Am. J. Hum. Genet. 57:1445–1456, 1995

Evolution of Haplotypes at the DRD2 Locus

C. M. Castiglione,[1] A. S. Deinard,[1] W. C. Speed,[1] G. Sirugo,[1] H. C. Rosenbaum,[2] Y. Zhang,[3,4] D. K. Grandy,[3] E. L. Grigorenko,[1] B. Bonne-Tamir,[5] A. J. Pakstis,[1] J. R. Kidd,[1] and K. K. Kidd[1]

Departments of [1]Genetics and [2]Biology, Yale University School of Medicine; [3]Vollum Institute for Advanced Biomedical Research and Department of Cell Biology and Anatomy, Oregon Health Sciences University, Portland; [4]Department of Anatomy, University of California, San Francisco, San Francisco; and [5]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv

Summary

We present here the first evolutionary perspective on haplotypes at DRD2, the locus for the dopamine D_2 receptor. The dopamine D_2 receptor plays a critical role in the functioning of many neural circuits in the human brain. If functionally relevant variation at the DRD2 locus exists, understanding the evolution of haplotypes on the basis of polymorphic sites encompassing the gene should provide a powerful framework for identifying that variation. Three DRD2 polymorphisms (*TaqI* "A" and "B" RFLPs and the $(CA)_n$ short tandem repeat polymorphism) encompassing the coding sequences have been studied in 15 populations; these markers are polymorphic in all the populations studied, and they display strong and significant linkage disequilibria with each other. The common haplotypes for the two *TaqI* RFLPs are separately derived from the ancestral haplotype but predate the spread of modern humans around the world. The knowledge of how the various haplotypes have evolved, the allele frequencies of the haplotypes in human populations, and the physical relationships of the polymorphisms to each other and to the functional parts of the gene should now allow proper design and interpretation of association studies.

Introduction

Because the dopamine D_2 receptor has an important role in the functioning of the CNS and has a high affinity for many antipsychotic drugs, neuroscientists, psychiatrists, and geneticists are very interested in DRD2, the gene for this receptor. Following the cloning of the gene (Grandy et al. 1989*a*, 1989*b*; Dal Toso et al. 1989; Araki et al. 1992), several polymorphisms were identified in the DNA encompassing the coding sequences of this gene (Bolos et al. 1990; Hauge et al. 1991; Parsian et al. 1991; Sarkar et al. 1991; Seeman et al. 1993; Suarez et al. 1994). These polymorphisms were used to map the locus to the long arm of chromosome 11 (Gelernter et al. 1989; Grandy et al. 1989*b*; Litt et al. 1995) and for linkage studies that have generally excluded this locus as causative for several neuropsychiatric disorders (e.g., Gelernter et al. 1990, 1994). Several amino acid variants, some frequent enough to be polymorphisms, have been identified, but none have been shown to alter function (Itokawa et al. 1993; Gejman et al. 1994). The role of genetic variation at this locus in susceptibility to alcoholism and other addictive disorders is highly controversial and is currently under intense scrutiny (see Gejman et al. 1994, for review).

The most studied polymorphisms have been two RFLPs that flank the coding region of the gene: the *TaqI* "B" system (Hauge et al. 1991) upstream of the initiation codon and the *TaqI* "A" system (Grandy et al. 1989*a*) downstream from the polyadenylation signal; and one short tandem repeat polymorphism (STRP), a $(CA)_n$ dinucleotide repeat located in intron 2 (Hauge et al. 1991). Studies of linkage disequilibrium among these and other polymorphisms at DRD2 are few. The three studies that included both the *TaqI* "B" and "A" sites (Hauge et al. 1991; O'Hara et al. 1993; Suarez et al. 1994) found very similar haplotype frequencies for the samples of mixed European ancestry. By standard measures, linkage disequilibrium between alleles at these sites was highly significant in all studies and at or near the maximum possible in light of the allele frequencies. In the one study of U.S. blacks (O'Hara et al. 1993) haplotype frequencies were somewhat different but linkage disequilibrium was also highly significant and near the maximum possible. Random genetic drift could explain the differences, but that gives little insight into how those haplotypes arose by mutation and/or recombination or which haplotypes might be associated with any functional differences. Hauge et al. (1991) found no significant linkage disequilibrium between the $(CA)_n$ STRP and the *TaqI* "A" RFLP; they did not have data to test directly but surmised that no linkage disequilibrium would exist between the STRP and the *TaqI* "B" RFLP, since the alleles at the two RFLPs were so highly correlated. Among the possible interpretations of their analyses was that the STRP alleles were randomized within

Received May 17, 1995; accepted for publication August 23, 1995.

Address for correspondence and reprints: Dr. Kenneth K. Kidd, Department of Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520-8005.

0002-9297/95/5706-0023$02.00

the encompassing frameworks defined by the alleles at the two RFLP sites, possibly as a result of a high mutation rate at the STRP.

Suarez et al. (1994) also studied three other RFLPs located between the *Taq*I "B" and "A" sites; they found strong linkage disequilibrium all across the region in their sample of mixed European ancestry. The only other study of linkage disequilibrium was the work of Goldman et al. (1993), who typed the *Taq*I "A" and the 3′ SSCP polymorphisms in three populations—Finns, American Caucasians, and Cheyenne Amerindians. Allele frequencies varied among the three populations, but the linkage disequilibrium was significant in all three. Thus, linkage disequilibrium has been found in all studies, but little information exists for non-European populations and the explanation for the disequilibrium is not clear.

We present here an evolutionary perspective on DRD2 haplotypes on the basis of two *Taq*I RFLPs that flank the coding sequences—*Taq*I "B" and *Taq*I "A"—and the $(CA)_n$ dinucleotide STRP located between them in intron 2. In the course of our research we have generated a more precise mapping of both the *Taq*I "B" RFLP and the STRP, developed new PCR primers to facilitate their typing, collected allele and haplotype frequencies in several diverse human populations, and identified the ancestral allele at the two RFLPs by PCR typing of several other hominoid species. The similar pattern of disequilibrium seen in all non-African populations is most simply explained by random genetic drift at the time modern humans migrated out of Africa.

Material and Methods

Sequencing

Three different templates were used for DNA sequencing: (1) for the region encompassing the STRP, the Sau 3A subclone 1:90 (Hauge et al. 1991); (2) for the region encompassing *Taq*I "B," our 5.1-kb *Eco*RI-*Bam*HI subclone of cosmid 24,25; and (3) to test sequences for restriction sites, PCR product amplified from genomic DNA of several different individuals of known phenotype. All sequence was generated using the *Taq* Dye-Deoxy Terminator cycle sequencing kit and protocol (part 901497, revision E; Applied Biosystems 1993), with the following modification: a 30-s denaturing step was used to ensure proper priming. The product was purified in Sephadex G50 spin columns (Centrisep) and was electrophoresed on an ABI 373A DNA Sequencer. Primers used were standard vector primers or were designed from initial sequence using the program PRIMER written by S. E. Lincoln, M. J. Daly, and E. S. Lander at MIT's Whitehead Institute (unpublished computer program).

