

## Research Article

# Parkinson's Disease in Relation to Pesticide Exposure and Nuclear Encoded Mitochondrial Complex I Gene Variants

Elizabeth H. Corder<sup>1</sup> and George D. Mellick<sup>2</sup>

<sup>1</sup> Center for Demographic Studies, Duke University, 2117 Campus Drive, PO Box 90408, Durham, NC 27708-0408, USA

<sup>2</sup> Eskitis Institute for Cell and Molecular Therapies, Griffith University, Nathan, QLD 4111 Australia

Received 1 December 2005; Revised 28 March 2006; Accepted 17 April 2006

Parkinson's disease (PD) is a common age-related neurodegenerative disorder thought to result from the integrated effects of genetic background and exposure to neuronal toxins. Certain individual nuclear-encoded mitochondrial complex I gene polymorphisms were found to be associated with ~ 2-fold risk variation in an Australian case-control sample. We further characterized this sample of 306 cases and 321 controls to determine the mutual information contained in the 22 SNPs and, additionally, level of pesticide exposure: five distinct risk sets were identified using grade-of-membership analysis. Of these, one was robust to pesticide exposure (I), three were vulnerable (II, III, IV), and another (V) denoted low risk for unexposed persons. Risk for individual subjects varied > 16-fold according to level of membership in the vulnerable groups. We conclude that inherited variation in mitochondrial complex I genes and pesticide exposure together modulate risk for PD.

Copyright © 2006 E. H. Corder and G. D. Mellick. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## INTRODUCTION

Parkinson's disease (PD) is a common age-related neurodegenerative disorder thought to result from the integrated effects of genetic background and exposure to neuronal toxins (eg, MPTP, rotenone and other insecticides). Rarely, it is caused by mendelian mutations ( $\alpha$ -synuclein, DJ-1, PINK1, and LRRK2). Tremor, rigidity, slowness of movement and postural instability are the predominant symptoms due to the selective loss of pigmented dopamine-producing neurons in the substantia nigra pars compacta region of the brain. Whether familial, age-related, or consequent to neuronal toxin exposure, impaired mitochondrial complex I (NADH:ubiquinoneoxidoreductase; EC 1.6.5.3) is found [1–17]. Mitochondrial complex I consists of at least 46 subunits, seven encoded in the mitochondrial genome and the remainder, at least 39, encoded in the nucleus.

One testable hypothesis is that PD risk is modulated by inherited sequence variation in complex I genes. Mellick et al [1] addressed this possibility: they screened for single nucleotide polymorphisms (SNPs) in nuclear encoded complex I genes. A total of 22 SNPs (16 genes) polymorphic among Australians were investigated (306 PD patients; 321 control

subjects). Statistically significant associations, and ~ 2-fold variation in risk, were observed for *NDUF* genes A1, A10, A6, and S4, when taken individually. None of these associations would have survived correction for multiple comparisons. Although information on pesticide exposure, sex, and age at onset was available, it was not used in the association analysis due to the limited sample size and the larger number of multiple comparisons that would have been generated.

We extend this work by jointly investigating the 22 SNP genotypes found for these 306 PD patients and 321 control subjects, avoiding multiple comparisons, and considering level of pesticide exposure and age at onset. The goals were to identify combinations of alleles robust to pesticide exposure, others that are especially vulnerable, and to quantify risk for individuals. This was accomplished using grade-of-membership analysis or GoM.

Using GoM, two sets of parameters are simultaneously estimated by maximum likelihood (see "Methods"). One set represents a specified number of extreme pure type groups. Here, each of the five groups has distinct frequencies for the SNP genotypes, level of toxin exposure, and PD status according to age. The other set of parameters represents the degree of similarity of each subject to the groups. These

graded membership scores range from zero, that is, denoting no resemblance to the group, to one, that is, the individual matches the group exactly, and sum to one for each person. The scores can be input into logistic models to quantify disease risk and produce 95% CI. This approach has identified highly predictive sufficient genetic risk sets for Alzheimer's disease [18] and multilocus genotypes specific to breast cancer and fibroadenoma [19].

Five model-based groups relevant to PD status were identified. Of these, one set of complex I polymorphisms was robust to pesticide exposure (I), three sets were vulnerable (II, III, IV), and another (V) denoted low risk for unexposed persons. Risk for individuals varied > 16-fold according to level of membership in the vulnerable groups. We conclude that inherited variation in mitochondrial complex I genes and pesticide exposure together modulate risk for PD.

## METHODS

The specific details of the case-control sample and the genotyping methods used have been reported previously [1].

### Study subjects

Onset ranged over six decades of age among the 306 case subjects (decade of age: %): 30–39: 3%, 40–49: 12%, 50–59: 31%, 60–69: 36%, 70–79: 16%, 80+ : 2%. The cases did not have symptoms of other neurological conditions, for example, change in cognition suggestive of Alzheimer's disease onset. The 321 control subjects did not have any symptom of parkinsonism (decade of age: %): 30–39: 1%, 40–49: 6%, 50–59: 22%, 60–69: 36%; 70–79: 31%, 80+ : 4%.

