renal disease.

Materials and methods

Subjects

Association studies were conducted in five distinct sample groups. End-stage renal disease patient DNA samples were from our ongoing studies of diabetes-associated and non-diabetes-associated renal failure. Caucasian samples consisted of 95 healthy unrelated individuals as controls and 75 unrelated ESRD patients with a diagnosis of Type 2 diabetes. African American samples consisted of 86 healthy unrelated employees at North Carolina Baptist Hospital (Winston-Salem, NC, USA) as controls, 92 samples from ESRD patients with a diagnosis of Type 2 diabetes, and 95 samples from ESRD patients with a diagnosis of non-diabetes-associated renal failure (primarily chronic glomerular disease or unknown etiologies of ESRD). Recruitment and clinical characteristics of these patients have been previously described [6].

DNA preparation

Total genomic DNA was purified from peripheral blood leukocytes using the PureGene DNA isolation kit (Genetra, Minneapolis, MN, USA). Quantification of DNA was determined using standardized fluorimetric readings on a Hoefer DyNA Quant 200 fluorimeter (Hoefer Pharmacia Biotech, Inc., San Francisco, CA, USA).

SNPs identified

Due to a lack of complete and ordered DNA sequence within this region, 38 polymorphisms (92% SNPs), primarily located within the known interleukin-1 cluster gene, were identified through one of the following methods (Appendix 2, see "Names"):

1. published reports,
2. NCBI dbSNP,
3. SSCP exon screening.

Additional polymorphisms were identified through direct sequencing of DNA flanking one of the 38 original polymorphism PCR products (Table 1). Subsequent to PCR amplification, polymorphisms were genotyped using one of the following methods:

1. agarose electrophoresis,
2. SSCP (ABI 377 DNA sequence verified),
3. PAGE,
4. direct sequencing (ABI 3700),
5. Sequenom.

In addition, approximately 10% of all samples were amplified in the reverse direction and genotypes verified by direct DNA sequencing.

Agarose gel electrophoresis

DNAs were PCR amplified in 20-μl final volume PCRs using 60 ng of human genomic DNA (see PCR conditions in the next section for SSCP analysis) and the appropriate annealing temperature indicated in Appendix 2. Amplified DNA products with 2.5 μl agarose loading buffer (0.25% xylene cyanol, 0.25% bromophenol blue, and 25% Ficoll Type 400) were separated in 1× TAE buffer using 2% high-melting agarose (Continental Laboratory Products) gels containing 30 μg ethidium bromide for 1 h and 15 min at 170 V and 105 mA.

SSCP and DNA sequence analysis

Oligonucleotide primers are listed in Appendix 2 along with amplimer size, annealing temperature, method of identification/verification/genotyping, and gene location. Primers were end-labeled with [γ-^{32}P]dATP (4500 Ci/mmol, 2.2 μM; ICN Radiochemicals) in a 37.5-μl reaction (105 pmol of oligonucleotide primer, 700 mM Tris–HCl (pH 7.6 at 25°C), 100 mM $MgCl_2$, 50 mM DTT, and 7.5 units of T4 polynucleotide kinase; Promega) at 37°C for 30 min, followed by heat inactivation at 70°C for 10 min. PCR amplification for SSCP was performed using 60 ng of genomic DNA in a 10-μl reaction (50 mM KCl, 100 mM Tris–HCl, pH 8.8, 1.2 mM $MgCl_2$, 0.2 mM each dNTP, 4 pmol each of unlabeled oligonucleotide, 0.9 pmol each of end-labeled oligonucleotide, 8 nmol of spermidine, and 1.0 unit of *Taq* polymerase. The PCR mixture was denatured for 5 min at 94°C and then cycled for 35 cycles [45 s at 94°C, 45 s at designated annealing temperature (see Appendix 2), and 45 s at 72°C], with a final extension at 72°C for 7 min. The PCR products were then diluted 1:8 with loading dye (95% formamide, 10 mM NaOH, 0.25% bromophenol blue, 0.25% xylene cyanol), denatured at 95°C for 2 min, and placed immediately on ice. Heat-denatured products (1.6 μl) were then separated by electrophoresis on polyacrylamide gels containing 0.5× MDE acrylamide (FMC Bioproducts), 0.6× TBE, and 5% glycerol at room temperature for 17 h at 11 W. Gels were exposed to X-ray film (Kodak) for 18–24 h.

DNA samples displaying variant banding patterns in SSCP were prepared for sequencing using 100 ng of human genomic DNA in a reaction volume of 60 μl containing 10 mM Tris–HCl, pH 8.2; 50 mM KCl; 1.5 mM $MgCl_2$; 0.2 mM each dATP, dCTP, dGTP, dTTP; 24 pmol of forward and reverse primer; and 1.5 units of *Taq* polymerase. PCR cycling conditions were identical to those used in SSCP-PCR described above. PCR products were separated on a 1% agarose gel containing ethidium bromide (220 μg/L) in a 1× TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) and visualized on an ultraviolet transilluminator. The DNA product was gel extracted using the QIAquick gel extraction kit (Qiagen, Inc., Valencia, CA, USA) following the manufacturer's recommendations. Approximately 50–150 ng of purified DNA from each samples and 3.6 pmol of the appropriate oligonucleotide primer were submitted to the Wake Forest University Baptist Medical Center DNA Sequencing Core Facility for fluorescence sequencing using an ABI Prism 377 automated sequencer (Applied Biosystems, Foster City, CA, USA).

PAGE

PAGE was performed similar to SSCP methods (see SSCP section) except for the following:

1. A single primer was end-labeled,
2. 1× TBE running buffer was used,
3. PCR product was diluted 1:1 with loading dye [80% formamide, 10 mM EDTA, 1 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue],
4. 6% PAGE was prerun for 30 min at 60 W, and
5. PCR products were heat denatured but NOT placed on ice prior to loading.

Direct DNA sequencing and allele genotyping

All PCRs for direct DNA sequencing were performed in a total volume of 30 μl using 50 ng of human genomic DNA (see PCR conditions in the above section for SSCP analysis). DNA sequencing was performed using the ABI BigDye Terminator sequencing kit (Applied Biosystems, Inc.). Each 10-μl sequencing reaction contained 10–50 ng of purified PCR product, 1.5 pmol of sequencing primer, 1 μl of BigDye Terminator mix, 1.5 μl of 5× sequencing dilution buffer (400 mM Tris, pH 9.0, 10 mM $MgCl_2$), and water to volume. Cycling conditions were 94°C for 1 min and 25 cycles of 94°C for 30 s, 50°C for 30 s, and 60°C for 4 min, finishing with a single 72°C extension step for 5 min. Sequencing products were ethanol precipitated, air-dried, resuspended in 25 μl ddH_2O, and analyzed on an ABI 3700 DNA analyzer. DNA sequencing data were aligned and polymorphisms verified and genotyped using Sequencher DNA analysis software (Gene Codes Corp., Ann Arbor, MI, USA) and Phred/Phrap/PolyPhred software [22–24]. While sequence alignment was performed and polymorphic ranking scores were identified using Phred/Phrap/PolyPhred software, genotypes were read by eye.