Table 1

Primers Employed for Generating Sequencing Template and for Polymorphism Typing

Primer	Sequence
DRD2REV1	5′ GGA GGG CGG TGC GT 3′
DRD2FOR	5′ TGT CCT TGG TAG GAT GGG A 3′
DRD2EX3F	5′ ATG ATG TGC ACG GCG AGC 3′
DRD2BACK1	5′ GAT ACC CAC TTC AGG AAG TC 3′
DRD2DWST3	5′ GAT GTG TAG GAA TTA GCC AGG 3′

PCR Conditions and Typing

DNA preparation.—DNA was purified from cultured cells by using standard proteinase-phenol-chloroform extraction and ethanol precipitation (Sambrook et al. 1989).

$(CA)_n$ STRP.—As an alternative to the original primers published for the CA repeat, we substituted the primer DRD2FOR (see table 1) for primer 509 and used it in conjunction with DRD2REV1, shorter than but contained within the original DRD2.419. This "new" primer pair, which results in a longer fragment (~120 bp), worked better in our hands for typing the STRP, but the more extensive sequence now available on either side allows many other primer pairs to be considered. Quite standard PCR conditions similar to those in Grandy et al. (1993) produced readily interpretable patterns. Alleles were read against an allelic ladder run on the same gel; the allelic ladder had previously been standardized against an M13 sequencing ladder and PCR product from the 1:90 clone.

TaqI "B" RFLP.—Two hundred nanograms of genomic DNA as template were PCR amplified in a volume of 25 μl, with 10 pmol of primers DRD2BACK1 and DRD2DWST3 (table 1) under standard conditions as described above. After amplification, whole reactions were digested in a total volume of 30 μl, with 6–8 U of *Taq*I restriction enzyme (New England Biolabs) at the manufacturer's recommended conditions. Following electrophoresis in a 1.5% agarose gel and EtBr staining, *Taq*I "absent" (B1) and *Taq*I "present" (B2) alleles were visualized as fragments of 459 bp (*Taq*I "absent") and of 267 bp and 192 bp (*Taq*I "present").

In addition, the DRD2BACK1-DRD2DWST3 PCR product from genomic DNA or from a clone was used as a probe in standard Southern blotting on preexisting *Taq*I filters to type some individuals in some populations. When genomic template was used to generate the probe, a small aliquot from the first amplification was used for a second amplification. The product of the second amplification was labeled with ^{32}P (Feinberg and Vogelstein 1983). This procedure minimizes back-

Table 2

Description of Population Samples

Population	No. of Cell Lines	Source	Citations
Biaka	68	L. Cavalli-Sforza	Cavalli-Sforza et al. 1986; Bowcock et al. 1987, 1991*a*, 1991*b*; Vigilant et al. 1989; Mountain et al. 1992
Mbuti	39	L. Cavalli-Sforza	Cavalli-Sforza et al. 1986; Bowcock et al. 1987, 1991*a*, 1991*b*; Vigilant et al. 1989; Mountain et al. 1992
Yemenites	47	B. Bonne-Tamir	Barr and Kidd 1993; Lichter et al. 1993
Druze	61	B. Bonne-Tamir	Hirschberg 1969
Danes	46	J. Parnas	Description not yet published.
Finns	35	L. Peltonen	Description not yet published.
Adygei	53	E. Grigorenko	Description not yet published.
Europeans (mixed)	49	K. Kidd	Kidd et al. 1991
Han(S) Chinese	49	L. Cavalli-Sforza and K. Kidd	Cavalli-Sforza et al. 1986; Bowcock et al. 1987, 1991*a*, 1991*b*; Balazs et al. 1992; Mountain et al. 1992
Han(T)	46	R. Lu and H. Ko	Lu et al. 1995
Maya	53	K. Weiss	Giuffra et al. 1990; Kidd et al. 1991; Torroni et al. 1992; Horai et al. 1993
Rondonian Surui	48	F. Black	Kidd et al. 1991; Lichter et al. 1993
Karitiana	55	F. Black	Kidd et al. 1991; Kidd et al. 1993
Ami	41	R. Lu and H. Ko	Lu et al. 1995
Atayal	42	R. Lu and H. Ko	Lu et al. 1995

ground from the total genomic sequence still in the initial reaction mixture without gel purifying large amounts of the specific PCR product.

TaqI "A" RFLP.—The majority of the samples were already typed using Southern blotting by Barr and Kidd (1993) as described by them; additional samples were typed by PCR using the primers and methods described by Grandy et al. (1993).

Populations Studied

Samples from 12 populations were selected for an initial survey of allele and haplotype frequencies. Most of these samples have been described elsewhere. Table 2 is a list of the populations, the sample sizes (N = number of cell lines established), sources of the samples, and citations providing background and/or referring to previously published research utilizing these same samples.

In brief, the Biaka are Aka-speaking Pygmies from the Central African Republic; they are considerably admixed with neighboring Bantu farmers. The Mbuti are Nilosaharan- and Niger Kordofanian–speaking Pygmies from the Ituri Forest of Zaire. The Yemenite Jews were living in Yemen from the time of the diaspora and have relatively recently immigrated to Israel. The Druze are from several villages of the Galilee of Israel, and the sample includes members of a very large kindred. Data on the large sibships were utilized to infer many otherwise ambiguous haplotypes in the parents; haplotype frequencies are based on only those parents and other unrelated individuals. This specific sample of Druze comes from towns other than those studied by Nevo et al. (1992) for classical markers. The Finns are unrelated individuals from Finland who were not of Swedish origin. The Adygei are Adygei-speaking people from the Krasnador region of the Adygei Republic in southeastern European Russia (Caucasus mountains region). They are culturally and linguistically distinct from neighboring Russians. The mixed European sample is composed of individuals of European ancestry living in the United States and Canada who are unrelated, "peripheral" members of large kindreds. Family typings are available on biologically related members of these kindreds and have been used to resolve otherwise ambiguous phenotypes. The Danes are a random sample of Danish-surname donors to the Copenhagen Blood Bank. The Han(S) are unrelated Chinese born in China and currently living in San Francisco or New Haven. The Maya are Yucatec speakers from a remote village, Xmaben, in central Yucatan. The Rondonian Surui speak a Tupi language (Surui) and are from southeast Rondonia, Brazil. The Karitiana are a very small tribal group speaking another Tupi language, Ariken (or Karitiana). In addition, DRD2 haplotype frequencies for three Taiwanese populations have been reported elsewhere (Lu et al. 1995) but are included here to provide a more complete representation of worldwide frequencies and disequilibrium. The Ami and Atayal are aboriginal populations from eastern Taiwan and the Han(T) are descendants of the original Chinese who immigrated to Taiwan from Fukien province nearly 400 years ago.

All these population samples exist as Epstein-Barr vi-

rus–transformed, lymphoblastoid cell lines (Anderson and Gusella 1984), most of which were established by us under approved human subjects protocols. The Corielle Institute for Medical Research (NIGMS Human Genetic Mutant Cell Line Repository) in Camden, New Jersey has available for distribution 5–10 cell lines from nine of the populations in this study: Ami, Atayal, Biaka, Mbuti, Druze, Han(S), Maya, Karitiana, and R. Surui.