### Environmental exposures

A structured questionnaire was used to probe for exposures to environmental toxins including insecticides, herbicides, fungicides, solvents, heavy metals [20]. However, this study limited exposure assessment to self-reported exposures to pesticides (ie, insecticides, herbicides, and fungicides). Responses were coded as 0 = no exposure, 1 = limited exposure, and 2 = regular exposure at least weekly for six months. Pesticide exposure was more common among men and cases (men: 63% vs 55%; women: 46% vs 39%), especially regular exposure (men: 20% vs 9%; women: 5% vs 3%).

### Genetic determinations

SNPs in nuclear genes that encode mitochondrial complex I proteins were identified from the HGVBbase as of July 2001 [1]. The 22 of 70 identified SNPs polymorphic among 16 randomly selected healthy Australian subjects are investigated (Table 1). SNP determinations were made using the DASH method [21, 22]. The major allele at each locus was coded as "a" and the minor allele "b" yielding genotypes *aa*, *ab*, and *bb*. Multilocus genotypes were coded, for example, *aa:ab*, for genes having more than one SNP. Infrequent (0.018) missing values were ignored in the data analysis.

TABLE 1: The genes and SNPs investigated.

Number	Gene	SNP (HGV base ID)
1	DLST	SNP000002340 (A/G)
2	NDUFA1	SNP000005157 (G/C)
3	NDUFA1	SNP000008196 (T/C)
4	NDUFA1	SNP000008197 (T/G)
5	NDUFA10	SNP000015174 (G/A)
6	NDUFA10	SNP000020002 (A/G)
7	NDUFA6	SNP000005146 (C/T)
8	NDUFA7	SNP000005158 (C/T)
9	NDUFA8	SNP000005147 (A/G)
10	NDUFA8	SNP000008968 (G/A)
11	NDUFB4	SNP000019034 (C/T)
12	NDUFB7	SNP000005144 (C/G)
13	NDUFB8	SNP000005127 (C/A)
14	NDUFB9	SNP000005142 (C/T)
15	NDUFS1	SNP000005158 (G/T)
16	NDUFS1	SNP000005159 (A/G)
17	NDUFS2	SNP000018866 (T/A)
18	NDUFS4	SNP000005133 (A/G)
19	NDUFS4	SNP000005178 (G/A)
20	NDUFS7	SNP000005156 (T/C)
21	NDUFS8	SNP000005155 (C/T)
22	NDUFV2	SNP000000182 (C/T)

### The data analytic approach

Detailed clinical genetic profiles were identified using grade-of-membership analysis or GoM [23–25]. Case subjects were considered according to age (< 60, 60–69, 70+) and environmental exposure (0, 1, 2); control subjects were coded according to exposure regardless of age, that is, 12 categories total.

GoM can be described after first identifying four indices. One is the number of subjects  $I$  ( $i = 1, 2, \dots, I$ ). Here,  $I = 627$  subjects were identified. The second index is the number of variables  $J$  ( $j = 1, 2, \dots, J$ ). There are  $J = 17$  variables. Our third index is  $L_j$ : the set of response levels for the  $J$ th variable. This leads to the definition of the basic GoM model where the probability that the  $i$ th subject has the  $L_j$ th level of the  $J$ th variable is defined by a binary variable (ie,  $y_{ijl} = 0, 1$ ). The model with these definitions is

$$\text{Prob}(y_{ijl} = 1.0) = \sum_k g_{ik} \lambda_{kjl}, \quad (1)$$

where the  $g_{ik}$  are convexly constrained scores (ie,  $0.0 \leq g_{ik} \leq 1.0$ ;  $\sum_k g_{ik} = 1.0$ ) for subjects and the  $\lambda_{kjl}$  are probabilities that, for the  $K$ th latent group, the  $L_j$ th level is found for the  $J$ th variable. The procedure thus uses this expression to identify  $K$  profiles representing the pattern of  $J \times L_j$  responses found for  $I$  subjects.

The parameters  $g_{ik}$  and  $\lambda_{kjl}$  are estimated simultaneously using the likelihood function (in its most basic form).

$$L = \prod_i \prod_j \prod_l \left( \sum_k g_{ik} \cdot \lambda_{kjl} \right)^{y_{ijl}} \quad (2)$$

In the likelihood  $y_{ijl}$  is 1.0 if the  $L$ jth level is present and 0.0 if it is not present. Decade of age provided starting values.

Information on sex was available to further characterize the groups. One option in the likelihood is to separate calculations for “internal” (here, clinical and genetic) and “external” (here, sex) variables. For internal variables, maximum likelihood estimations (MLE) of  $g_{ik}$  and  $\lambda_{kjl}$  are generated and the information in internal variables is used to define the  $K$  groups. For external variables the likelihood is evaluated (and MLE of  $\lambda_{kjl}$ ; generated) but the information is not used to redefine the  $K$  groups, that is, the likelihood calculations for likelihood equations involving the  $g_{ik}$  are disabled for external variables so that the  $g_{ik}$ , and the definition of the  $K$  groups, is not changed.

Next, three age-specific logistic models (< age 60, 60–69, 70+) were constructed to estimate the risk for PD according to membership in the vulnerable groups II, III, and IV. For this use, the graded membership scores were categorized from 1 (< 0.20 membership) to 5 (> 0.80 membership) representing 0.20 increments.

## RESULTS

Five GoM groups represent the data, displayed in Table 2. Group I was robust to toxin exposure. Groups II, III, and IV were vulnerable. Group V had limited toxin exposure and was at low risk. Each group had a distinctive set of SNP genotypes for nuclear genes that encode mitochondrial complex I subunits.

