Am J Hum Genet 32:555–563, 1980

Linkage Relationships and Multipoint Mapping of the Human Parotid Salivary Proteins (Pr, Pa, Db)

PAO-LO YU,[1] ROBERT C. KARN,[1] A. DONALD MERRITT,[2] EDWIN A. AZEN,[3] AND P. MICHAEL CONNEALLY[1]

SUMMARY

Based on data from 76 informative families, linkages between *Pa* and *Pr* and between *Pr* and *Db* have been established by two-point linkage analysis. In both pairs of loci, there were no significant sex heterogeneity in recombination fractions. Linkage between *Pa* and *Db* cannot be established based on two-point analyses, but a significant sex difference in the recombination fraction between *Pa* and *Db* was observed. Strong confirming evidence was obtained from three-point analysis to place *Pa*, *Pr*, and *Db* in one linkage group. The most likely order is *Pa-Pr-Db*, but the relative odds over second order *Pr-Pa-Db* are small. Haplotype frequencies of *Pr*, *Pa*, and *Db* were obtained based on the phenotypes of the 685 random Caucasians, providing evidence for marked linkage disequilibrium among the three loci.

INTRODUCTION

Human parotid saliva contains many proteins, including a variety which are rich in proline, glycine, and glutamic acid. The genetics of three proline-rich proteins, Pr (proline-rich), Pa (acidic), and Db (double-band), have been studied by several groups [1–6]. Those proteins are polymorphic in Caucasian, black, and Oriental populations, and each one follows a simple Mendelian mode of inheritance. Based on family

Received September 18, 1979; revised December 3, 1979.

This study is publication no. 79-9 of the Department of Medical Genetics, Indiana University School of Medicine, and was supported in part by grant P50-GM-21054 from the U.S. Public Health Service in connection with the Indiana University Human Genetics Center. R. C. K. was supported by Career Development Award 1-K04-AM-00284 from the U.S. Public Health Service.

[1] Department of Medical Genetics, Indiana University School of Medicine, Indianapolis.

[2] Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda, Maryland.

[3] Department of Medicine, University of Wisconsin, Madison.

studies, we have shown that Pr, Pa, and Db are the expressions of three separate but linked loci [7]. Immunologic [8] and biochemical [9] evidence indicates that at least two (*Pr* and *Pa*) of the three loci have very similar structures.

In this report, we substantiate the evidence for the linkage of the three loci, *Pr*, *Pa*, and *Db*, with additional family data and attempt to determine their mapping order.

MATERIALS AND METHODS

Parotid saliva samples were collected on ice from 3,000 individuals in 400 families using a modified Carlsen-Crittenden collector. The Pr, Pa, and Db phenotypes were scored following electrophoresis in a vertical, 5% thin-layer acrylamide gel, a modification [10] of the technique of Azen and Oppenheim [1]. The bridge buffer consisted of 0.15 M Tris and 0.01 M boric acid and had a final pH of 8.6. The gel was run for approximately 2 hrs at a constant voltage (180 V), at which time the bromophenol blue marker dye reached the anodal end of the gel. The gel was fixed in 20% trichloroacetic acid for 1 hr (minimum), rinsed with water, and stained in 0.05% Coomassie Blue R250 in 10% methanol for 4–16 hrs. The gel was then rinsed in water (no destaining was necessary) and photographed for permanent record.

The Pa phenotypes were also scored following electrophoresis in a horizontal starch-urea gel, pH 2.4, using a continuous aluminum-lactate buffer as described by Sung and Smithies [11] and Azen [12]. The cathodal gel was run for 18 hrs at a constant 300 V, sliced lengthwise, and each half placed in 2 mM Amido Black in 1% acetic acid for 40 min. The stained gel was rinsed well, placed in 0.5 M sulfuric acid to destain for 2–4 hrs, and then photographed.

Linkage analysis between *Pr*, *Pa*, and *Db* and their locus order was done on 76 informative families, using the computer programs LIPED [13] and ZEXMAX [14]. The gene and haplotype frequencies were estimated on 685 random Caucasians using Kaplan and Elston's MAXLIK [15].