Data Analyses

Gene and haplotype frequencies.—Allele frequencies for the individual systems were estimated by simple gene counting because each has simple codominant inheritance. The haplotype frequencies were estimated using the program HAPLO (Hawley and Kidd 1995), which implements the EM algorithm (Dempster et al. 1977) for frequency estimation and the jackknife procedure to estimate the standard error. HAPLO allows incorporation of data on some individuals with missing data for any one system but with typings done on the other two systems. Known genotypes for individuals can also be incorporated. The program also calculates a likelihood ratio χ^2 that is an overall test for the presence of linkage disequilibrium at the locus, comparing the likelihoods of the observed phenotypes based on the estimated haplotype frequencies with that based on expected haplotype frequencies corresponding to equilibrium.

Disequilibrium.—Various pairwise disequilibrium coefficients were calculated by the program LINKDISEQ (A. J. Pakstis, unpublished computer program). The measure presented is the standardized D' value first proposed by Lewontin (1964). D' is calculated by dividing the raw D coefficient by either the minimum or the maximum value possible, depending on whether the sign of D is negative or positive, respectively. Devlin and Risch (in press) have shown this is less sensitive than some other measures to changes in marginal allele frequencies and thus can be used for comparisons across populations.

Results

Mapping and Sequencing

TaqI "B".—A 5.1-kb *Eco*RI-*Bam*HI subclone from the cosmid 24,25 (a gift from G. Evans and J. Eubanks) recognizes the *Taq*I "B" polymorphism by Southern blotting (data not shown). (We have deposited the encompassing 5.6-kb *Eco*RI-*Eco*RI subclone with the American Type Culture Collection.) The cosmid and consequently the subclones derive from the allele lacking the *Taq*I "B" recognition sequence at the polymorphic site. The precise sequence difference between alleles was determined by sequencing PCR product from individuals homozygous for the alternative alleles. Sequence from an individual homozygous for absence of the recognition sequence was identical to sequence from the clone. Hauge et al. (1991) mapped the *Taq*I "B" RFLP upstream of exon 2 and the STRP into intron 2 of the DRD2 gene but did not give precise physical locations. We can now locate the polymorphic *Taq*I "B" site 0.91 kb upstream from the first codon (*Bam*HI site) in exon 2 (fig. 1). Primers DRD2BACK1 and DRD2DWST3 (table 1) were designed to amplify the region encompassing the *Taq*I "B" site and to type the polymorphism as described. The entire 1.1 kb of relevant sequence from DRD2BACK1 to the exon 2 *Bam*HI site has been deposited with GenBank (accession number U23026). The sequence deposited is that in the clone; the polymorphic base is the "A" at position 192, which is a "G" in the alternative (*Taq*I site present) allele.

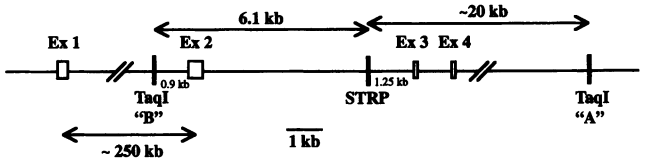

Figure 1 Molecular locations of polymorphisms at DRD2. Precise physical distances from exon 2 to the polymorphic *Taq*I site and from exon 3 to the STRP, on the basis of new subclones and genomic DNA sequence across the polymorphic sites. The *Taq*I "A" system has been previously mapped and converted to PCR (Grandy et al. 1989*a*, 1993).

STRP.—PCR was attempted with a primer designed from published sequence pointing upstream from Exon 3 (DRD2EX3F) and each of the original primers (DRD2.419 and DRD2.509) flanking the STRP. Primers DRD2.509 and DRD2EX3F did not consistently amplify any product on any template under any set of conditions. Primer DRD2.419 and DRD2EX3F amplify a fragment slightly smaller than 1.3 kb on human genomic DNA and on the phage clone λhD2G2. This PCR product and clone 1:90 (Hauge et al. 1991) served as templates for sequencing by primer walking, which yielded 1.65 kb of contiguous sequence, mapping the $(CA)_n$ repeat 1.25 kb upstream of exon 3 (fig. 1). The full 1.65 kb of sequence, including the full 467 bp sequence of clone 1:90, which encompasses the $(CA)_n$ repeat, and the sequence between this clone and exon 3, has been deposited with GenBank (accession number U23027). The sequenced clone has the 16-repeat allele. The previously published intergenic sequence of the STRP (Hauge et al. 1991; GenBank accession number X54392) corresponds to the complement of positions 221–358. The sequence is oriented in the genome with primer DRD2.419 at the upstream end.

In addition, examination of the sequence and digestion of various genomic fragments amplified across the region makes it clear that the *Taq*I "D" polymorphism (Parsian et al. 1991; Suarez et al. 1994) is also located

in intron 2 but upstream of this STRP. The PCR product of primers DRD2FOR and DRD2REV1 can be used as a probe to type the *Taq*I "D" polymorphism by Southern blotting (data not shown).

Allele Frequencies

Data were collected for the *Taq*I "B" system using primarily the PCR-based methods described. Data were collected as described on the STRP for all population samples. Most of the individuals in some of these samples had already been typed for the *Taq*I "A" system by standard RFLP analysis by Barr and Kidd (1993); additional typings were done by PCR using the primers of Grandy et al. (1993). Some individuals were typed by both methods, and the results were completely concordant (data not shown). The frequencies of the alleles at the three systems are given in table 3. STRP alleles both larger and smaller than those reported by Hauge et al. (1991) were detected. (For future comparative purposes we note the phenotypes for the STRP in terms of repeat number for the following CEPH family members: 1331-01 is 15/16; 884-01 is 14/15; 1331-13 is 16/16; 1331-14 is 15/15; and 1331-15 is 13/16. Individuals 884-01 and 1331-15 can serve to generate an allelic ladder.)

Haplotypes and their standard errors are given in table 4; linkage disequilibrium values are given in table 5. Tables 4 and 5 also include the haplotype data on these three systems originally analyzed with regard to alcoholism in three Taiwanese populations (Lu et al. 1995); there was no association of DRD2 haplotypes with alcoholism in these populations, and the total data for each population are presented here for comparison of the haplotype frequencies and linkage disequilibrium analyses with the other populations.

Primate Alleles

The PCR primers for both the *Taq*I "B" and "A" sites were used to amplify samples of genomic DNA from five chimpanzees (three common and two pygmy), two gorillas, and two orangutans, and one gibbon using the same protocols as for human samples. All reactions produced single products of the same size as the respective human product. The *Taq*I "B" products of all the apes were cut by *Taq*I into fragments of the same sizes as those of the human B2 allele; none of the *Taq*I "A" products was cut at the site polymorphic in humans (data not shown). Since it is exceedingly improbable that "identical" polymorphisms or independent parallel mutations would exist in four hominoid genera at restriction sites in noncoding sequence, we infer the ancestral primate alleles are B2 and A1 at the respective sites. Based on this information, figure 2 shows schematically how the human haplotypes are related, assuming no recurrent mutation.

Unrelated individuals from four species of nonhuman primates were tested with the STRP primers. These included 27 common chimpanzees, 7 pygmy chimpanzees, 8 gorillas, and 8 orangutans. No product was obtained with these primers on any orangutan sample, suggesting some sequence difference occurs at one of the primer sites since these DNA samples do give good PCR product with primer pairs at other loci. All eight gorilla samples gave product of a single size corresponding to only one or two copies of the CA repeat when compared to a human allelic ladder, but we have not determined the nature of the region by sequencing. In both common and pygmy chimpanzees the fragment was polymorphic. The alleles in the common chimpanzees corresponded in size to 6 repeats (6 chromosomes), 7 repeats (21 chromosomes), 9 repeats (13 chromosomes), 11 repeats (13 chromosomes), and 12 repeats (1 chromosome), while in the pygmy chimpanzees the alleles corresponded to 11 repeats (13 chromosomes) and 12 repeats (1 chromosome) when compared to the human allelic ladder.