MALDI-TOF MassARRAY genotyping

Amplified and purified DNA samples were robotically dispensed onto a silicon chip (SpectroCHIP, Sequenom, San Diego, CA, USA) and analyzed using a MassARRAY system (Sequenom) generally as described [25], with the following specifications for our evaluation of *IL1F10*.

PCR primer design. Primers were designed based on our genomic sequence (AC016724) of novel gene *IL1F10* using SpectroDESIGNER (Sequenom) to produce products of approximately 100 bp with 50 bp flanking the identified SNP (Table 7). Extension primers again using SpectroDESIGNER (Sequenom; see Table 7) were designed for use in

Table 7
IL1F10 SNP *Sequenom MassARRAY* amplification and extension primers

SNP	NTP mix[a]	Amplification primers (5′–3′)	Extension primer
Ex2	ddACT	F ACGTTGGATGTTCTAACTGCCCTTCTCTCC	AGGGAACACATTCCTGCA
	dG	R ACGTTGGATGGACATGCCAGGACAACTTAC	
Ex4-1	ddACT	F ACGTTGGATGTCTCTTCCCTCCTAGAGAAG	CCAAGCCTCTGTTAGGAAGT
	dG	R ACGTTGGATGTGGATCCCCAGGAAAATGGG	
Ex4-2	ddCGT	F ACGTTGGATGTCTCTTCCCTCCTAGAGAAG	TTCCTAACAGAGGCTTGG
	dA	R ACGTTGGATGTGGATCCCCAGGAAAATGGG	
Ex5-1	ddACT	F ACGTTGGATGCAGGAAACTGCGTTTTAGCC	CCTGAGCAGGATGAGCTT
	dG	R ACGTTGGATGGGGACATTATTCTGCCTACC	
Ex5-2	ddACT	F ACGTTGGATGCTTGGGATTAGGATGTGGAC	AGGCAGAATAATGTCCCCC
	dG	R ACGTTGGATGAAACCAAGCTCATCCTGCTC	

[a] Terminator mix of three dideoxy NTPs (listed) with the remaining deoxy NTP (listed below).

MALDI-TOF MassARRAY genotyping of five of the six confirmed *IL1F10* SNPs. SNP IL1F10-2 Ex3, a noncoding SNP, was evaluated by SSCP and direct sequencing in a subset of case–control subjects due to difficulties in primer design using the Sequenom software.

Amplification reaction. The PCR amplification (5 μl total volume) contained 1× HotStar *Taq* PCR buffer 2.5 mM $MgCl_2$, 200 μM each dNTP (Promega), 0.1 U Enzyme HotStar *Taq* polymerase (Qiagen), 100 nM each forward and reverse extension primer, and 5 ng of genomic DNA. The PCR amplification was performed in 384-well plates using the following conditions: 95°C for 15 min and 45 cycles of 95°C for 20 s, 56°C for 30 s, 72°C for 1 min, followed by a final extension step of 72°C for 3 min.

Statistical Analysis

Linkage disequilibrium

For each of the SNPs we tested whether the observed allele frequencies departed from Hardy–Weinberg proportions. We estimated the degree of linkage disequilibrium between all pairs of SNPs using the classic linkage disequilibrium coefficient *D*, where for alleles *a* and *b* at the respective loci we have:

$$\hat{D}_{ab} = \hat{p}_{ab} - \hat{p}_a\hat{p}_b\,. \tag{1}$$

Here, $\hat{p}_{ab}$ is the probability of the haplotype *ab* estimated by the expectation-maximization (EM) algorithm [26], and $\hat{p}_a$ and $\hat{p}_b$ are the respective maximum likelihood estimates of allele frequency. The normalized estimate of D_{ab}, $\hat{D}'_{ab}$, is obtained by dividing $\hat{D}_{ab}$ by its maximal value as determined from allele frequencies [27]. Thus, the statistic $\hat{D}_{ab}$ estimates the magnitude of linkage disequilibrium and $\hat{D}'_{ab}$ estimates the signed proportion of the maximal value that $\hat{D}_{ab}$ could obtain conditional on allele frequencies and sample size.

Case–control analyses

We computed ethnic-specific tests of association for each SNP and each tandem two- and three-marker haplotype. Specifically, we made the following three comparisons: (1) T2DM/ESRD Caucasian cases with Caucasian controls, (2) T2DM/ESRD African American cases with African American controls, and (3) non-T2DM/ESRD African American cases with African American controls. For each SNP we tested for an association using a permutation test based on the standard multinomial likelihood ratio test statistic and 10,000 permutations. We tested for a haplotypic association by estimating via the EM algorithm all tandem two- and three-marker haplotypes and computing the haplotype-based permutation test of association using the multinomial likelihood ratio test statistic and 10,000 permutations of the data [28].

Disease associated markers—haplotype and LD analysis

Genotype data for each disease-associated SNP in both the African American T2DM/ESRD and the African American non-T2DM/ESRD group were used to reconstruct three- and seven-marker haplotypes, respectively, using PHASE [29]. We utilized the PHASE algorithm (version 1.0.1) that estimates population haplotypes based on all possible haplotype comparisons (PHASE.small). Haplotype frequencies between cases and controls were compared using the RXC program (George Carmody, Carleton University, Ottawa, ON, Canada), which carries out Monte Carlo simulations [30,31] to calculate the statistical significance of contingency tables. Pair-wise LD measured by D' was calculated between associated markers (see Linkage disequilibrium).

Acknowledgments

This work was supported by National Institutes of Health Grants RO1 DK53591 (D.W.B.) and RO1 HL56266 (B.I.F.).