RESULTS

Figure 1 shows the segregation for *Pr*, *Pa*, and *Db* in a mating between two individuals heterozygous for the three loci. The Pr proteins are the expression of two autosomal codominant alleles Pr^1 and Pr^2 [7]. The Db proteins are an autosomal dominant system with dominant allele Db^1 and recessive allele Db^{0*}. Similarly, Pa proteins are an autosomal dominant system whose dominant and recessive alleles are Pa^1 and Pa^0, respectively.

The results of two-point linkage analysis are presented in table 1. A significant linkage relationship ($z > 3$) was obtained between *Pa* and *Pr* and between *Pr* and *Db*. *Pa* and *Pr* are closely linked with a recombination fraction estimated at .03 (male) to .04 (female). The linkage between *Pr* and *Db* is slightly looser with a recombination fraction estimated at .08 to .14. In both cases, the observed sex difference in the recombination fractions is not significant ($\chi^2_1 = 0.19$ and 0.47, $P > .4$). For *Pa* and *Db*, none of the lod scores reached 3, and there is a significant sex difference in the recombination fraction ($\chi^2_1 = 4.18$, $P < .05$). Figure 2 shows the lod score curves by sex for each pair of loci.

Since linkage has been established between *Pa* and *Pr* and between *Pr* and *Db*, an attempt was made to determine the relative order of these three loci. A computer program ZEXMAX [14], which incorporated Kaplan and Elston's MAXLIK [15], was used

* Db^0 is used here to replace Db^2 reported in an earlier article [7], since Db^0 is more consistent for a recessive allele.

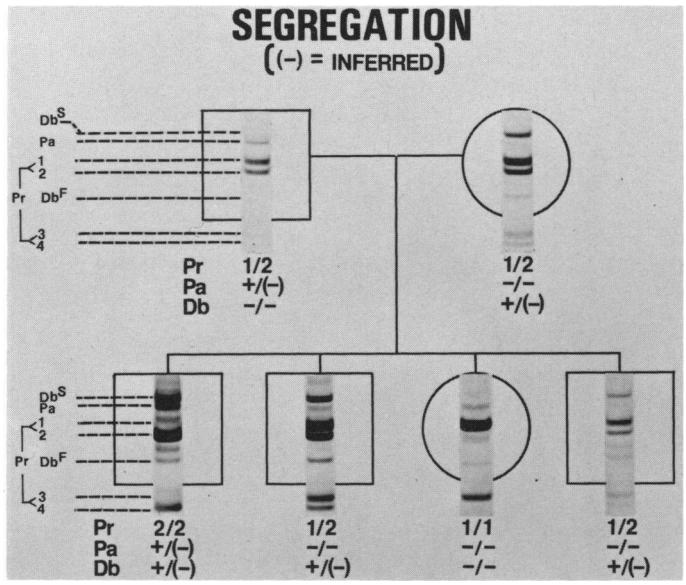

FIG. 1.—Segregation for *Pr*, *Pa*, and *Db* in a mating between two individuals heterozygous for the three loci.

to estimate the recombination fractions between the loci and to calculate the maximum lod score (logarithm of the likelihood ratio of multiple linkage, three-point, vs. nonlinkage of the loci) for each order. These lod scores were then compared to give the relative odds for each order. The result of this multipoint analysis is given in table 2. Both orders *Pa-Pr-Db* and *Pr-Pa-Db* are more likely than order *Pr-Db-Pa*, with the relative odds ranging from 3112:1 to 299×10^{12}:1. Order *Pa-Pr-Db* is slightly favored over the order *Pr-Pa-Db*. The relative odds range from 1.004:1 (male) to 3.02:1 (female). When the informative parent was male, the estimates of the recombination frequency were smaller and lod scores were larger compared with data derived from informative females. Figure 3 shows the relative distance and order of *Pr*, *Pa*, and *Db* in male and female maps based on two- and three-point analyses.

Gene frequencies for *Pr*, *Pa*, and *Db* alleles were estimated based on 685 random Caucasians and are shown in table 3. The haplotype frequencies of these three loci were also estimated using MAXLIK [15]. Table 4 shows both estimated haplotype frequencies and expected frequencies assuming independence among these three loci. A chi-square test for goodness of fit revealed a significant discrepancy between the observed and expected values ($\chi^2_4 = 999.94$, $P < .005$).