Discussion

Mapping and Sequencing

Precise physical mapping of the polymorphisms in the DRD2 region is important for understanding the patterns of linkage disequilibrium seen at this locus and interpreting those patterns with respect to any possible functional relevance (Templeton and Sing 1993; Jorde et al. 1994; Watkins et al. 1994). The *Taq*I "B" and STRP polymorphisms are separated by just over 6 kb, including exon 2 and almost 5 kb of intron 2. The *Taq*I "A" polymorphism is ~20 kb downstream from the STRP (fig. 1). The common, expressed polymorphism, located at amino acid 311 (Itokawa et al. 1993; Gejman et al. 1994), is in the middle of exon 7 ~5 kb downstream from the STRP and 11 kb upstream from the *Taq*I "A" site. The SSCP marker in exon 8 is just over 2 kb closer to the *Taq*I "A" site (Bolos et al. 1990).

Our sequence for the 5′ untranslated portion of exon 2 and the adjacent downstream end of intron 1 (Genbank accession number U23026) differs slightly from some previously published sequences (Gandelman et al. 1991; Araki et al. 1992). The sequence we obtained is based on several runs from both directions on the clone; only a few differences existed among those runs and could easily be reconciled. Thus, we are confident of our sequence and do not know to what degree the differences from some of the other published sequences represent sequencing errors, polymorphisms, or cloning artifacts.

Allele Frequency Variation

Table 3 shows that all three systems are polymorphic in all populations studied (except that the *Taq*I "B" site

Table 3

Allele Frequencies and Standard Errors for Three DRD2 Polymorphisms

	*Taq*I "B"				STRP								*Taq*I "A"			
	Allele				Allele								Allele			
	B1	B2	2*N*	Heterozygosity	17	16	15	14	13	12	2*N*	Heterozygosity	A1	A2	2*N*	Heterozygosity
	A. Frequency															
Biaka	.07	.93	136	.13	.01	.04	.09	.62	.18	.06	130	.57	.27	.73	132	.39
Mbuti	0	1.00	66	0	0	.03	.06	.52	.39	0	66	.57	.30	.70	64	.42
Druze	.08	.92	106	.15	0	.17	.46	.14	.23	0	110	.69	.11	.89	102	.20
Yemenites	.07	.93	88	.13	0	.06	.51	.21	.22	0	86	.64	.12	.88	92	.21
Danes	.12	.88	78	.21	0	.23	.52	.12	.13	0	84	.65	.16	.84	76	.27
Finns	.25	.75	56	.38	0	.20	.39	.30	.11	0	46	.71	.24	.76	54	.36
Adygei	.12	.88	96	.21	0	.20	.51	.12	.17	0	100	.66	.17	.83	96	.28
Europeans	.13	.87	78	.23	0	.15	.53	.17	.14	.01	80	.65	.13	.87	68	.23
Han(S)	.47	.53	40	.51	0	.51	.04	.38	.07	0	74	.59	.42	.58	64	.49
Maya	.71	.29	98	.41	.05	.63	.01	.26	.04	.01	96	.53	.69	.31	90	.43
Surui	.65	.35	94	.46	0	.58	.01	.35	.06	0	88	.54	.56	.44	88	.49
Karitiana	.62	.38	106	.47	0	.59	0	.41	0	0	106	.48	.60	.40	98	.48
	B. Standard Error															
Biaka	.02				.01	.02	.03	.04	.03	.02			.04			
Mbuti	0				0	.02	.03	.06	.06	0			.06			
Druze	.03				0	.04	.05	.03	.04	0			.03			
Yemenites	.03				0	.03	.05	.04	.04	0			.03			
Danes	.04				0	.05	.05	.04	.04	0			.04			
Finns	.06				0	.06	.07	.07	.05	0			.06			
Adygei	.03				0	.04	.05	.03	.04	0			.04			
Europeans	.04				0	.04	.06	.04	.04	.01			.04			
Han(S)	.08				0	.06	.02	.06	.03	0			.06			
Maya	.05				.02	.05	.01	.04	.02	.01			.05			
Surui	.05				0	.05	.01	.05	.02	0			.05			
Karitiana	.05				0	.05	0	.05	0	0			.05			

NOTE.—Data on the Danes, Finns, Adygei, and mixed Europeans are completely new. The data on the other populations are new for the *Taq*I "B" system and the STRP. The data for the *Taq*I "A" system for those populations are updated, but the frequencies remain largely the same as reported elsewhere (Barr and Kidd 1993). The B1 and A1 alleles correspond to the *absence* of the respective *Taq*I sites. Alleles at the STR system are labeled as the number of dinucleotide repeats.

Table 4

DRD2 Haplotype Frequencies and Standard Errors for 15 Populations Based on Three Polymorphisms: *Taq*I "B," STRP, and *Taq*I "A"