### *Robust to pesticide exposure (I)*

Low risk for PD despite limited toxin exposure carried a distinctive genetic signature of infrequent genotypes: X-linked  $A1$   $aa:ab:ab$  or  $bb:aa:aa$ ,  $A6$   $bb$ ,  $A8$   $bb:a-$ ,  $A10$   $ab:aa$ ,  $B4$   $ab$ ,  $B8$   $ab$  or  $bb$   $B9$   $ab$ ,  $S1$   $ab:aa$ , and  $S4$   $bb:ab$ . The group consisted predominantly of females (84% chance). There was some chance of being affected at ages 70 or older after regular exposure to toxins despite this protective signature.

### *Early onset, regular exposure (II)*

Group II was affected before age 60 and vulnerable to regular pesticide exposure. Its genetic signature consisted of  $A6$   $aa$ ,  $A8$   $ab:ab$ ,  $A10$   $aa:bb$ ,  $S1$   $bb:a$ , the common  $S2$   $ab$  genotype found also for group I, and  $DLST$   $bb$ .

### *Early onset PD, limited exposure (III)*

Group III also had high risk for PD with onset before age 60 (86%), at limited pesticide exposure (43%). This vulnerability was associated with  $A6$   $ab$ ,  $B7$   $bb$ , and  $DLST$   $aa$ . Homozygous  $V2$   $bb$  was found (22%) for this group only.

### *Late onset PD, limited exposure (IV)*

Group IV was at risk for PD (63%) at ages 60 to 69, at limited pesticide exposure (70%). It had genetic signature  $A8$   $aa:aa$ ,  $B7$   $ab$ ,  $S1$   $bb:bb$ ,  $S2$   $bb$ ,  $S4$   $bb:aa$ , and  $S8$   $bb$ .

### *Low risk, no exposure (V)*

Group V represents low risk for PD when not exposed to pesticides. Two SNP genotypes stand out as determinants:  $V2$   $ab$  (QRF = 1.47—the highest genetic influence score) and  $S4$   $aa:aa$  (QRF = 1.25). QRF stands for “question relevance score” denoting the relative importance of the variable in determining the group.

### *Informative variables*

No one genetic variable dominated. An information statistic  $H$ , related to Shannon’s information statistic (Bell Laboratories) was estimated for each variable: values close to zero indicate that the variable was not useful. Three of four SNPs deemed statistically significant in chi-square analyses ( $A6$  :  $H$  = 0.92,  $A10$  :  $H$  = 0.68,  $S4$  :  $H$  = 1.13) [1] were identified as being highly informative. The fourth,  $A1$ , had limited heterozygosity and low  $H$  score (0.19). Nonetheless,  $8196b + 8197b$  distinguished robust group I from the other groups. Additional loci were highly informative:  $B7$   $bb$  was associated with risk (III),  $aa$  with protection (I, V) ( $H$  = 1.07);  $A8$   $5147$   $bb$  was protective (I); risk was associated with  $8968$   $ab$  for persons exposed to pesticides (II) ( $H$  = 0.84), among others as shown in Table 1.

## RISK FOR PARKINSON’S DISEASE

Figure 1 shows the membership distributions of case and control subjects in each age group (< age 60, 60 to 69, 70+). Few subjects exactly matched the respective groups ( $N$  = 0, 1, 1, 3, 1). Most divided membership, for example, had SNP genotypes found for several of the extreme pure type groups shown in Table 1. Relatively few subjects had membership scores of 0.60 or higher (0.60–0.79: green; 0.80–1.00: gray). Nonetheless, several trends were apparent: cases occurring before age 60 tended to resemble groups II and III more than the control subjects. Cases at ages 60 to 69 resembled group IV. Cases at age 70 and older did not over-represent II, III, and IV. Instead, both cases and controls most strongly resembled group V, that is, there may be a survival advantage for this set of polymorphisms. Note the consonance of membership distributions for control subjects in each age group.

The odds of PD, for subjects in each age group, were predicted by membership in groups II, III, and IV in logistic models (Figure 2). Early onset PD was significantly predicted by groups II “early onset, regular exposure” and III “early onset, limited exposure” (OR (95% CI): 2.7 (1.8 to 4.0) and 4.0 (2.5 to 6.5)), respectively. Note that the confidence limits do not include the neutral reference value of one, which would denote no risk. Hence, even limited resemblance of subjects to either group, that is, membership score 0.20–0.39 versus < 0.20, carries statistically significantly increased risk.

TABLE 2: Probabilities for each variable outcome found for GoM groups I to V. Group I: “robust to pesticide exposure,” group II “early onset PD, regular exposure,” group III: “early onset PD, limited exposure,” group IV “late onset PD, limited exposure,” group V “low risk, no exposure.”