DISCUSSION

The salivary protein expressions, Pr, Pa, and Db, constitute a complex human polymorphism. Our estimates of gene frequencies among Caucasians (table 3) are in agreement with those reported by Azen and Denniston [2] and by Friedman et al. [3].

TABLE 1

LINKAGE RELATIONSHIPS OF *Pr*, *Pa*, AND *Db*

MARKER PAIR	NO. FAMILIES	TOTAL NO. INDIVIDUALS	LOD(z) SCORES AT VARIOUS RECOMBINATION FRACTIONS θ = 0.00	0.05	0.10	0.20	0.30	0.40	$\hat{z}_m$	$\hat{\theta}_m$
Pr:Pa	59	564								
Male			$-\infty$	16.19	14.58	10.43	6.06	2.17	16.53	0.03
Female			$-\infty$	9.49	8.75	6.37	3.73	1.40	9.53	0.04
Total*			$-\infty$	26.70	24.30	17.43	10.11	3.65	26.95	0.03
Pr:Db	39	454								
Male			$-\infty$	2.36	2.43	1.91	1.13	0.39	2.45	0.08
Female			$-\infty$	1.05	1.65	1.61	1.07	0.41	1.74	0.14
Total*			$-\infty$	2.74	3.74	3.44	2.20	0.80	3.83	0.12
Pa:Db	37	421								
Male			$-\infty$	2.22	2.27	1.77	1.04	0.36	2.30	0.08
Female			$-\infty$	−2.99	−1.46	−0.29	0.06	0.05	0.08	0.34
Total*			$-\infty$	−0.52	1.03	1.62	1.16	0.42	1.63	0.19

* Males plus females plus intercrosses.