Appendix 1A

IL1 gene cluster variant allele frequencies—Caucasian controls

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs16347-3	0.9896	95.0	1					
rs16347-2	0.9896	95.0	0.6684	0.3316				
rs16347-1	0.9896	95.0	1					
rs1516792-3	0.9479	91.0	1					
rs1516792-2a	0.9479	91.0	1					
rs1516792-1	0.9479	91.0	1					
rs17561-3	0.9896	95.0	0.7526	0.2474				
rs17561-2	0.9896	95.0	0.7526	0.2474				
rs17561-1	0.9896	95.0	0.7053	0.2947				
rs20540	0.9896	95.0	1					
rs1609682-4	0.9896	95.0	0.5789	0.4211				
rs1609682-3	0.9896	95.0	0.6737	0.3263				
rs1609682-2	0.9844	94.5	0.7513	0.2487				
rs1609682-1	0.9896	95.0	0.6737	0.3263				
SNP-889	0.9896	95.0	0.2526	0.7474				
Ex5 at 4336 SNP-2	0.9844	94.5	0.8095	0.1905				
Ex5 at 4338 SNP-1	0.9896	95.0	0.3895	0.6105				
rs1133558-2	0.9896	95.0	1					
rs1133558-1	0.9896	95.0	0.3526	0.6474				
SNP-511	0.9896	95.0	0.6632	0.3368				
Gaatp33330	1	96.0	0.6719	0.2396	0	0.0885		
IL1F10-1 Ex2 5′UTR	0.9271	89.0	1					
IL1F10-2 Ex3 no code	0	0.0						
IL1F10-3 Ex4 cSNP-1	0.9792	94.0	0.5745	0.4255				
IL1F10-4 Ex4 cSNP-2	0.9062	87.0	0.6207	0.3793				
IL1F10-5 Ex5 3′UTR-1	0.8958	86.0	0.7558	0.2442				
IL1F10-6 Ex5 3′UTR-2	0.9688	93.0	1					
rs315929-1	0.9896	95.0	0.4579	0.5421				
rs315929-2	0.9896	95.0	1					
rs315929-3	0.9688	93.0	0.1774	0.8226				
rs315929-4	0.9896	95.0	0.9579	0.0421				
rs315929-5	0.9896	95.0	0.7263	0.2737				
rs315929-6	0.9844	94.5	0.5291	0.4709				
rs315929-7	0.9896	95.0	0.9158	0.0842				
rs315931-1	0.9583	92.0	0.7337	0.2663				
rs315931-2	0.9583	92.0	0.2609	0.7391				
rs315932	0.9583	92.0	0.9511	0.0489				
rs315933-1	0.9427	90.5	0.7403	0.2597				
rs315933-2	0.9479	91.0	0.8901	0.1099				
rs315921-1	0.9375	90.0	0.9611	0.0389				
rs315921-2	0.9271	89.0	0.8202	0.1798				
rs315921-3	0.9375	90.0	1					
rs315921-4	0.875	84.0	0.8452	0.1548				
1731-1934-1	0.9896	95.0	0.7368	0.2632				
1731-1934-2	0.8229	79.0	0.8608	0.1392				
1731-1934-3	0.9792	94.0	0.7340	0.2660				
1731-1934-4	0.9583	92.0	0.7446	0.2554				
1731-1934-5	0.9792	94.0	0.7340	0.2660				
1731-1934-6	0.9583	92.0	0.7337	0.2663				
1731-1934-7	0.9635	92.5	0.7405	0.2595				
rs315919-1	0.974	93.5	0.7540	0.2460				
rs315919-2	0.974	93.5	1					
rs315919-3	0.974	93.5	0.7647	0.2353				
rs315919-4	0.9792	94.0	0.2447	0.7553				
rs315919-5	0.9792	94.0	0.6649	0.3351				

Appendix 1A *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs315935-1	0.9271	89.0	1					
rs315935-2	0.9271	89.0	1					
rs315935-3	0.9271	89.0	1					
Ex2 +8006-1	0.9896	95.0	0.7316	0.2684				
Ex2 +8006-2	0.9896	95.0	0.2684	0.7316				
Ex2 +8006-3	0.9896	95.0	1					
Ex2 +8006-4	0.9896	95.0	0.7316	0.2684				
Ex2 +8006-5	0.9792	94.0	0.7500	0.2500				
Ex2 +8006-6	0.9792	94.0	0.2340	0.7660				
Ex2 +8006-7	0.9792	94.0	0.7394	0.2606				
Ex2 +8006-8	0.9792	94.0	1					
Ex2 +8006-9	0.9896	95.0	0.7316	0.2684				
Ex2 +8006-10	0.9792	94.0	0.2606	0.7394				
Ex2 +8006-11	0.0104	1.0	1					
Int2 VNTR	0.9896	95.0	0.7053	0.2684	0.0211	0.0053		
rs315955-1	0.9896	95.0	1					
rs315955-2	0.9844	94.5	0.7302	0.2698				
rs315955-3	0.9896	95.0	0.7211	0.2789				
rs315954-1	0.974	93.5	1					
rs315954-2	0.9792	94.0	1					
rs314953-1	0.9792	94.0	0.7394	0.2606				
rs314953-2	0.9323	89.5	1					
rs314953-3	0.9896	95.0	1					
rs315952-1	0.9896	95.0	0.2632	0.7368				
rs315952-2	0.9896	95.0	1					
rs315952-3	0.9896	95.0	1					
rs315951-1	0.9896	95.0	0.7316	0.2684				
rs315951-2	0.9635	92.5	0.9784	0.0216				
rs315951-3	0.9896	95.0	1					
rs9005-1	0.9844	94.5	0.7090	0.2910				
rs9005-2	0.9792	94.0	1					
rs315950-1	0.9896	95.0	0.7053	0.2947				
rs315950-2	0.9792	94.0	1					
rs315950-3	0.0521	5.0	1					
rs315950-4	0.9896	95.0	1					
rs315950-5	0.0521	5.0	1					
rs315950-6	0.0625	6.0	0.6667	0.3333				
rs315949-1	0.9896	95.0	0.9105	0.0895				
rs315949-2	0.9896	95.0	1					
rs315949-3	0.9896	95.0	1					
rs315949-4	0.9896	95.0	0.5684	0.4316				

[a] Not polymorphic in any study group.

Appendix 1B

IL1 gene cluster variant allele frequencies—Caucasian T2DM/ESRD

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs16347-3								
rs16347-2	1	75.0	0.7067	0.2933				
rs16347-1	1	75.0	1					
rs1516792-3	0.9333	70.0	1					
rs1516792-2a	0.9333	70.0	1					
rs1516792-1	0.9333	70.0	1					
rs17561-3	0.9867	74.0	0.7365	0.2635				
rs17561-2	0.9867	74.0	0.7432	0.2568				
rs17561-1	0.9867	74.0	0.7297	0.2703				
rs20540	0.9733	73.0	1					
rs1609682-4	0.94	70.5	0.5603	0.4397				
rs1609682-3	0.96	72.0	0.7153	0.2847				

(continued on next page)