	B1 17 A1	B1 16 A1	B1 15 A1	B1 14 A1	B1 13 A1	B1 12 A1	B2 16 A1	B2 15 A1	B2 14 A1	B2 13 A1	B2 12 A1	B1 16 A2	B1 15 A2	B1 14 A2	B1 13 A2	B2 17 A2	B2 16 A2	B2 15 A2	B2 14 A2	B2 13 A2	N^a	$\chi^{2\,b}$	df	Heterozygosity
A. Frequency																								
Biaka	0	.04	0	0	.01	.01	0	0	.10	.05	.05	0	0	.01	0	.01	0	.09	.51	.12	68	60.5	16	.70
Mbuti	0	0	0	0	0	0	0	0	.09	.22	0	0	0	0	0	0	.03	.06	.43	.18	33	11.3	3	.72
Druze	0	.06	0	0	.01	0	0	.04	0	0	0	0	0	0	0	0	.11	.43	.14	.21	55	43.5	10	.74
Yemenites	0	.06	0	0	0	0	0	.03	0	.02	0	0	0	0	0	0	.02	.47	.20	.19	47	38.3	10	.70
Danes	0	.13	0	0	0	0	0	.04	0	0	0	0	0	0	0	0	.10	.48	.12	.13	42	53.0	10	.71
Finns	0	.15	.02	0	0	0	.02	0	.05	0	0	0	.07	0	0	0	.02	.30	.24	.11	28	32.5	10	.81
Adygei	0	.11	0	0	0	0	.02	.03	.02	0	0	.01	0	0	0	0	.06	.49	.10	.17	50	62.5	10	.71
Europeans	0	.06	0	0	0	.01	.02	0	.06	0	0	0	.03	0	.02	0	.07	.50	.12	.11	41	34.3	13	.71
Han(S)	0	.43	0	.03	0	0	0	0	0	0	0	.02	0	.01	0	0	.07	.04	.35	.07	37	58.7	10	.69
Han(T)	0	.37	0	0	0	0	0	0	0	0.	0	.01	0	0	0	0	0	.07	.52	.03	46	163.9	10	.59
Ami	0	.30	0	.01	0	0	0	0	.01	0	0	.05	0	0	0	0	.01	.01	.60	0	40	99.1	7	.55
Atayal	0	.51	0	0	0	0	0	0	.04	0	0	.05	0	.01	.01	0	0	0	.19	.19	42	98.7	7	.67
Maya	.04	.52	.01	.10	0	.01	0	0	.01	0	0	.01	0	.01	0	.01	.10	0	.13	.04	49	78.4	16	.69
R. Surui	0	.55	.01	.01	0	0	.02	0	0	0	0	0	0	.07	.02	0	.02	0	.27	.03	47	110.5	10	.63
Karitiana	0	.59	0	0	0	0	0	0	0	0	0	0	0	.03	0	0	0	0	.38	0	53	185.4	4	.55
B. Standard Error																								
Biaka	...	.02	...	...	.01	.01	...	...	.03	.02	.02	...	...	.01	...	.01	...	.02	.05	.03				
Mbuti	...	...	...	...	...	...	...	...	.05	.06	...	...	...	...	...	...	.03	.03	.07	.05				
Druze	...	.02	.01	...	.01	...	...	.02	...	...	...	...	...	...	...	...	.03	.05	.03	.05				
Yemenites	...	.03	...	...	.01	...	.01	.02	...	.02	...	...	...	...	...	...	.02	.06	.04	.04				
Danes	...	.03	...	...	...	...	...	.02	...	...	...	...	...	...	...	...	.04	.05	.03	.04				
Finns	...	.05	.02	...	...	...	.02	.03	.05	...	...	...	.08	.05	...	...	.02	.07	.09	.04				
Adygei	...	.03	...	...	...	...	.01	.02	.02	...	...	.01	...	...	...	...	.02	.05	.03	.04				
Europeans	...	.03	...	...	...	.01	.02	...	.03	...	...	...	.02	...	.02	...	.03	.05	.04	.03				
Han(S)	...	.05	...	.03	...	...	...	...	...	.01	...	.02	...	.02	...	...	.04	.03	.05	.03				
Han(T)	...	.05	...	...	...	...	...	...	...	...	...	.01	...	...	...	...	...	.03	.05	.02				
Ami	...	.13	...	.01	...	...	...	.03	.07	...	...	.13	...	...	...	...	.03	.03	.08	...				
Atayal	...	.05	...	...	...	...	...	...	.04	...	...	.04	...	...	.01	...	...	...	.03	.04				
Maya	.03	.06	.01	.04	...	.01	...	...	.01	...	...	.02	...	.02	...	...	.04	...	.04	.02				
R. Surui	...	.05	.01	.01	...	...	.02	...	.01	...	...	...	...	.03	.02	...	.02	...	.05	.02				
Karitiana	...	.05	...	...	...	...	...	...	...	...	...	...	...	.02	...	...	...	...	.05	...				

NOTE.—Four of the 24 possible haplotypes were not seen in any of these populations and have not been included in the table: B2-17-A1, B1-17-A2, B1-12-A2, and B2-12-A2. Data for the three Taiwanese population samples (Han(T), Ami, and Atayal) were reported elsewhere in a different context (Lu et al. 1995).

[a] N = no. of individuals with typings for at least two of the three systems and used to estimate haplotype frequencies.

[b] The likelihood ratio testing for disequilibrium. All values are significant at $P < .001$ (except $P < .01$ for Mbuti).

Table 5

Linkage Disequilibrium Values at DRD2 for 15 Populations

			STRP							
			$(CA)_{16}$		$(CA)_{15}$		$(CA)_{14}$		$(CA)_{13}$	
2N	POPULATION	*Taq*I "B,A"	*Taq*I "B"	*Taq*I "A"	*Taq*I "B"	*Taq*I "A"	*Taq*I "B"	*Taq*I "A"	*Taq*I "B"	*Taq*I "A"
136	Biaka	.83	1.00	1.00	1.00 (ns)	1.00 (ns)	.81 (++)	.19 (ns)	.24 (ns)	.15 (ns)
66	Mbuti	...	...	1.00 (ns)	...	1.00 (ns)	...	.45 (ns)	...	.53 (++)
110	Druze	1.00	.84	.48 (++)	1.00 (ns)	.31 (ns)	1.00 (ns)	1.00 (ns)	.42 (ns)	.49 (ns)
94	Yemenites	1.00	1.00	.71	1.00 (ns)	.46 (ns)	1.00 (ns)	1.00 (ns)	1.00 (ns)	.16 (ns)
84	Danes	1.00	1.00	.70	1.00 (++)	.55 (ns)	1.00 (ns)	1.00 (ns)	1.00 (ns)	1.00 (ns)
56	Finns	.60	.70	.84	.02 (ns)	.77 (ns)	1.00 (ns)	.26 (ns)	1.00 (ns)	1.00 (ns)
100	Adygei	.87	1.00	.68	1.00	.72 (+)	1.00 (ns)	.05 (ns)	1.00 (ns)	1.00 (ns)
76	Europeans	.50	.49	.53	.55 (ns)	1.00 (++)	1.00 (ns)	.27 (ns)	.06 (ns)	1.00 (ns)
74	Hans(S)	1.00	.86	.88	1.00 (ns)	1.00 (ns)	.81	.85	1.00 (ns)	1.00 (ns)
92	Han(T)	1.00	1.00	1.00	1.00 (ns)	1.00 (ns)	1.00	1.00	1.00 (ns)	1.00 (ns)
80	Ami	.94	.95	.87	1.00 (ns)	1.00 (ns)	.94	.87	...	...
84	Atayal	.83	1.00	.84	...	...	.91	.70	.90	1.00
98	Maya	.95	.50	.46	1.00 (ns)	1.00 (ns)	.38 (++)	.37 (++)	1.00 (+)	1.00 (+)
92	Surui	.89	.78	.91	1.00 (ns)	1.00 (ns)	.66	.95	.40 (ns)	1.00 (ns)
106	Karitiana	1.00	1.00	1.00	...	...	1.00	1.00	...	...

NOTE.—Each column gives one of the nine possible pairwise disequilibrium values (D′) for each population. The 17- and 12-repeat alleles at the STRP are rare and hence omitted. All values are significantly different from zero at $P < .001$ *unless* otherwise noted. Values that are not significantly different from zero are labeled "ns"; values with $.05 < P < .01$ are labeled "+" and those with $.01 < P < .001$ are marked "++". When a group lacks a particular allele, ellipses (. . .) appear in place of D′.

was not polymorphic in the Mbuti) and show considerable frequency variation among populations, a result in accord with other studies of DNA polymorphisms on diverse populations (Bowcock et al. 1991*b*; Kidd et al. 1991, 1993). At the *Taq*I "B" site there is significant variation even among European and Middle Eastern populations: the Yemenites and Druze from the Middle East have a lower frequency of the B1 (*Taq*I site missing) allele than the Danes and Finns from northern Europe—.07 and .08 versus .12 and .25—with an overall heterogeneity χ^2 among the six European and Middle Eastern populations equal to 13.5 (df = 4; $P < .01$).

The STRP system is apparently limited to only six different alleles. The four common alleles—13, 14, 15, and 16 repeats—are those originally described by Hauge et al. (1991) as C1, C2, C3, and C4, respectively. The 15-repeat allele appears to be common only in European and Middle East populations; the 13-repeat common in those and African populations; the 16-repeat common in all but Africans; and the 14-repeat common everywhere. The frequency variation is significant overall, though not necessarily for any pair of populations. Three groups of populations seem to be defined by the patterns of allele frequencies: (1) the two African populations, (2) the European and Middle Eastern samples, and (3) the Chinese, Taiwanese, and New World populations.