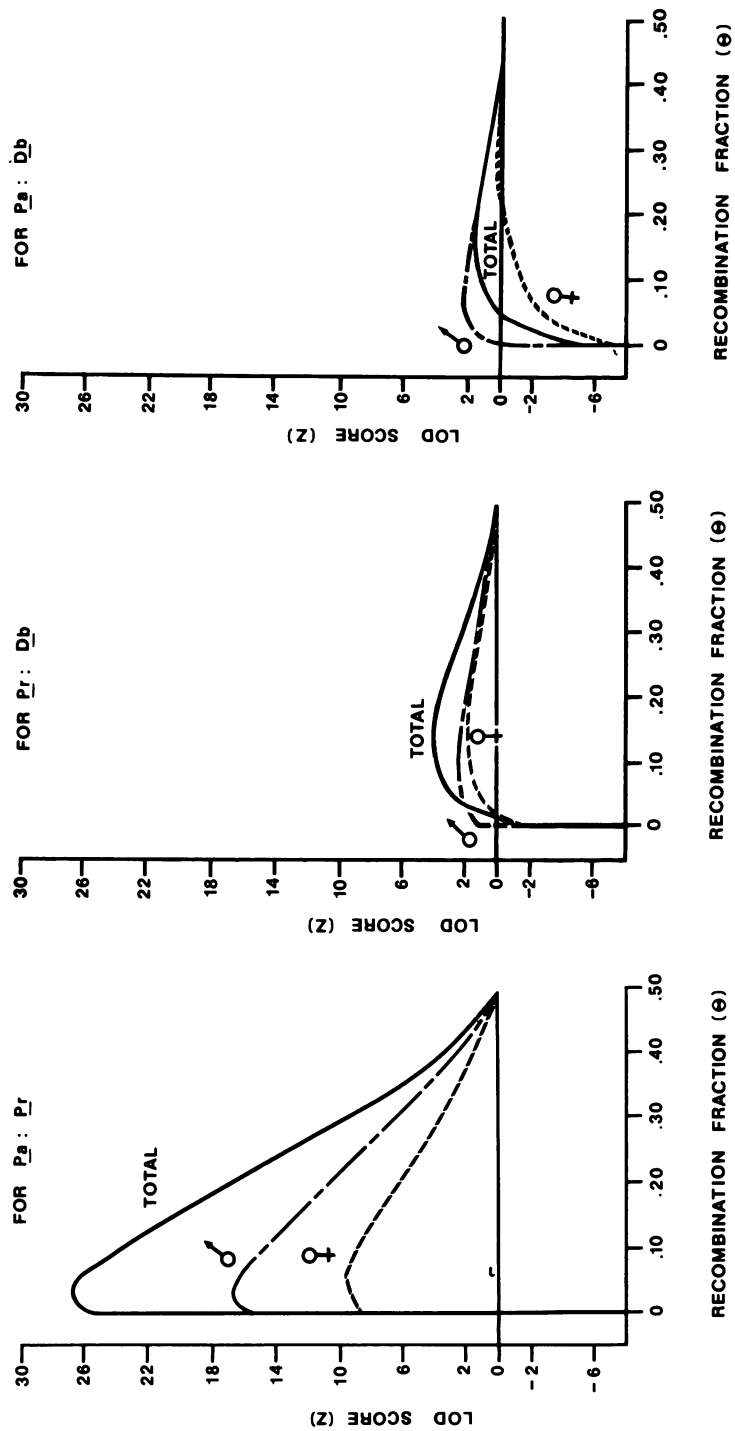

FIG. 2.—Lod score (z) curves at various recombination fractions (θ) by sex for each pair of loci

TABLE 2

THREE-POINT LINKAGE FOR *Pa*, *Pr*, AND *Db*

Sex	$\hat{c}$*	$\hat{\theta}_I$†	$\hat{\theta}_{II}$‡	$\hat{\theta}_{III}$§	$\hat{z}$‖	Odds#
Males:						
Pa-Pr-Db	1.0	.03	.08	.11	21.2	3126
Pr-Pa-Db	1.0	.03	.08	.11	21.2	3112
Pr-Db-Pa	1.0	.08	.08	.15	17.8	1
Females:						
Pa-Pr-Db	1.0	.04	.14	.18	10.8	129×10^5
Pr-Pa-Db	1.0	.04	.34	.35	10.4	427×10^4
Pr-Db-Pa	1.0	.14	.34	.39	3.7	1
Combined:						
Pa-Pr-Db	1.0	.03	.12	.15	32.3	299×10^{12}
Pr-Pa-Db	1.0	.03	.19	.21	32.0	126×10^{12}
Pr-Db-Pa	1.0	.12	.19	.26	17.9	1

* Coefficient of coincidence.
† Recombination fraction between first and second loci for that order.
‡ Recombination fraction between second and third loci for that order.
§ Recombination fraction between first and third loci for that order. $\hat{\theta}_{III} = \hat{\theta}_I + \hat{\theta}_{II} - 2\hat{c}\hat{\theta}_I\hat{\theta}_{II}$.
‖ Log-likelihood.
\# Relative odds within classes.

We previously established close linkage between *Pr* and *Pa* ($\hat{\theta} = .04$) [7]. Our present study provides additional data from 40 more families, increasing the lod scores substantially. Maximum lod scores were obtained at $\hat{\theta} = .03$ for *Pr* and *Pa*. Additionally, linkage between *Pr* and *Db* is established. A maximum lod score of 3.83 at $\hat{\theta} = .12$ was obtained when the data from both informative male and female parents and intercrosses were combined. In both pairs of loci, no significant sex heterogeneity in recombination was observed.

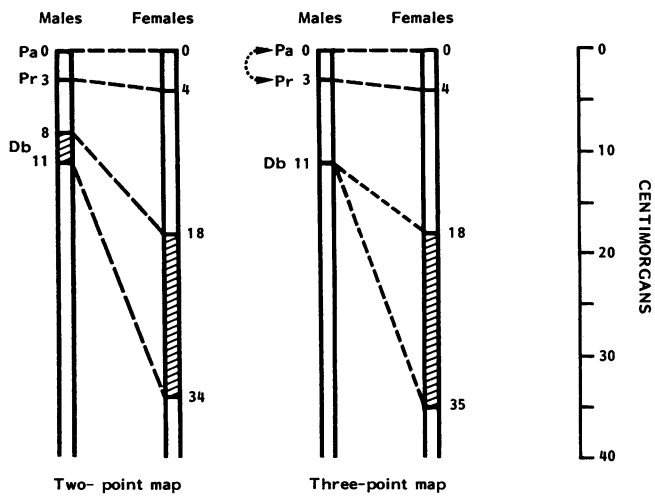

FIG. 3.—Relative distance and order of *Pr*, *Pa*, and *Db* in male and female maps based on two- or three-point analyses of family data.