Appendix 1B *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs1609682-2	0.96	72.0	0.7361	0.2639				
rs1609682-1	0.96	72.0	0.7083	0.2917				
SNP -889	1	75.0	0.2533	0.7467				
Ex5 at 4336 SNP-2	0.9733	73.0	0.8082	0.1918				
Ex5 at 4336 SNP-1	0.9733	73.0	0.3630	0.6370				
rs1133558-2	0.96	72.0	1					
rs1133558-1	0.9733	73.0	0.3973	0.6027				
SNP -511	0.9867	74.0	0.5946	0.4054				
Gaatp33330	0.9867	74.0	0.6486	0.2703	0	0.0811		
IL1F10-1 Ex2 5′UTR	0.9333	70.0	1					
IL1F10-2 Ex3 No code	0	0.0						
IL1F10-3 Ex4 cSNP-1	0.9867	74.0	0.6014	0.3986				
IL1F10-4 Ex4 cSNP-2	0.88	66.0	0.6742	0.3258				
IL1F10-5 Ex5 3′UTR-1	0.8667	65.0	0.7462	0.2538				
IL1F10-6 Ex5 3′UTR-2	0.9333	70.0	1					
rs315929-1	0.9333	70.0	0.5071	0.4929				
rs315929-2	0.9733	73.0	1					
rs315929-3	0.9733	73.0	0.1712	0.8288				
rs315929-4	0.9733	73.0	0.9315	0.0685				
rs315929-5	0.9733	73.0	0.6644	0.3356				
rs315929-6	0.9733	73.0	0.5890	0.4110				
rs315929-7	0.9667	72.5	0.9034	0.0966				
rs315931-1	0.9467	71.0	0.6549	0.3451				
rs315931-2	0.9533	71.5	0.3427	0.6573				
rs315932	0.96	72.0	0.9236	0.0764				
rs315933-1	0.96	72.0	0.6736	0.3264				
rs315933-2	0.96	72.0	0.9097	0.0903				
rs315921-1	0.9067	68.0	0.9265	0.0735				
rs315921-2	0.8933	67.0	0.8209	0.1791				
rs315921-3	0.92	69.0	1					
rs315921-4	0.8933	67.0	0.9030	0.0970				
1731-1934-1	0.92	69.0	0.6884	0.3116				
1731-1934-2	0.04	3.0	1					
1731-1934-3	0.9067	68.0	0.6912	0.3088				
1731-1934-4	0.88	66.0	0.6818	0.3182				
1731-1934-5	0.8933	67.0	0.6866	0.3134				
1731-1934-6	0.88	66.0	0.6894	0.3106				
1731-1934-7	0.8933	67.0	0.6866	0.3134				
rs315919-1	0.9067	68.0	0.6838	0.3162				
rs315919-2	0.8933	67.0	0.9925	0.0075				
rs315919-3	0.8933	67.0	0.6791	0.3209				
rs315919-4	0.92	69.0	0.3188	0.6812				
rs315919-5	0.9133	68.5	0.5693	0.4307				
rs315935-1	0.96	72.0	0.9931	0.0069				
rs315935-2	0.96	72.0	0.9931	0.0069				
rs315935-3	0.96	72.0	0.9792	0.0208				
Ex2 +8006-1	0.9733	73.0	0.6781	0.3219				
Ex2 +8006-2	0.9733	73.0	0.3219	0.6781				
Ex2 +8006-3	0.9733	73.0	1					
Ex2 +8006-4	0.9733	73.0	0.6781	0.3219				
Ex2 +8006-5	0.9733	73.0	0.6781	0.3219				
Ex2 +8006-6	0.9733	73.0	0.1986	0.8014				
Ex2 +8006-7	0.9667	72.5	0.6828	0.3172				
Ex2 +8006-8	0.1067	8.0	1					
Ex2 +8006-9	0.9733	73.0	0.6781	0.3219				

Appendix 1B *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
Ex2 +8006-10	0.96	72.0	0.3125	0.6875				
Ex2 +8006-11	0.9067	68.0	0.8824	0.1176				
Int2 VNTR	0.9867	74.0	0.6554	0.3176	0.0270			
rs315955-1	0.96	72.0	1					
rs315955-2	0.9467	71.0	0.7183	0.2817				
rs315955-3	0.9467	71.0	0.7183	0.2817				
rs315954-1	0.9733	73.0	0.9932	0.0068				
rs315954-2	0.9733	73.0	1					
rs314953-1	0.9733	73.0	0.7397	0.2603				
rs314953-2	0.0933	7.0	1					
rs314953-3	0.9733	73.0	1					
rs315952-1	0.9467	71.0	0.2394	0.7606				
rs315952-2	0.9733	73.0	1					
rs315952-3	0.9733	73.0	1					
rs315951-1	0.9733	73.0	0.7397	0.2603				
rs315951-2	0.9467	71.0	0.9859	0.0141				
rs315951-3	0.9733	73.0	1					
rs9005-1	0.9733	73.0	0.6918	0.3082				
rs9005-2	0.9733	73.0	1					
rs315950-1	0.9733	73.0	0.6849	0.3151				
rs315950-2	0.96	72.0	1					
rs315950-3	0.04	3.0	1					
rs315950-4	0.9467	71.0	1					
rs315950-5	0.04	3.0	1					
rs315950-6	0.04	3.0	0.5000	0.5000				
rs315949-1	0.0933	7.0	0.8571	0.1429				
rs315949-2	0.0933	7.0	1					
rs315949-3	0.0933	7.0	1					
rs315949-4	0.9733	73.0	0.5753	0.4247				

[a] Not polymorphic in any study group.

Appendix 1C

IL1 gene cluster variant allele frequencies—African American controls

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs16347-3	1							
rs16347-2	1	86.0	0.7035	0.2965				
rs16347-1	1	86.0	0.8953	0.1047				
rs1516792-3	0.9186	79.0	0.9810	0.0190				
rs1516792-2a	0.9419	81.0	1					
rs1516792-1	0.9593	82.5	0.9818	0.0182				
rs17561-3	0.9767	84.0	0.8095	0.1905				
rs17561-2	0.9826	84.5	0.8639	0.1361				
rs17561-1	0.9884	85.0	0.7765	0.2235				
rs20540	0.8895	76.5	0.9477	0.0523				
rs1609682-4	0.9767	84.0	0.6905	0.3095				
rs1609682-3	0.9767	84.0	0.8690	0.1310				
rs1609682-2	0.9767	84.0	0.6548	0.3452				
rs1609682-1	0.9767	84.0	0.7024	0.2976				
SNP −889	0.9884	85.0	0.3529	0.6471				
Ex5 at 4336 SNP-2	0.9651	83.0	0.8735	0.1265				
Ex5 at 4336 SNP-1	0.9651	83.0	0.1627	0.8373				
rs1133558-2	0.9767	84.0	1					
rs1133558-1	0.9535	82.0	0.5976	0.4024				
SNP −511	0.9884	85.0	0.4412	0.5588				
Gaatp33330	0.9651	83.0	0.8795	0.0843	0.006	0.0181	0.012	
IL1F10-1 Ex2 5′UTR	0.9302	80.0	0.9937	0.0063				

(continued on next page)