Variable & outcome	Sample	freq.	I	II	III	IV	V	H-value
Disease status & pesticide exposure								1.41
PD, < age 60	None	8.77	—	—	57	—	—	
	Limited	10.21	—	62	29	—	—	
	Regular	3.51	—	38	—	—	—	
PD, ages 60–69	None	8.61	—	—	—	31	—	
	Limited	7.02	—	—	—	25	—	
	Regular	2.07	—	—	—	7	—	
PD, ages 70+	None	4.63	—	—	—	—	<b>13</b>	
	Limited	3.03	—	—	—	—	<b>8</b>	
	Regular	0.96	8	—	—	—	—	
Control	None	28.07	—	—	—	—	<b>79</b>	
	Limited	20.26	85	—	—	37	—	
	Regular	2.87	7	—	14	—	—	
A1:5157-8196-8197	aa:aa:aa	90.16	57	100	100	90	94	0.19
	aa:ab:ab	5.41	3.7	—	—	—	—	
	aa:bb:bb	2.46	—	—	—	10	—	
	ab:aa:aa	1.15	—	—	—	—	6	
	bb:aa:aa	0.82	6	—	—	—	—	
A6:5146	aa	46.62	—	<b>100</b>	—	—	82	0.92
	ab	42.93	—	—	<b>100</b>	84	—	
	bb	10.45	100	—	—	16	18	
A7:5148	aa	68.76	100	100	32	100	—	0.57
	ab	27.21	—	—	49	—	100	
	bb	4.03	—	—	20	—	—	
A8:5147-8968	aa:aa	36.16	—	—	—	<b>88</b>	55	0.84
	ab:aa	34.36	42	57	100	—	—	
	ab:ab	14.50	—	43	—	—	32	
	bb:aa	7.17	46	—	—	—	—	
	bb:ab	6.35	12	—	—	<b>6</b>	13	
	bb:bb	1.47	—	—	—	<b>5</b>	—	
A10:15174-20002	aa:aa	21.64	—	<b>33</b>	31	15	29	0.68
	aa:ab	30.66	<b>23</b>	—	52	56	34	
	aa:bb	17.21	—	<b>67</b>	—	—	—	
	ab:aa	16.18	<b>77</b>	—	—	—	—	
	ab:ab	14.31	—	—	17	29	37	
B4:19034	aa	79.39	—	100	100	97	100	0.43
	ab	19.97	<b>100</b>	—	—	—	—	
	bb	0.64	—	—	—	3	—	
B7:5144	aa	25.69	100	—	—	—	100	1.07
	ab	50.57	—	100	—	100	—	
	bb	23.74	—	—	100	—	—	
B8:5127	aa	62.86	—	100	100	45	65	0.39
	ab	32.15	<b>71</b>	—	—	55	35	
	bb	4.98	<b>29</b>	—	—	—	—	
B9:5142	aa	92.08	50	100	100	100	100	0.16
	ab	7.92	50	—	—	—	—	

TABLE 2: Continued.

S1:5158-5159	<i>aa:aa</i>	30.54	13	35	—	58	31	0.65
	<i>ab:aa</i>	12.48	48	—	32	—	—	
	<i>ab:ab</i>	37.60	40	25	68	—	69	
	<i>bb:aa</i>	1.64	—	7	—	—	—	
	<i>bb:ab</i>	7.72	—	33	—	—	—	
	<i>bb:bb</i>	10.02	—	—	—	—	42	—
S2:18866	<i>aa</i>	44.64	—	—	100	53	100	0.81
	<i>ab</i>	45.62	100	<b>100</b>	—	—	—	
	<i>bb</i>	9.74	—	—	—	47	—	
S4:5133-5178	<i>aa:aa</i>	25.28	—	—	—	—	<b>97</b>	1.13
	<i>ab:aa</i>	26.09	—	50	68	—	—	
	<i>ab:ab</i>	25.77	—	38	32	57	—	
	<i>bb:aa</i>	7.62	—	—	—	43	—	
	<i>bb:ab</i>	11.67	100	—	—	—	—	
	<i>bb:bb</i>	3.57	—	12	—	—	—	<b>3</b>
S7:5156	<i>aa</i>	28.62	—	48	60	29	—	0.40
	<i>ab</i>	55.79	100	52	9	28	100	
	<i>bb</i>	15.59	—	—	31	42	—	
S8:5155	<i>aa</i>	69.03	100	83	66	25	85	0.21
	<i>ab</i>	25.84	—	17	34	52	15	
	<i>bb</i>	5.13	—	—	—	22	—	
V2:182	<i>aa</i>	66.07	100	100	78	100	—	0.60
	<i>ab</i>	30.68	—	—	—	—	<b>100</b>	
	<i>bb</i>	3.25	—	—	22	—	—	
DLST:2340	<i>aa</i>	24.23	—	—	<b>86</b>	—	—	0.72
	<i>ab</i>	57.84	100	18	<b>14</b>	100	100	
	<i>bb</i>	17.93	—	82	—	—	—	

Note: influential genotypes in determining the group are indicated in bold. The question relevance factor (QRF) score for that variable and group was > 1.20.

Each 0.20 increment multiplies risk: successive increments in group II membership carry risks of 2.67, 7.13, and 29.0. Successive increments in group III membership carry risks of 4, 16, and 64. Higher levels of risk are predicted by, for example, 0.5 membership in each of groups II and III.

Onset at ages 60 to 69 was significantly predicted by groups II “early onset, regular exposure” and IV “late onset, limited exposure” (OR (95% CI): 1.6 (1.1 to 2.3) and 2.63 (1.8 to 3.8)). Successive increments in group II membership carry risks of 1.6, 2.6, and 4.0. Successive increments in group IV membership carry risks of 2.63, 6.9, and 18.2. Higher levels of risk are predicted by, for example, 0.5 membership in each of groups II and IV.