TABLE 3

ESTIMATION OF GENE FREQUENCIES

Locus	Allele	$p \pm SE$
Pr	Pr^1	.718 ± .012
	Pr^2	.282 ± .012
Db	Db^1	.179 ± .011
	Db^0	.821 ± .011
Pa	Pa^1	.211 ± .012
	Pa^0	.789 ± .012

NOTE. —No. = 685.

The linkage relationship between *Pa* and *Db* cannot yet be established, based on two-point analysis, as none of the lod scores reached 3. However, a significant sex difference in the recombination fraction suggests that linkage between *Pa* and *Db* is relatively closer in males ($\hat{\theta} = .08$ and $\hat{z} = 2.3$) than in females ($\hat{\theta} = .34$ and $\hat{z} = 0.08$).

The three-point analysis shows the orders *Pa-Pr-Db* or *Pr-Pa-Db* as definitely more likely than order *Pr-Db-Pa*. Though *Pa-Pr-Db* is slightly favored over *Pr-Pa-Db* in all three types of data, the relative odds (ranging from 1.004:1 to 3.02:1) are not large enough to establish the definite order at this time. In addition to determining the order of the loci, the three-point analysis also provides a measure of linkage. To give further support to the two-point results (table 1), three-point lod scores for all three of the loci being linked vs. only the *Pa* and *Pr* loci being linked were calculated to be 4.7 for males, 1.3 for females, and 5.4 for combined. The lod scores for male and for combined data provide strong evidence for placing *Pa*, *Pr*, and *Db* in one linkage group.

Having established that these salivary proteins are expressions of three closely linked loci, they may be treated jointly as haplotypes similar to the treatment of the Rhesus (*Rh*) and histocompatibility (*HLA*) loci. As shown in table 4, there is a significant discrepancy between the estimated haplotype frequencies and expected frequencies. The haplotypes Pr^1-Pa^0-Db^1, Pr^1-Pa^0-Db^0, and Pr^2-Pa^1-Db^1 are more frequent than expected, while other haplotypes show less frequent occurrence than expected. This discrepancy between the two frequency distributions suggests linkage disequilibrium.

TABLE 4

FREQUENCIES OF HAPLOTYPES OF THE THREE LOCI, *Pr*, *Pa*, AND *Db*

HAPLOTYPES	OBSERVED (%)	EXPECTED (%)
$Pr^1\ Pa^1\ Db^1$	0	2.7
$Pr^1\ Pa^1\ Db^0$	0.6	12.4
$Pr^1\ Pa^0\ Db^1$	14.8	10.1
$Pr^1\ Pa^0\ Db^0$	56.3	46.5
$Pr^2\ Pa^1\ Db^1$	0.6	1.1
$Pr^2\ Pa^1\ Db^0$	19.7	4.9
$Pr^2\ Pa^0\ Db^1$	2.5	4.0
$Pr^2\ Pa^0\ Db^0$	5.4	18.3

However, there are no instances in man of linkage disequilibrium for loci 3 or more centimorgans apart. Thus, other explanations such as high selection pressure or genetic interaction must be entertained. We are presently investigating such possibilities.

The gene loci for the proline-rich salivary proteins, *Pr*, *Pa*, and *Db*, appear to constitute a complex of three genes having products of intriguing structural similarities. Such a system is reminiscent of the HLA complex on human chromosome 6 [16]. Azen et al. [17] and Azen and Denniston [18] have reported preliminary evidence that gene loci of three other parotid salivary proteins, major glycoprotein (Gl), parotid size variant (Ps), and middle band (Pm), are linked to *Pr* and *Db*. The core proteins of these three proteins (Gl, Ps, and Pm) are known to be proline-rich. Thus, the salivary protein gene complex may contain other proteins related to *Pr*, *Pa*, or *Db*. Existing data indicate distinct immunologic [8] and biochemical [9] relationships between the closely linked *Pr* and *Pa* protein gene loci; thus, the loci may be the result of gene duplication.