Appendix 1C *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
IL1F10-2 Ex3 No code	0.1163	10.0	0.9500	0.0500				
IL1F10-3 Ex4 cSNP-1	0.9884	85.0	0.5941	0.4059				
IL1F10-4 Ex4 cSNP-2	0.9302	80.0	0.7250	0.2750				
IL1F10-5 Ex5 3′UTR-1	0.814	70.0	0.8786	0.1214				
IL1F10-6 Ex5 3′UTR-2	0.9651	83.0	0.9458	0.0542				
rs315929-1	0.9419	81.0	0.6543	0.3457				
rs315929-2	0.9535	82.0	0.9451	0.0549				
rs315929-3	0.9477	81.5	0.1534	0.8466				
rs315929-4	0.936	80.5	0.9130	0.0870				
rs315929-5	0.9419	81.0	0.5123	0.4877				
rs315929-6	0.9535	82.0	0.7622	0.2378				
rs315929-7	0.936	80.5	0.8758	0.1242				
rs315931-1	0.936	80.5	0.5155	0.4845				
rs315931-2	0.9419	81.0	0.4877	0.5123				
rs315932	0.9419	81.0	0.9383	0.0617				
rs315933-1	0.936	80.5	0.7578	0.2422				
rs315933-2	0.936	80.5	0.9193	0.0807				
rs315921-1	0.9419	81.0	0.9321	0.0679				
rs315921-2	0.9244	79.5	0.9371	0.0629				
rs315921-3	0.936	80.5	0.9255	0.0745				
rs315921-4	0.9419	81.0	0.9259	0.0741				
1731-1934-1	0.8953	77.0	0.9416	0.0584				
1731-1934-2	0.8023	69.0	0.8333	0.1667				
1731-1934-3	0.8837	76.0	0.9408	0.0592				
1731-1934-4	0.8721	75.0	0.9533	0.0467				
1731-1934-5	0.8837	76.0	0.9408	0.0592				
1731-1934-6	0.8837	76.0	0.9408	0.0592				
1731-1934-7	0.8837	76.0	0.9408	0.0592				
rs315919-1	0.907	78.0	0.9231	0.0769				
rs315919-2	0.8953	77.0	0.8636	0.1364				
rs315919-3	0.8953	77.0	0.9416	0.0584				
rs315919-4	0.8953	77.0	0.0584	0.9416				
rs315919-5	0.8779	75.5	0.5629	0.4371				
rs315935-1	0.907	78.0	0.8654	0.1346				
rs315935-2	0.907	78.0	1					
rs315935-3	0.9012	77.5	0.9871	0.0129				
Ex2 +8006-1	0.9302	80.0	0.9312	0.0688				
Ex2 +8006-2	0.9186	79.0	0.0696	0.9304				
Ex2 +8006-3	0.9186	79.0	1					
Ex2 +8006-4	0.9186	79.0	0.9304	0.0696				
Ex2 +8006-5	0.9186	79.0	0.9304	0.0696				
Ex2 +8006-6	0.9186	79.0	0.0506	0.9494				
Ex2 +8006-7	0.8953	77.0	0.9286	0.0714				
Ex2 +8006-8	0.9128	78.5	0.9299	0.0701				
Ex2 +8006-9	0.9186	79.0	0.9304	0.0696				
Ex2 +8006-10	0.9186	79.0	0.0696	0.9304				
Ex2 +8006-11	0.1628	14.0	0.7500	0.2500				
Int2 VNTR	1	86.0	0.8895	0.0988	0	0.0058	0.0058	
rs315955-1	0.9012	77.5	0.9290	0.0710				
rs315955-2	0.8953	77.0	0.9091	0.0909				
rs315955-3	0.9012	77.5	0.8968	0.1032				
rs315954-1	0.907	78.0	0.8590	0.1410				
rs315954-2	0.907	78.0	0.9808	0.0192				
rs314953-1	0.8953	77.0	0.5584	0.4416				
rs314953-2	0.8953	77.0	0.9935	0.0065				
rs314953-3	0.8953	77.0	1					

Appendix 1C *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs315952-1	0.8779	75.5	0.4901	0.5099				
rs315952-2	0.9419	81.0	0.9506	0.0494				
rs315952-3	0.9419	81.0	0.9938	0.0062				
rs315951-1	0.907	78.0	0.5321	0.4679				
rs315951-2	0.9244	79.5	0.9811	0.0189				
rs315951-3	0.907	78.0	0.9487	0.0513				
rs9005-1	0.8953	77.0	0.8312	0.1688				
rs9005-2	0.8837	76.0	0.9803	0.0197				
rs315950-1	1	86.0	0.8895	0.1105				
rs315950-2	1	86.0	0.8547	0.1453				
rs315950-3	0	0.0						
rs315950-4	1	86.0	0.8895	0.1105				
rs315950-5	0	0.0						
rs315950-6	0	0.0						
rs315949-1	0.1512	13.0	1					
rs315949-2	0.1512	13.0	1					
rs315949-3	0.1512	13.0	1					
rs315949-4	0.9826	84.5	0.6450	0.3550				

[a] Not polymorphic in any study group.

Appendix 1D

IL1 gene cluster variant allele frequencies—African American T2DM/ESRD

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs16347-3	1	92.0	0.9946	0.0054				
rs16347-2	1	92.0	0.7554	0.2446				
rs16347-1	1	92.0	0.8478	0.1522				
rs1516792-3	0.9565	88.0	1					
rs1516792-2a	0.962	88.5	1					
rs1516792-1	0.9783	90.0	0.9944	0.0056				
rs17561-3	0.9783	90.0	0.8278	0.1722				
rs17561-2	0.9783	90.0	0.8833	0.1167				
rs17561-1	0.9783	90.0	0.7000	0.0003				
rs20540	0.9946	91.5	0.9563	0.0437				
rs1609682-4	0.962	88.5	0.6328	0.3672				
rs1609682-3	0.9783	90.0	0.9000	0.0001				
rs1609682-2	0.9783	90.0	0.6222	0.3778				
rs1609682-1	0.9783	90.0	0.7611	0.2389				
SNP −889	1	92.0	0.3533	0.6467				
Ex5 at 4336 SNP-2	0.9891	91.0	0.8956	0.1044				
Ex5 at 4336 SNP-1	0.9891	91.0	0.2033	0.7967				
rs1133558-2	1	92.0	0.9946	0.0054				
rs1133558-1	1	92.0	0.5707	0.4293				
SNP −511	0.9783	90.0	0.4333	0.5667				
Gaatp33330	0.9728	89.5	0.8883	0.0559	0	0.0279	0.0056	0.0223
IL1F10-1 Ex2 5′UTR	0.913	84.0	1					
IL1F10-2 Ex3 No code	0.1957	18.0	0.9167	0.0833				
IL1F10-3 Ex4 cSNP-1	1	92.0	0.5761	0.4239				
IL1F10-4 Ex4 cSNP-2	0.9022	83.0	0.7229	0.2771				
IL1F10-5 Ex5 3′UTR-1	0.8587	79.0	0.8797	0.1203				
IL1F10-6 Ex5 3′UTR-2	0.9348	86.0	0.9419	0.0581				

(continued on next page)