The model was not predictive at ages 70 and older, that is, the global hypothesis that the parameter values were zero could not be rejected ( $P = .15$ ).

## DISCUSSION

The object of this study is the mutual information contained in multiple SNPs located in nuclear genes that encode

mitochondrial complex I subunits, level of toxin exposure, and Parkinson’s disease status according to age. A prior investigation of individual SNPs in the study sample [1] found relative risks of about two for PD associated with certain SNPs located in *NDUF* genes A1, A10, A6, and S4. This more inclusive analysis replicated these findings and yielded better estimates of risk in relation to the available information. Specifically, five model-based groups were identified that represented robustness to pesticide exposure (I), vulnerability to regular (II) and limited exposure (III, IV) and low risk in the absence of exposure (V). The robust group consisted predominantly of females and carried a set of less frequent alleles including one on the X chromosome. The vulnerable groups differed according to age at onset (< age 60 for II and III; age 60 to 69 for group IV) and level of toxin exposure (limited or regular for II, none or limited for groups III and IV). Even the low risk groups I and V had some level of risk for PD at ages 70 and older. Thus the mutual information investigated using GoM was more informative in terms of age, sex, and toxin exposure compared to straightforward association analysis.

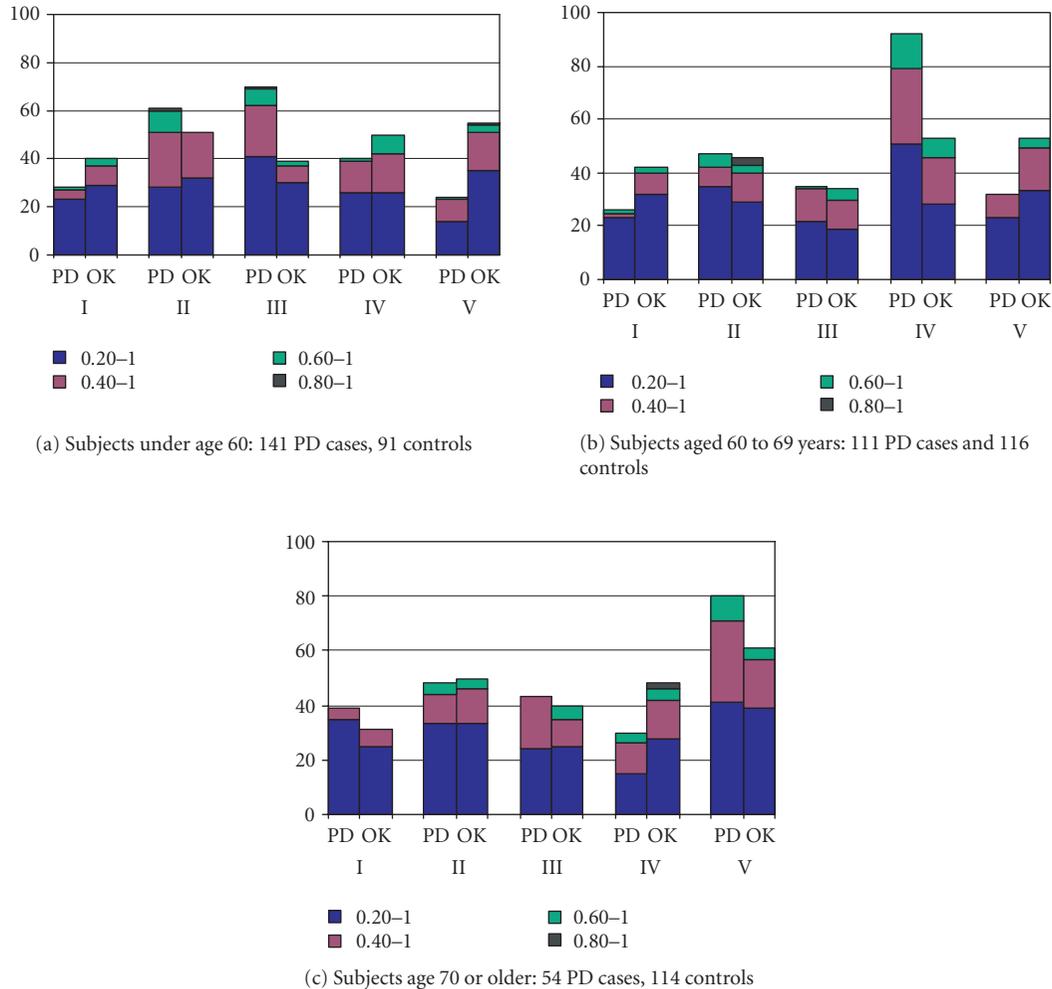


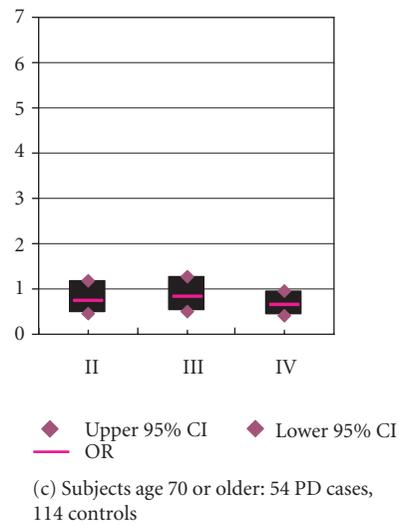
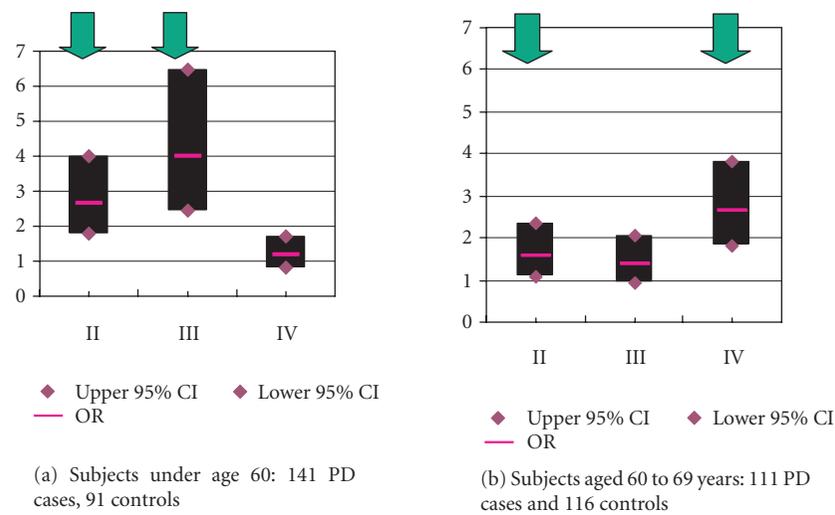
FIGURE 1: Individuals have assigned membership scores in the groups. These continuous scores ranging from 0 to 1 have been categorized as  $< 0.20$ ,  $0.20-0.39$ ,  $0.40-0.59$ ,  $0.60-0.79$ , and  $0.80$  to  $1.00$ . Control subjects in each age group have similar frequency distributions of membership. Case subjects  $<$  age 60 over-represent membership in groups II and III, group IV at ages 60 to 69. Cases at age 70 and older over-represent the groups I and V.

This approach, avoiding multiple comparisons which would have decimated each of the individual associations [1], was able to estimate risk for individuals according to the level of membership in the vulnerable groups. At ages  $<$  60, statistically significant 3-fold and 4-fold elevation in risk was found for persons who had limited ( $0.20-0.39$ ) resemblance to groups II and III, respectively, compared to those having very little ( $< 0.20$ ) resemblance. Successive increments in group II membership carry risks of 2.67, 7.13, and 29.0. Successive increments in group III membership carry risks of 4, 16, and 64. Higher levels of risk are predicted by, for example, 0.5 membership in each of groups II and III. Hence, clinically relevant and statistically significant results were obtained.