ACKNOWLEDGMENTS

We express our appreciation for the technical assistance provided by Judy E. Retz and Marge A. Craft and for the computing facilities of the Indiana University computing network.

REFERENCES

1. AZEN EA, OPPENHEIM FG: Genetic polymorphism of proline-rich human salivary proteins. *Science* 180:1067–1069, 1973
2. AZEN EA, DENNISTON CL: Genetic polymorphism of human salivary proline-rich proteins: further genetic analysis. *Biochem Genet* 12:109–120, 1974
3. FRIEDMAN RD, MERRITT AD, RIVAS ML: Genetic studies of human acidic salivary protein (Pa). *Am J Hum Genet* 27:292–303, 1975
4. MINAGUCHI K, IKEMOTO S, NAKAJAMA I, SUZUKI K: Studies of genetic markers in human saliva (I) frequencies of Pa and Pb systems from parotid saliva of Japanese in Tokyo. *Bull Tokyo Dent Coll* 17:185–190, 1976
5. MINAGUCHI K, IKEMOTO S, IIDA K, SUZUKI K: Studies of genetic markers in human saliva (II) frequencies of Pr and Db systems from parotid saliva of Japanese in Tokyo. *Bull Tokyo Dent Coll* 17:191–197, 1976
6. IKEMOTO S, MINAGUCHI K, HINOHAVA H: Genetic polymorphisms of human parotid salivary proteins (Pa, Pb, Pr, Db and Pm) and salivary amylase isozyme in Japanese population. *Hum Hered* 27:328–331, 1977
7. YU PL, SCHWARTZ RC, MERRITT AD, ET AL.: Linkage relationships of the proline-rich salivary proteins (Pr, Pa, Db), in *Human Gene Mapping 4*. National Foundation. *Cytogenet Cell Genet* 22:655–658, 1978
8. FRIEDMAN RD, KARN RC: Immunological relationships and a genetic interpretation of major and minor acidic proteins in human parotid saliva. *Biochem Genet* 15:549–562, 1977
9. FRIEDMAN RD, MERRITT AD: Partial purification and characterization of a polymorphic protein (Pa) in human parotid saliva. *Am J Hum Genet* 27:304–314, 1975
10. TAGGART RT, MILLER RB, KARN RC, ET AL.: Vertical thin layer slab polyacrylamide gel electrophoresis of selected human polymorphic proteins, in *Electrophoresis '78*, edited by CATSIMPOOLAS. New York, Elsevier North Holland, 1978, pp 231–242
11. SUNG M, SMITHIES O: Differential elution of histone from gel-trapped nuclei. *Biopolymers* 7:39–58, 1969
12. AZEN EA: Genetic polymorphism of basic proteins from parotid saliva. *Science* 176:673–674, 1972
13. OTT J: Estimation of the recombination fraction in human pedigrees: efficient computation

of the likelihood for human linkage studies. *Am J Hum Genet* 26:588–597, 1974
14. MEYERS DA: Multipoint mapping of linkage group I. Ph.D. thesis, Bloomington, Indiana Univ., 1977
15. KAPLAN EB, ELSTON RC: A subroutine package for maximum likelihood estimation (MAXLIK). Institute of Statistics mimeo series no. 823, Chapel Hill, Univ. North Carolina, 1973
16. BODMER WF, CHAIRMAN: Report of the Committee on the Genetic Constitution of Chromosome 6, in *Human Gene Mapping 3*. National Foundation. *Birth Defects: Orig Art Ser* XII:24–30, 1975
17. AZEN EA, HURLEY CK, DENNISTON C: Genetic polymorphism of the major parotid salivary glycoprotein (Gl) with linkage to the genes for Pr, Db and Pa. *Biochem Genet* 17:257–279, 1979
18. AZEN EA, DENNISTON C: Polymorphism of Ps (parotid size variant) and detection of a protein (PmS) related to the Pm (parotid middle band) system with genetic linkage of *Ps* and *Pm* to *Gl*, *Db*, and *Pr* genetic determinants. *Biochem Genet*. In press, 1980