Appendix 1D *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs315929-1	0.9891	91.0	0.6429	0.3571				
rs315929-2	1	92.0	0.9402	0.0598				
rs315929-3	1	92.0	0.1359	0.8641				
rs315929-4	1	92.0	0.8804	0.1196				
rs315929-5	1	92.0	0.5000	0.5000				
rs315929-6	1	92.0	0.8098	0.1902				
rs315929-7	0.9674	89.0	0.8315	0.1685				
rs315931-1	1	92.0	0.4946	0.5054				
rs315931-2	1	92.0	0.5000	0.5000				
rs315932	1	92.0	0.9457	0.0543				
rs315933-1	1	92.0	0.7337	0.2663				
rs315933-2	1	92.0	0.9728	0.0272				
rs315921-1	0.9022	83.0	0.9578	0.0422				
rs315921-2	0.9348	86.0	0.9535	0.0465				
rs315921-3	0.9348	86.0	0.9070	0.0930				
rs315921-4	0.913	84.0	0.9048	0.0952				
1731-1934-1	0.9565	88.0	0.9432	0.0568				
1731-1934-2	0.0109	1.0	1					
1731-1934-3	0.9565	88.0	0.9432	0.0568				
1731-1934-4	0.9348	86.0	0.9535	0.0465				
1731-1934-5	0.9565	88.0	0.9432	0.0568				
1731-1934-6	0.9565	88.0	0.9432	0.0568				
1731-1934-7	0.913	84.0	0.9464	0.0536				
rs315919-1	0.962	88.5	0.9266	0.0734				
rs315919-2	0.962	88.5	0.8192	0.1808				
rs315919-3	0.9674	89.0	0.9382	0.0618				
rs315919-4	0.9674	89.0	0.0674	0.9326				
rs315919-5	0.9674	89.0	0.4831	0.5169				
rs315935-1	0.9565	88.0	0.8466	0.1534				
rs315935-2	0.9565	88.0	1					
rs315935-3	0.9457	87.0	1					
Ex2 +8006-1	1	92.0	0.9239	0.0761				
Ex2 +8006-2	1	92.0	0.0761	0.9239				
Ex2 +8006-3	1	92.0	1					
Ex2 +8006-4	1	92.0	0.9239	0.0761				
Ex2 +8006-5	1	92.0	0.9239	0.0761				
Ex2 +8006-6	1	92.0	0.0272	0.9728				
Ex2 +8006-7	1	92.0	0.9239	0.0761				
Ex2 +8006-8	1	92.0	0.9620	0.0380				
Ex2 +8006-9	1	92.0	0.9239	0.0761				
Ex2 +8006-10	1	92.0	0.0761	0.9239				
Ex2 +8006-11	0.1304	12.0	0.7083	0.2917				
Int2 VNTR	1	92.0	0.8696	0.0924	0.0217	0.0163		
rs315955-1	1	92.0	0.9239	0.0761				
rs315955-2	1	92.0	0.9022	0.0978				
rs315955-3	1	92.0	0.8804	0.1196				
rs315954-1	0.9891	91.0	0.8352	0.1648				
rs315954-2	0.9891	91.0	0.9835	0.0165				
rs314953-1	1	92.0	0.6141	0.3859				
rs314953-2	1	92.0	1					
rs314953-3	1	92.0	0.9837	0.0163				
rs315952-1	0	0.0						
rs315952-2	1	92.0	0.9239	0.0761				
rs315952-3	1	92.0	0.9891	0.0109				
rs315951-1	0.9946	91.5	0.5464	0.4536				
rs315951-2	0.9891	91.0	1					
rs315951-3	0.9891	91.0	0.9231	0.0769				
rs9005-1	1	92.0	0.7935	0.2065				
rs9005-2	0.9891	91.0	0.9835	0.0165				
rs315950-1	0.9674	89.0	0.9101	0.0899				
rs315950-2	0.9511	87.5	0.8286	0.1714				

Appendix 1D *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs315950-3	0.9565	88.0	0.9943	0.0057				
rs315950-4	0.9511	87.5	0.8400	0.1600				
rs315950-5	0.9511	87.5	0.9429	0.0571				
rs315950-6	0.9457	87.0	0.8563	0.1437				
rs315949-1	0.9946	91.5	0.9781	0.0219				
rs315949-2	1	92.0	0.9837	0.0163				
rs315949-3	1	92.0	0.9728	0.0272				
rs315949-4	1	92.0	0.6467	0.3533				

[a] Not polymorphic in any study group.

Appendix 1E

IL1 gene cluster variant allele frequencies—African American Non-T2DM/ESRD

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs16347-3	0.9896	95.0	0.9947	0.0053				
rs16347-2	0.9896	95.0	0.8316	0.1684				
rs16347-1	0.9896	95.0	0.8526	0.1474				
rs1516792-3	0.9896	95.0	1					
rs1516792-2a	0.9896	95.0	1					
rs1516792-1	0.9896	95.0	0.9684	0.0316				
rs17561-3	0.9688	93.0	0.8172	0.1828				
rs17561-2	0.9792	94.0	0.8723	0.1277				
rs17561-1	0.974	93.5	0.7059	0.2941				
rs20540	0.9323	89.5	0.9441	0.0559				
rs1609682-4	0.9688	93.0	0.5968	0.4032				
rs1609682-3	0.9792	94.0	0.9415	0.0585				
rs1609682-2	0.9792	94.0	0.6117	0.3883				
rs1609682-1	0.9792	94.0	0.8298	0.1702				
SNP −889	0.9896	95.0	0.3895	0.6105				
Ex5 at 4336 SNP-2	0.9792	94.0	0.8617	0.1383				
Ex5 at 4336 SNP-1	0.9792	94.0	0.1915	0.8085				
rs1133558-2	1	96.0	1					
rs1133558-1	0.9896	95.0	0.6105	0.3895				
SNP −511	0.9896	95.0	0.4158	0.5842				
Gaatp33330	0.9479	91.0	0.8352	0.1099	0.0165	0.0385		
IL1F10-1 Ex2 5′UTR	0	0.0						
IL1F10-2 Ex3 No code	0	0.0						
IL1F10-3 Ex4 cSNP-1	1	96.0	0.5625	0.4375				
IL1F10-4 Ex4 cSNP-2	0.8021	77.0	0.7597	0.2403				
IL1F10-5 Ex5 3′UTR-1	0.4375	42.0	0.7857	0.2143				
IL1F10-6 Ex5 3′UTR-2	0.8854	85.0	0.9235	0.0765				
rs315929-1	1	96.0	0.6302	0.3698				
rs315929-2	1	96.0	0.9323	0.0677				
rs315929-3	0.9792	94.0	0.0798	0.9202				
rs315929-4	0.9896	95.0	0.8737	0.1263				
rs315929-5	0.9479	91.0	0.4725	0.5275				
rs315929-6	0.9479	91.0	0.8242	0.1758				
rs315929-7	0.9375	90.0	0.7944	0.2056				
rs315931-1	0.9896	95.0	0.4632	0.5368				
rs315931-2	0.9688	93.0	0.5430	0.4570				
rs315932	0.9896	95.0	0.9474	0.0526				

(continued on next page)