Taking the Rotterdam cohort as a guide, incidence increases from 0.3 per 1000 person-years at ages 55 to 65 years to 4.4 per 1000 person-years at ages 85 years and older [27]. The incidence of symptoms of parkinsonism was similar for men and women, but men more often met diagnostic criteria

(male-to-female ratio, 1.54; 95% CI, 0.95 to 2.51), hence, the great care taken in this study to consider age at onset. Because men were more often exposed to pesticides compared to women, sex, per se, was not used to determine the risk groups.

Biological interpretation is not straightforward, yet the results lend further credence to the belief that faulty combinations of mitochondrial complex I subunits pose significant risk for age-related PD, presumably, by reduced ATP production and increased production of reactive oxygen species. The results imply that certain persons are robust and others vulnerable to PD when exposed to pesticides. Measurement of pesticide exposure information was structured, but imperfect: memory fades; duration beyond six months was not investigated. One feature of GoM, the identification of extreme types, minimizes this problem by filtering a lack of fidelity in the data. To the extent that groups are misidentified, risk estimates would be expected to be biased toward the null,



↓ Denotes statistical significance: the 95% CI for the odds ratio does not include the neutral value of one.

(d)

FIGURE 2: The odds of PD were estimated for each age group (< age 60, 60 to 69, 70+) in logistic models. The predictors were membership scores in the vulnerable groups II, III, and IV. The scores were coded categorically from 1 to 5 representing 0.20 increments. Risk multiplies for each increment of 0.20. For example, at ages < 60, successive increments in group III membership carry risks of 4 (0.20–0.39), 16 (0.40 to 0.59), and 64 (0.60 to 0.79).

that is, underestimate risk. The ability of GoM to interpret mixtures of genetic, clinical, and pathologic data is further demonstrated in these referenced papers [28–34].

In summary, fuzzy latent class analysis was employed to identify sets of polymorphisms located in nuclear genes encoding mitochondrial complex I proteins associated with PD, and effect modification by toxin exposure. Even partial resemblance to the identified risk sets carried appreciable risk for PD. This form of analysis may prove a particularly useful way for hypothesis generation and subsequent investigation

of specific gene x gene and gene x environment interactions in relation to common sporadic PD.