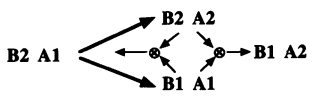

Figure 2 Schematic representation of the evolutionary relationships of DRD2 haplotypes defined by the *Taq*I "B" and "A" sites. The sequence in other higher primates corresponds to B2A1 making this the ancestral haplotype. Separate mutations, indicated by the thicker arrows, gave rise to the B2A2 and B1A1 haplotypes. These presumably occurred only once, and drift has increased their frequencies. Recombination—crossovers indicated by thinner arrows and ⊗—can generate the B1A2 haplotype and reconstitute the ancestral B2A1 haplotype. This will be an ongoing process with its rate dependent on the frequency of B2A2/B1A1 heterozygotes in the population.

The frequency variation for the *Taq*I "A" system is much as was shown by Barr and Kidd (1993): the A1 allele has moderate frequencies in African populations, has the lowest frequencies seen in Middle Eastern populations, has moderate frequencies in European populations, has frequencies approaching 50% in East Asian populations, and becomes more common than the A2 allele in Amerindian populations. We note that the particularly low frequency of the A1 allele in Middle Eastern populations originally reported by Barr and Kidd (1993) is confirmed by the larger Yemenite and Druze samples studied here. Our sample of Finns is independent of the one studied by Goldman et al. (1993), but the allele frequencies are quite similar. Similarly, the three Amerindian populations we have studied have a

range of allele frequencies similar to those reported by Goldman et al. (1993). The allele frequencies in Taiwanese populations (Lu et al. 1995) are similar to the previously reported Asian frequencies (Barr and Kidd 1993). Thus, whenever specific populations or regions are studied, similar frequency estimates have resulted, confirming the overall global pattern of frequency variation for the "A" system. The variation among the six European and Middle Eastern populations we have studied is almost as much as was seen at the "B" system, but the heterogeneity test does not reach statistical significance ($\chi^2 = 6.0$; df = 4).

Variation among populations is measured as F_{st}, the standardized inbreeding coefficient. Since F_{st} is the ratio of between-group variance over the total possible variation in the population, this measure of population structure can vary from zero to one, where one indicates complete fixation within groups of the allelotype measured. For the twelve populations in table 3 plus the three Taiwanese populations from Lu et al. (1995) (see table 4) the *Taq*I "B" polymorphism shows more variation among groups—$F_{st} = .27$—than the *Taq*I "A" polymorphism, $F_{st} = .15$. For these same 15 populations the STRP has an F_{st} of .18. Thus, all three sites show substantial amounts of variation among populations. Within populations, in contrast, the STRP shows higher average heterozygosity, as expected, with $\bar{H} = .59$, while the average heterozygosities for the *Taq*I "B" and "A" systems are $\bar{H} = .31$ and $\bar{H} = .38$, respectively. These average values for F_{st} and $\bar{H}$ apply only to these particular 15 populations and are not directly comparable to other sets of populations. However, we note that an overlapping set of 12 populations with the same broad geographic coverage—Africa to South America—had an average F_{st} for 35 RFLP systems of ~.15 and an average heterozygosity for the same 35 systems of just over .3 (Kidd et al. 1993). Thus, there is nothing unusual about the inter- or intrapopulation variation at the individual polymorphic systems at DRD2.

Haplotype Frequency Variation

For the haplotypes (table 4) the average heterozygosity is high; $\bar{H} = .67$. Eight different haplotypes occur with a frequency of 10% in at least one population; all eight show significant frequency variation among the 15 populations, but the standardized variance averaged across haplotypes is about the same as for the individual sites: $F_{st} = .16$. As yet there are no haplotype data at other loci analyzed for a fully comparable set of populations, so it is not possible to comment on how typical these F_{st} and $\bar{H}$ values might be. Overall linkage disequilibrium at the locus, measured as a likelihood ratio χ^2, is highly significant for all populations. Most of the pairwise disequilibrium coefficients (table 5) are also highly significant. Strong positive associations of the A1 and B1 alleles and the A2 and B2 alleles were seen elsewhere in samples of mixed European ancestry (Hauge et al. 1991; O'Hara et al. 1993; Suarez et al. 1994) and a sample of U.S. blacks who are primarily of mixed African ancestry (O'Hara et al. 1993). That same pattern of highly significant disequilibrium that is at or near the maximum possible is now extended to specific European and Middle Eastern populations and to most other populations studied (table 5).

Pairwise disequilibrium measures of individual STRP alleles with the B and A sites also show levels of disequilibrium at or near the maximum possible. However, for most STRP alleles in most populations the disequilibrium is not statistically significant. The most notable exception is the 16-repeat allele, which is in significant disequilibrium with both the B and A sites in all but the Mbuti. Another exception is the 14-repeat allele, which is in significant disequilibrium with both the B and A sites in the East Asian and New World populations.

The pairwise D' values do not fully capture the pattern of disequilibrium. Inspection of the haplotypes in the various populations (table 4) shows that the 16-repeat STRP allele occurs primarily on the B1-A1 haplotype (defined by the two *Taq*I polymorphisms) while the 15-, 14-, and 13-repeat alleles occur primarily on the B2-A2 haplotype. This strong association between STRP allele and haplotype background is easier to see in the condensed table of regional haplotype frequencies (table 6). This suggests that the mutation rate of the STRP is quite low and that haplotype frequencies differ primarily as a result of random genetic drift. The strong linkage disequilibrium also indicates that the recombination rate is low, although crossovers may be responsible for haplotypes such as B2-14-A1 in the Maya and low-frequency haplotypes in other populations. The two most frequent haplotypes outside of Africa, B1-16-A1 and B2-14-A2, show the largest and most significant amounts of deviation from the expected frequencies. The STRP is located much closer to the *Taq*I "B" system (~6 kb) than to the *Taq*I "A" system (~20 kb). However, the two physically closer systems show essentially the same amount and pattern of disequilibrium as do the STRP and *Taq*I "A" systems. This appears to argue that in this region, frequency variation of entire haplotypes, at least in non-African populations, is the primary source of different levels of disequilibrium, not any systematic role of recombination. This is consistent with the findings of Jorde et al. (1994) and Watkins et al. (1994) that in such short regions recombination is not generally a systematic factor in disequilibrium. Similarly, the relatively small number of STRP alleles, the low frequency and sparse distribution of the "extreme" alleles, and the significant disequilibrium of some alleles with specific haplotypes suggest that the mutation rate at this STRP is not exceedingly high and that the different distribu-

Table 6

Frequencies of the Most Common Haplotypes by Geographic Region

	DRD2 Haplotype Frequencies												
Region (no. of populations pooled)	B1 16 A1	B1 14 A1	B1 ... A1	B2 14 A1	B2 13 A1	B2 ... A1	B1 ... A2	B2 16 A2	B2 15 A2	B2 14 A2	B2 13 A2	B2 ... A2	*N*
Africa (2)	.027	0	.014	.097	.106	.034	.007	.010	.080	.481	.140	.007	101
Mideast (2)	.061	0	.005	0	.009	.035	0	.069	.450	.168	.203	0	102
Europe (4)	.109	0	.006	.030	0	.035	.028	.066	.457	.135	.134	0	161
East Asia (4)	.401	.009	0	.013	0	0	.040	.018	.031	.416	.072	0	165
Americas (3)	.554	.036	.022	.004	0	.006	.045	.039	0	.264	.023	.004	149

tions of STRP alleles in non-African populations are primarily the result of genetic drift of entire haplotypes and not of mutation at the STRP.