Appendix 1E *(continued)*

SNP name	Data present	No. of people w/data	Allele frequency					
			Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
rs315933-1	0.9896	95.0	0.6947	0.3053				
rs315933-2	0.9896	95.0	0.9526	0.0474				
rs315921-1	0.9583	92.0	0.9457	0.0543				
rs315921-2	0.9427	90.5	0.9613	0.0387				
rs315921-3	0.9427	90.5	0.9448	0.0552				
rs315921-4	0.9323	89.5	0.8603	0.1397				
1731-1934-1	0.9375	90.0	0.9444	0.0556				
1731-1934-2	0	0.0						
1731-1934-3	0.9375	90.0	0.9444	0.0556				
1731-1934-4	0.8854	85.0	0.9647	0.0353				
1731-1934-5	0.9115	87.5	0.9486	0.0514				
1731-1934-6	0.8646	83.0	0.9458	0.0542				
1731-1934-7	0.8854	85.0	0.9529	0.0471				
rs315919-1	0.9896	95.0	0.9368	0.0632				
rs315919-2	0.9896	95.0	0.8053	0.1947				
rs315919-3	0.9896	95.0	0.9526	0.0474				
rs315919-4	0.9896	95.0	0.0579	0.9421				
rs315919-5	0.9896	95.0	0.4789	0.5211				
rs315935-1	0.9844	94.5	0.8571	0.1429				
rs315935-2	0.9896	95.0	1					
rs315935-3	0.9844	94.5	0.9947	0.0053				
Ex2 +8006-1	0.9792	94.0	0.9255	0.0745				
Ex2 +8006-2	0.9792	94.0	0.0745	0.9255				
Ex2 +8006-3	0.9792	94.0	0.9947	0.0053				
Ex2 +8006-4	0.9792	94.0	0.9255	0.0745				
Ex2 +8006-5	0.9792	94.0	0.9255	0.0745				
Ex2 +8006-6	0.9792	94.0	0.0106	0.9894				
Ex2 +8006-7	0.9792	94.0	0.9255	0.0745				
Ex2 +8006-8	0.9792	94.0	0.9415	0.0585				
Ex2 +8006-9	0.9792	94.0	0.9255	0.0745				
Ex2 +8006-10	0.9792	94.0	0.0745	0.9255				
Ex2 +8006-11	0.1875	18.0	0.7778	0.2222				
Int2 VNTR	0.9896	95.0	0.8842	0.1	0.0105	0.0053		
rs315955-1	0.9896	95.0	0.8895	0.1105				
rs315955-2	1	96.0	0.9062	0.0938				
rs315955-3	0.9792	94.0	0.8989	0.1011				
rs315954-1	0.9792	94.0	0.8351	0.1649				
rs315954-2	0.9896	95.0	0.9842	0.0158				
rs314953-1	0.9792	94.0	0.6117	0.3883				
rs314953-2	0.0104	1.0	1					
rs314953-3	0.9792	94.0	0.9947	0.0053				
rs315952-1	0.9792	94.0	0.4628	0.5372				
rs315952-2	0.9896	95.0	0.9368	0.0632				
rs315952-3	0.9896	95.0	0.9789	0.0211				
rs315951-1	0.9896	95.0	0.5579	0.4421				
rs315951-2	0.9896	95.0	0.9895	0.0105				
rs315951-3	0.9896	95.0	0.9368	0.0632				
rs9005-1	1	96.0	0.7917	0.2083				
rs9005-2	0.9792	94.0	0.9840	0.0160				
rs315950-1	0.9948	95.5	0.9215	0.0785				
rs315950-2	0.9896	95.0	0.8316	0.1684				
rs315950-3	0.9896	95.0	1					
rs315950-4	0.9792	94.0	0.8457	0.1543				
rs315950-5	0	0.0						
rs315950-6	0	0.0						
rs315949-1	0.25	24.0	1					
rs315949-2	0.25	24.0	1					
rs315949-3	0.25	24.0	0.9792	0.0208				
rs315949-4	0.9688	93.0	0.6667	0.3333				

[a] Not polymorphic in any study group.

Appendix 2

Gene/primer set	Name[a]	Primers		Variant/method	Temp (°C)	Size (bp)	Location
ILIA							
1	−889	F	ggcttaaactccaactggga	SNP/SSCP	56	270	Promoter
		R	ccaaggtgtgtcttctgtag				
2	rs1609682	F	tgtgacccacaactatcatgg	SNP/DS	56	646	Intron 3
		R	ctgaagacatactaggcagtac				
3	rs20540	F	agaatggagatcctcctcacta	SNP/DS	54	660	Exon 4
		R	gtaacatgaccaaggaggcata				
4	rs17561	F	actggtagtcttccttgtcagt	SNP/DS	60	694	Exon 5
		R	ctctcttgtggaacttccagat				
5	rs1516792	F	tcaatcaggttgctacgttggt	SNP/DS	56	713	Intron 5
		R	tcttcatcttgggcagtcacat				
6	rs16347	F	cttgggagacctgtaatcat	Deletion/SSCP	58	182	3′UTR
		R	gtggtctcatggttgtcaaa				
ILIB							
7	−511	F	tggcattgatctggttcatc	SNP/SSCP	53	304	Promoter
		R	gtttaggaatcttcccactt				
8	rs1133558	F	ctgcacaacgattgtcaggaaa	SNP/DS	58	671	Exon 1 5′UTR
		R	gagaagtcctttgggtctagag				
9	Ex5 + 4336 (old +3953)	F	agcagtaatagacctgaagctg	SNP/DS	58	654	Exon 5
		R	tgacattgcactatgcccaaga				
Intergenic *ILIB–ILIRN*							
10	Gaatp33330	F	gaggcgtgagaatctcaaga	SSTR/PAGE	56	175–199	Intergenic
		R	gtatcctcaagtggatctgg				
IL1F10[b]							
11	SNP 68	F	caggccaattatagacgaatgg	SNP/SSCP/S[c]	58	262	Ex2 5′UTR
		R	actcagcattcctttcctctt				
12	SNP 147	F	caaatgctcaaggtggtgattc	SNP/SSCP	58	254	Ex3 no change
		R	ggaaagggaagaagagaaggg				
13	SNP 206	F	ttccttcctgcctgagccttt	SNP/SSCP/S[c]	58	295	Ex4 I to T
		R	agataggcacaaagatgccaga				
14	SNP 227	F	Same as SNP 206	SNP/SSCP/S[c]	58	295	Ex4 D to A
		R	Same as SNP 206				
15	SNP 574	F	tagggagacaggaaactgcgt	SNP/SSCP/S[c]	58	272	Ex5 3′UTR
		R	cttccattctctctctgactct				
16	SNP621	F	Same as SNP 574	SNP/SSCP/S[c]	58	272	Ex5 3′UTR
		R	Same as SNP 574				
IL1RN							
17	rs315929	F	ccatagtctcaagaggtcaca	SNP/DS	58	620	Promoter
		R	gaagcattcagaccatgggatt				
18[d]	rs315931	F	aatcccatggtctgaatgcttc	SNP/DS	58	650	Promoter
	rs315932						Promoter
	rs315933						Promoter
		R	tcatagtcttcacacctccaga				
19	rs315921	F	ctgcctctatctgtttgcttgt	SNP/DS	58	662	Promoter
		R	taagacctcagctcagactctg				
20[d]	SNP 1731	F	gaatgtgtgcacacatgcatga	SNP/DS	58	669	Promoter
	SNP 1812						Ex lic 5′UTR[e]
	SNP 1868						Exon lic 5′UTR
	SNP 1887						Exon lic 5′UTR
	SNP 1934						Intron lic long
		R	ttctgtcacagcagttacacct				
21	rs315919	F	tgtgagggtcattttccactgt	SNP/DS	58	647	Intron lic long
		R	agcaattctcctgccttagcta				
22	rs315935	F	ctcctctgaatgatctcaagtc	SNP/DS	58	646	Intron 11ic long
		R	tgtatctagtggcttccatgtg				
23	Ex2 +8006 long	F	tctcagatgggaagcaagtaag	SNP/DS	57	635	Exon 2ic long

(continued on next page)

Appendix 2 *(continued)*

Gene/primer set	Name[a]	Primers		Variant/method	Temp (°C)	Size (bp)	Location
		R	aatggtcaaccttgccttctgg				
24	Int2 VNTR	F	ctcagcaacactcctat	VNTR/EF	50	240 to 595	Intron 2ic long
		R	tcctggtctgcaggtaa				
25	rs315955	F	tgtgagaggagactgcataaga	SNP/DS	58	620	Intron 3ic long
		R	ccaggtttcttaggacagcaaa				
26[e]	rs315954	F	cagagctcaaaccctgacagaa	SNP/DS	60	681	Intron 3ic long
	rs315953						Intron 3ic long
		R	ggtcagtgatgttaactgcctg				
27[d]	rs315952	F	agaacagaaagcaggacaagcg	SNP/DS	58	563	Exon 4ic long
	rs315951						Exon 4ic
							3′UTR
		R	atcaagtggcctgatggatcca				
28	rs9005	F	taatctggactcctctgtccag	SNP/DS	58	654	3′ genomic
		R	ttccagtagaggctgttgtgga				
29	rs315950	F	ggtatagagtgctgaggaaact	SNP/DS	58	662	3′ genomic
		R	ctcataatttcccagatgctcc				
30	rs315949	F	ttgaatttgtccccataggtgg	SNP/DS	58	676	3′ genomic
		R	aagctttgatgagttttctgcct				

Note. DS, direct sequencing; EF, agarose electrophoresis; L, literature— polymorphism identified through previous publications; PAGE, polyacrylamide gel electrophoresis.