**ACKNOWLEDGMENTS**

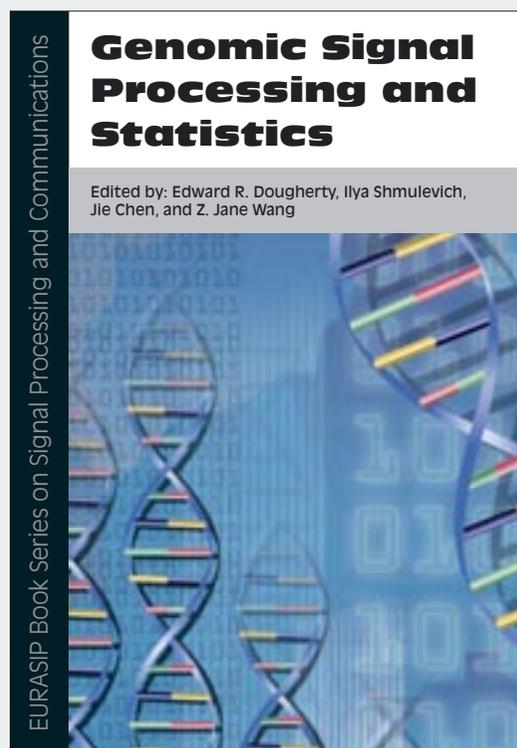
Dr Mellick is supported by the Geriatric Medical Foundation of Queensland. The authors acknowledge A. J. Brookes, J. A. Prince, and P. A. Silburn for their roles in producing the originally published case-control data which was supported by a STINT Foundation grant to Drs Mellick and Brookes.

## REFERENCES

- [1] Mellick GD, Silburn PA, Prince JA, Brookes AJ. A novel screen for nuclear mitochondrial gene associations with Parkinson's disease. *Journal of Neural Transmission*. 2004;111(2):191–199.
- [2] Liu Y, Yang H. Environmental toxins and alpha-synuclein in Parkinson's disease. *Molecular Neurobiology*. 2005;31(1–3):273–282.
- [3] Yu S, Ueda K, Chan P. Alpha-synuclein and dopamine metabolism. *Molecular Neurobiology*. 2005;31(1–3):243–254.
- [4] Son JH, Kawamata H, Yoo MS, et al. Neurotoxicity and behavioral deficits associated with Septin 5 accumulation in dopaminergic neurons. *Journal of Neurochemistry*. 2005;94(4):1040–1053.
- [5] Le Couteur DG, McLean AJ, Taylor MC, Woodham BL, Board PG. Pesticides and Parkinson's disease. *Biomedicine & Pharmacotherapy*. 1999;53(3):122–130.
- [6] Le Couteur DG, Muller M, Yang MC, Mellick GD, McLean AJ. Age-environment and gene-environment interactions in the pathogenesis of Parkinson's disease. *Reviews on Environmental Health*. 2002;17(1):51–64.
- [7] Mouradian MM. Recent advances in the genetics and pathogenesis of Parkinson disease. *Neurology*. 2002;58(2):179–185.
- [8] Meulener M, Whitworth AJ, Armstrong-Gold CE, et al. Drosophila DJ-1 mutants are selectively sensitive to environmental toxins associated with Parkinson's disease. *Current Biology*. 2005;15(17):1572–1577.
- [9] Ritz B, Yu F. Parkinson's disease mortality and pesticide exposure in California 1984–1994. *International Journal of Epidemiology*. 2000;29(2):323–329.
- [10] Smeitink JA, van den Heuvel LW, Koopman WJ, Nijtmans LG, Ugalde C, Willems PH. Cell biological consequences of mitochondrial NADH: ubiquinone oxidoreductase deficiency. *Current Neurovascular Research*. 2004;1(1):29–40.
- [11] Smeitink JA, van den Heuvel LW, DiMauro S. The genetics and pathology of oxidative phosphorylation. *Nature Reviews. Genetics*. 2001;2(5):342–352.
- [12] Sveinbjornsdottir S, Hicks AA, Jonsson T, et al. Familial aggregation of Parkinson's disease in Iceland. *The New England Journal of Medicine*. 2000;343(24):1765–1770.
- [13] Solano A, Roig M, Vives-Bauza C, et al. Bilateral striatal necrosis associated with a novel mutation in the mitochondrial ND6 gene. *Annals of Neurology*. 2003;54(4):527–530.
- [14] Vives-Bauza C, Andreu AL, Manfredi G, et al. Sequence analysis of the entire mitochondrial genome in Parkinson's disease. *Biochemical and Biophysical Research Communications*. 2002;290(5):1593–1601.
- [15] Hattori N, Yoshino H, Tanaka M, Suzuki H, Mizuno Y. Genotype in the 24-kDa subunit gene (NDUFV2) of mitochondrial complex I and susceptibility to Parkinson disease. *Genomics*. 1998;49(1):52–58.
- [16] Elbaz A, Levecque C, Clavel J, et al. CYP2D6 polymorphism, pesticide exposure, and Parkinson's disease. *Annals of Neurology*. 2004;55(3):430–434.
- [17] Deng Y, Newman B, Dunne MP, Silburn PA, Mellick GD. Further evidence that interactions between CYP2D6 and pesticide exposure increase risk for Parkinson's disease. *Annals of Neurology*. 2004;55(6):897.
- [18] Corder EH, Huang R, Cathcart HM, et al. Membership in genetic groups predicts Alzheimer disease. *Rejuvenation Research*. 2006;9(1):89–93.
- [19] Corder EH, Hefler LA. Multilocus genotypes spanning estrogen metabolism associated with breast cancer and fibroadenoma. *Rejuvenation Research*. 2006;9(1):56–60.
- [20] Gartner CE, Battistutta D, Dunne MP, Silburn PA, Mellick GD. Test-retest repeatability of self-reported environmental exposures in Parkinson's disease cases and healthy controls. *Parkinsonism & Related Disorders*. 2005;11(5):287–295.
- [21] Howell WM, Jobs M, Gyllensten U, Brookes AJ. Dynamic allele-specific hybridization. A new method for scoring single nucleotide polymorphisms. *Nature Biotechnology*. 1999;17(1):87–88.
- [22] Prince JA, Feuk L, Howell WM, et al. Robust and accurate single nucleotide polymorphism genotyping by dynamic allele-specific hybridization (DASH): design criteria and assay validation. *Genome Research*. 2001;11(1):152–162.
- [23] Woodbury MA, Clive J. Clinical pure types as a fuzzy partition. *Journal of Cybernetics*. 1974;4:111–121.
- [24] Woodbury MA, Clive J, Garson A. Mathematical typology: a grade of membership technique for obtaining disease definition. *Computers and Biomedical Research*. 1978;11(3):277–298.
- [25] Woodbury MA, Manton KG, Tolley DH. A general model for statistical analysis using fuzzy sets: sufficient conditions for identifiability and statistical properties. *Information Sciences*. 1994;1:149–180.
- [26] de Lau LM, Giesbergen PC, de Rijk MC, Hofman A, Koudstaal PJ, Breteler MM. Incidence of parkinsonism and Parkinson disease in a general population: the Rotterdam Study. *Neurology*. 2004;63(7):1240–1244.
- [27] Manton KG, Corder LS, Stallard E. Estimates of change in chronic disability and institutional incidence and prevalence rates in the U.S. elderly population from the 1982, 1984, and 1989 National Long Term Care Survey. *Journal of Gerontology*. 1993;48(4):S153–S166.
- [28] Jablensky A, Woodbury MA. Dementia praecox and manic-depressive insanity in 1908: a Grade of Membership analysis of the Kraepelinian dichotomy. *European Archives of Psychiatry and Clinical Neuroscience*. 1995;245(4–5):202–209.
- [29] Hallmayer JF, Kalaydjieva L, Badcock J, et al. Genetic evidence for a distinct subtype of schizophrenia characterized by pervasive cognitive deficit. *American Journal of Human Genetics*. 2005;77(3):468–476.
- [30] Corder EH, Woodbury MA, Manton KG, Field LL. Grade-of-membership sibpair linkage analysis maps the IDDM11 locus to chromosome 14q24.3-q31: a pattern recognition approach to linkage. *Annals of Human Genetics*. 2001;65:387–394.
- [31] Corder EH, Woodbury MA, Volkmann I, Madsen DK, Bogdanovic N, Winblad B. Density profiles of Alzheimer disease regional brain pathology for the Huddinge Brain Bank: pattern recognition emulates and expands upon Braak staging. *Experimental Gerontology*. 2000;35(6–7):851–864.
- [32] Corder EH, Ervin JF, Lockhart E, Szymanski MH, Schmechel DE, Hulette CM. Cardiovascular damage in Alzheimer disease: autopsy findings from the Bryan ADRC. *Journal of Biomedicine and Biotechnology*. 2005;2005(2):189–197.
- [33] Iivonen S, Corder EH, Lehtovirta M, et al. Polymorphisms in the CYP19 gene confer increased risk for Alzheimer disease. *Neurology*. 2004;62(7):1170–1176.
- [34] Helisalmi S, Hiltunen M, Vepsäläinen S, et al. Genetic variation in apolipoprotein D and Alzheimer's disease. *Journal of Neurology*. 2004;251(8):951–957.