The Evolutionary Pattern

The haplotype data fit into an evolutionary framework diagrammed in figure 2. The B2 and A1 alleles were homozygous in all other higher primates tested, as we would expect, on the basis of several considerations. First, only small, though significant, levels of fixed nucleotide changes between humans, chimpanzees, and gorillas exist at the many loci studied. Second, the expected lifetime of a polymorphism is far shorter than the 4 million years or more since the human lineage diverged from the other great apes (Takahata 1993). Third, the survey by Mountain et al. (1992), though not based on sequence, concluded that for most to all human RFLP polymorphisms one of the two alleles was present as the only form in chimpanzees and/or gorillas. These same data argue that recurrent mutation at these sites in humans can be considered to be absent. Thus, starting from the ancestral B2A1 haplotype, separate mutations generated the B1 allele and the A2 allele. The frequency and disequilibrium data argue for the "parallel" scenario in figure 2 rather than a sequential scenario such as B2A1→B2A2→B1A2→B1A1, with the first two steps being sequential mutations and the third recombination. However, in light of the stochastic nature of drift, even seemingly "impossible" scenarios cannot be absolutely ruled out and we cannot yet evaluate the relative likelihoods of the alternatives.

The allele sizes for the STRP in chimpanzees and gorillas indicate that a general increase in repeat number has occurred in hominid evolution, but it is not possible to identify any specific extant allele as ancestral to the others. The occurrence of multiple STRP alleles within the B2A2 framework argues that the B2A2 haplotype (the A1→A2 mutation) arose early in recent human evolution and has had time to "acquire" these STRP alleles by mutation or recombination with the ancestral haplotype. The ancestral "B2A1" haplotype for the two flanking RFLPs is uncommon to absent in all populations studied (except in Africa) as is the "double-derived" haplotype, "B1A2." Drift, probably associated with the exodus of modern humans from Africa, would appear to be the explanation for the low frequency of the ancestral B2A1 haplotype in all non-African populations. The very low frequency of the "double-derived" haplotype, B1A2, in all parts of the world could also be explained by drift. However, its somewhat higher frequency in East Asia and the Americas would seem to favor the scenario in figure 2. Only in these populations do the frequencies of both the B2A2 and B1A1 haplotypes become common, resulting in heterozygotes being more common. Even a low frequency of recombination might generate enough recombinants that, in conjunction with drift, these low but detectable frequencies of the recombinant occur. Different STRP alleles occur on this B1A2 haplotype in different populations, but always one of the two most common in the specific population. This also suggests that recent, separate crossover events are primarily responsible for the B1A2 haplotypes. Thus, the two globally most common haplotypes appear to have separate origins prior to the spread of modern humans around the world. The STRP alleles further subdivide the "B2A2" haplotype lineages in populations from Africa, the Middle East, and Europe but not appreciably in East Asia and the Americas.

The haplotypes give a new perspective on the population genetics of the A1 allele, the allele so prominently featured in disease-association studies. This allele is largely confined to the B1-16-A1 haplotype. Except in the two African populations sampled, the B1-16-A1 haplotype represents over half, usually 70%, of all A1 chromosomes. This suggests a common evolutionary origin of this haplotype quite early prior to human dispersal around the world and its subsequent "preservation" among those modern humans that left Africa. However, this is not the only A1-containing haplotype in the U.S. "white" population. In the Druze, for exam-

ple, the B2-15-A1 haplotype is nearly as frequent as the B1-16-A1. It is extremely improbable that a common functional variant would be equally associated with all A1-containing haplotypes. Thus, it should be useful, especially in Europeans and Africans, to include the STRP in haplotypes for association studies to resolve the controversy over a functional role of the DRD2 locus in alcoholism and other neuropsychiatric disorders.

These data provide an initial molecular and statistical framework for detecting common functionally relevant allelic variation at DRD2 and interpreting associations of that variation with behavioral disorders. They also provide additional support for the African origin of modern *Homo sapiens*. A clearer picture of the global distribution of DRD2 variation will require data on many more populations. We are in the process of collecting such data.

Acknowledgments

We thank R. A. Shimkets for his assistance in subcloning the *Eco*RI-*Bam*HI fragment and M. S. Lee for his help in resolving the restriction map of this region. We also appreciate the assistance of Audrey Shuster and Beth Wilson, who established many of the population cell lines. This work was supported in part by USPHS grants AA09379 from the NIAAA plus MH30929 and MH39029 from the NIMH as well as a grant from the Alfred P. Sloan foundation to K.K.K. and by grant MH48991 from the NIMH to D.K.G. Collection of the Danish samples was supported by grant MH44876 from the NIMH to Dr. P. Holzman.

References

Anderson MA, Gusella J (1984) Use of cyclosporin A in establishing Epstein-Barr virus-transformed human lymphoblastoid cell lines. In Vitro 20:856–858

Araki K, Kuwano R, Morii K, Hayashi S, Minoshima S, Shimizu N, Katagiri T, et al (1992) Structure and expression of human and rat D2 dopamine receptor genes. Neurochem Int 21:91–98

Balazs I, Neuweiler J, Gunn P, Kidd J, Kidd K, Kuhl J, Mingjun L (1992) Human population genetic studies using hypervariable loci. I. Analysis of Assamese, Australian, Cambodian, Caucasian, Chinese and Melanesian populations. Genetics 131:191–198

Barr C, Kidd KK (1993) Population frequencies of the A1 allele at the dopamine D_2 receptor locus. Biol Psychiatry 34:204–209

Bolos AM, Dean M, Lucas-Derse S, Ramsburg M, Brown GL, Goldman D (1990) Population and pedigree studies reveal a lack of association between the dopamine D_2 receptor gene and alcoholism. JAMA 264:3156–3160

Bowcock AM, Bucci C, Hebert JM, Kidd JR, Kidd KK, Friedlaender JS, Cavalli-Sforza LL (1987) Study of 47 DNA markers in five populations from four continents. Gene Geogr 1:47–64

Bowcock AM, Hebert JM, Mountain JL, Kidd JR, Rogers J, Kidd KK, Cavalli-Sforza LL (1991*a*) Study of an additional 58 DNA markers in five human populations from four continents. Gene Geogr 5:151–173

Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, Kidd KK, Cavalli-Sforza LL (1991*b*) Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. Proc Natl Acad Sci USA 88:839–843

Cavalli-Sforza LL, Kidd JR, Kidd KK, Bucci C, Bowcock AM, Hewlett BS, Friedlaender JS (1986) DNA markers and genetic variation in the human species. Cold Spring Harbor Symp Quant Biol 11:411–417

Dal Toso R, Sommer B, Ewert M, Herb A, Pritchett DB, Bach A, Shivers BD, et al (1989) The dopamine D2 receptor: two molecular forms generated by alternative splicing. EMBO J 8:4025–4034