[a] When available NCBI dbSNP reference SNP (rs) names were used; however, primer sets do not correspond to those provided in the NCBI SNP database.

[b] Novel gene *IL1F10* SNPs were identified by SSCP exon screening and then genotyped using Sequenom.

[c] IL1 receptor antagonist has 3 splice variants; ic, intracellular variant.

[d] Due to larger amplimers used in direct sequencing, multiple SNPs were amplified with a single primer set.

[e] Sequenome amplification primers are different from the SSCP primers listed here (see Table 1). S, Sequenom MassARRAY method.

References

[1] U.S. Renal Data System. USRDS 2002 Annual Data Report. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (NIH), DHHSAvailable at www.usrds.org.2002. Accessed May 29, 2003.

[2] K. Hiromura, M. Kurosawa, S. Yano, T. Naruse, Tubulointerstitial mast cell infiltration in glomerulonephritis, Am. J. Kidney Dis. 32 (1998) 593–599.

[3] T.A. Jacot, G.E. Striker, M. Stetler-Stevenson, L.J. Striker, Mesangial cells from transgenic mice with progressive glomerulosclerosis exhibit stable, phenotypic changes including undetectable MMP-9 and increased type IV collagen, Lab. Invest. 75 (1996) 791–799.

[4] W. Eberhardt, et al., Nitric oxide modulates expression of matrix metalloproteinase-9 in rat mesangial cells, Kidney Int. 57 (2000) 59–69.

[5] A.I. Blakemore, et al., Interleukin-1 receptor antagonist allele (IL1RN*2) associated with nephropathy in diabetes mellitus, Hum. Genet. 97 (1996) 369–374.

[6] B.I. Freedman, et al., Genetic linkage analysis of growth factor loci and end-stage renal disease in African Americans, Kidney Int. 51 (1997) 819–825.

[7] B.V. Loughrey, et al., An interleukin-1B allele, which correlates with a high secretor phenotype, is associated with diabetic nephropathy, Cytokine 10 (1998) 984–988, doi:10.1006/cyto.1998.0388.

[8] J. T. Bensen et al., Association of an IL1A 3′UTR polymorphism with end stage renal disease and IL1α expression. Kidney Int. 63 (2003) 1211–1219.

[9] L. Tarnow, et al., Polymorphisms in the interleukin-1 gene cluster do not contribute to the genetic susceptibility of diabetic nephropathy in Caucasian patients with IDDM, Diabetes 46 (1997) 1075–1076.

[10] L.B. Jorde, W.S. Watkins, J. Kere, D. Nyman, A.W. Eriksson, Gene mapping in isolated populations: new roles for old friends? Hum. Hered. 50 (2000) 57–65.

[11] K.K. Kidd, et al., A global survey of haplotype frequencies and linkage disequilibrium at the DRD2 locus, Hum. Genet. 103 (1998) 211–227.

[12] C.G. Kunst, et al., FMR1 in global populations, Am. J. Hum. Genet. 58 (1996) 513–522.

[13] C. Lonjou, et al., Allelic association under map error and recombinational heterogeneity: a tale of two sites, Proc. Natl. Acad. Sci. USA 95 (1998) 11366–11370.

[14] S.M. Purandare, et al., Genotyping of PCR-based polymorphisms and linkage disequilibrium analysis at the NF1 locus, Am. J. Hum. Genet. 59 (1996) 159–166.

[15] S.A. Tishkoff, et al., Global patterns of linkage disequilibrium at the CD4 locus and modern human origins, Science 271 (1996) 1380–1387.

[16] S.A. Tishkoff, et al., A global haplotype analysis of the myotonic dystrophy locus: implications for the evolution of modern humans and for the origin of myotonic dystrophy mutations, Am. J. Hum. Genet. 62 (1998) 1389–1402.

[17] J.K. Pritchard, M. Przeworski, Linkage disequilibrium in humans: models and data, Am. J. Hum. Genet. 69 (2001) 1–14.

[18] G.A. Huttley, M.W. Smith, M. Carrington, S.J. O'Brien, A scan for linkage disequilibrium across the human genome, Genetics 152 (1999) 1711–1722.

[19] L.B. Jorde, et al., Linkage disequilibrium predicts physical distance in the adenomatous polyposis coli region, Am. J. Hum. Genet. 54 (1994) 884–898.

[20] D.A. Nickerson, et al., DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene, Nat. Genet. 19 (1998) 233–240.

[21] A. Cox, N.J. Camp, M.J.H. Nicklin, F.S. di Giovine, G.W. Duff, An analysis of linkage disequilibrium in the interleukin-1 gene cluster, using a novel grouping method for multiallelic markers, Am. J. Hum. Genet. 62 (1998) 1180–1188.

[22] B. Ewing, L. Hillier, M.C. Wendl, P. Green, Base-calling of automated sequencer traces using Phred, I. Accuracy assessment. Genome Res 8 (1998) 175–185.

[23] B. Ewing, P. Green, Basecalling of automated sequencer traces using Phred, II. Error probabilities. Genome Res. 8 (1998) 186–194.
[24] D.A. Nickerson, V.O. Tobe, S.L. Taylor, PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing, Nucleic Acids Res. 25 (1997) 2745–2751.
[25] K.H. Buetow, et al., High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, Proc. Natl. Acad. Sci. USA 98 (2001) 581–584.
[26] K. Lange, Mathematical and statistical methods for genetic analysis, Springer, New York, 1997.
[27] B.S. Weir, Genetic data analysis II, Sinauer Associates, Inc., Sunderland, MA, 1996.
[28] D. Fallin, N.J. Schork, Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data, Am. J. Hum. Genet. 67 (2000) 947–59.
[29] M. Stephens, N.J. Smith, P. Donnelly, A new statistical method for haplotype reconstruction from population data, Am. J. Hum. Genet. 68 (2001) 978–989.
[30] D.A. Roff, P. Bentzen, The statistical analysis of mitochondrial DNA polymorphisms: χ and the problem of small samples, Mol. Biol. Evol. 6 (1989) 539–545.
[31] J.S. Waye, B. Eng, H.Y. Ni, M.A. Blajchman, G. Carmody, Trinucleotide repeat polymorphism within the human antithrombin gene (AT3): allele frequency data for three population groups, Mol. Cell. Probes 8 (1994) 149–154, doi:10.1006/mcpr.1994.1020.