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Proc R Stat Soc 39:1–38

Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine scale mapping. Genomics (in press)

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6

Gandelman K, Harman S, Todd RD, O'Malley KL (1991) Analysis of the structure and expression of the human dopamine $D2_A$ receptor gene. J Neurochem 56:1024–1029

Gejman P, Ram A, Gelernter J, Friedman E, Cao Q, Pickar D, Blum K, et al (1994) No structural mutation in the dopamine D_2 receptor gene in alcoholism or schizophrenia. JAMA 271:204–208

Gelernter J, Grandy DK, Bunzow J, Pakstis AJ, Civelli O, Retief AE, Litt M, et al (1989) D2 dopamine receptor locus (probe hD2G1) maps close to D11S29 (probe L7) and is also linked to PBGD (probe PBGD0.0) and D11S84 (probe p-2-7-ID6) on 11q. (A2669) Cytogenet Cell Genet 51:1002

Gelernter J, Pakstis AJ, Pauls DL, Kurlan R, Gancher ST, Civelli O, Grandy D, et al (1990) Tourette syndrome is not linked to D_2 dopamine receptor. Arch Gen Psychiatry 47:1073–1077

Gelernter J, Pauls DL, Leckman J, Kidd KK, Kurlan R (1994) D2 dopamine receptor alleles do not influence severity of Tourette's syndrome. Arch Neurol 51:397–400

Giuffra LA, Lichter P, Wu J, Kennedy JL, Pakstis AJ, Rogers J, Kidd JR, et al (1990) Genetic and physical mapping and population studies of a fibronectin receptor β-subunit-like sequence on human chromosome 19. Genomics 8:340–346

Goldman D, Brown GL, Albaugh B, Robin R, Goodson S, Trunzo M, Akhtar L, et al (1993) DRD2 dopamine receptor genotype, linkage disequilibrium, and alcoholism in American Indians and other populations. Alcohol Clin Exp Res 17:199–204

Grandy DK, Litt M, Allen L, Bunzow JR, Marchionni M, Makam H, Reed L, et al (1989*a*) The human dopamine D_2 receptor gene is located on chromosome 11 at q22-q23 and identifies a *Taq*I RFLP. Am J Hum Genet 45:778–785

Grandy DK, Marchionni MA, Makam H, Stofko RE, Alfano M, Frothingham L, Fischer JB, et al (1989*b*) Cloning of the cDNA and gene for a human D_2 dopamine receptor. Proc Natl Acad Sci USA 86:9762–9766

Grandy DK, Zhang Y, Civelli O (1993) PCR detection of the *Taq*I A RFLP at the DRD2 locus. Hum Mol Genet 2:2197

Hauge X, Grandy D, Eubanks J, Evans G, Civelli O, Litt M (1991) Detection and characterization of additional DNA polymorphisms in the dopamine D_2 receptor gene. Genomics 10:527–530

Hawley ME, Kidd KK (1995) HAPLO: a program using the EM algorithm to estimate frequencies of multi-site haplotypes. J Hered 86:409–411

Hirschberg HZ (1969) The Druzes. In: Arberry AJ (ed) Religion in the Middle East: three religions in concord and conflict, Vol 2, Islam, pp 330–348, University Press, Cambridge

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima J (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 10:23–47

Itokawa M, Arinami T, Futamura N, Hamaguchi N, Toru M (1993) A structural polymorphism of human dopamine D2 Receptor, D2(Ser^{311}→Cys). Biochem Biophys Res Commun 196:1369–1375

Jorde LB, Watkins WS, Carlson M, Groden J, Albertsen H, Thliveris A, Leppert M (1994) Linkage disequilibrium predicts physical distance in the adenomatous polyposis coli region. Am J Hum Genet 54:884–898

Kidd JR, Black FL, Weiss KM, Balazs I, Kidd KK (1991) Studies of three Amerindian populations using nuclear DNA polymorphisms. Hum Biol 63:775–794

Kidd JR, Pakstis AJ, Kidd KK (1993) Global levels of DNA variation. Proceedings of the 4th International Symposium on Human Identification. Promega, Madison, pp 21–30

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49: 49–67

Lichter JG, Barr CL, Kennedy JL, Van Tol HHM, Kidd KK, Livak KJ (1993) A hypervariable segment in the human dopamine receptor D4 (DRD4) gene. Hum Mol Genet 2:767–773

Litt M, Kramer P, Kort E, Fain P, Cox S, Root D, White R, et al (1995) The CEPH consortium linkage map of human chromosome 11. Genomics 27:101–112

Lu RB, Ko HC, Chang FM, Castiglione CM, Schoolfield G, Pakstis AJ, Kidd JR, et al (1995) No association between alcoholism and multiple polymorphisms at the dopamine D2 receptor gene (DRD2) in three distinct Taiwanese populations. Biol Psychiatry 38

Mountain JL, Lin AA, Bowcock AM, Cavalli-Sforza LL (1992) Evolution of modern humans: evidence from nuclear DNA polymorphisms. Phil Trans R Soc London B 337:159–165

Nevo S, Cleve H, Koller A, Eigel E, Patutschnick W, Kanaaneh H, Joel A (1992) Serum protein polymorphisms in Arab Moslems and Druze of Israel: BF, F13B, AHSG, GC, PLG, PI, and TF. Hum Biol 64:587–603

O'Hara BF, Smith SS, Bird G, Persico AM, Suarez BK, Cutting GR, Uhl GR (1993) Dopamine D2 receptor RFLPs, haplotypes and their association with substance use in Black and Caucasian research volunteers. Hum Hered 43:209-218

Parsian A, Fisher L, O'Malley K, Todd R (1991) A new *Taq*I RFLP within intron 2 of human dopamine D2 receptor gene. Nucleic Acids Res 19:6977

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2d ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sarkar G, Kapelner S, Grandy DK, Marchionni M, Civelli O, Sobell J, Heston L, et al (1991) Direct sequencing of the dopamine D2 receptor (DRD2) in schizophrenics reveals three polymorphisms but no structural change in the receptor. Genomics 11:8–14

Seeman P, Ohara K, Ulpian C, Seeman MV, Jellinger K, Van Tol HH, Niznik HB (1993) Schizophrenia: normal sequence in the dopamine D2 receptor region that couples to G-proteins. Neuropsychopharmacology 8:137–142

Suarez BK, Parsian A, Hampe CL, Todd RD, Reich T, Cloninger CR (1994) Linkage disequilibria at the D_2 dopamine receptor locus (DRD2) in alcoholics and controls. Genomics 19:12–20

Takahata N (1993) Allelic genealogy and human evolution. Mol Biol Evol 10:2–22

Templeton AR, Sing C (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. Genetics 134:659–669

Torroni A, Schurr TG, Yang G, Szathmary EJE, Williams RC, Schanfield MS, Troup GA, et al (1992) Native American mitochondrial DNA analysis indicates that the Amerind and NaDene populations were founded by two independent migrations. Genetics 130:153–162

Vigilant L, Pennington R, Harpending H, Kocher T, Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354

Watkins WS, Zenger R, O'Brien E, Nyman D, Eriksson AW, Renlund M, Jorde LB (1994) Linkage disequilibrium patterns vary with chromosomal location: a case study from the von Willebrand factor region. Am J Hum Genet 55:348–355