# GENOMIC SIGNAL PROCESSING AND STATISTICS

Edited by: Edward R. Dougherty, Ilya Shmulevich, Jie Chen, and Z. Jane Wang



## Genomic Signal Processing and Statistics

Edited by: Edward R. Dougherty, Ilya Shmulevich, Jie Chen, and Z. Jane Wang

Recent advances in genomic studies have stimulated synergetic research and development in many cross-disciplinary areas. Genomic data, especially the recent large-scale microarray gene expression data, represents enormous challenges for signal processing and statistics in processing these vast data to reveal the complex biological functionality. This perspective naturally leads to a new field, genomic signal processing (GSP), which studies the processing of genomic signals by integrating the theory of signal processing and statistics. Written by an international, interdisciplinary team of authors, this invaluable edited volume is accessible to students just entering this emergent field, and to researchers, both in academia and industry, in the fields of molecular biology, engineering, statistics, and signal processing. The book provides tutorial-level overviews and addresses the specific needs of genomic signal processing students and researchers as a reference book.

Limited-Time Promotional Offer.  
Buy this title NOW at  
**20%** discount plus  
**Free Shipping.**

The book aims to address current genomic challenges by exploiting potential synergies between genomics, signal processing, and statistics, with special emphasis on signal processing and statistical tools for structural and functional understanding of genomic data. The book is partitioned into three parts. In part I, a brief history of genomic research and a background introduction from both biological and signal processing/statistical perspectives are provided so that readers can easily follow the material presented in the rest of the book. In part II, overviews of state-of-the-art techniques are provided. We start with a chapter on sequence analysis, and follow with chapters on feature selection, clustering, and classification of microarray data. The next three chapters discuss the modeling, analysis, and simulation of biological regulatory networks, especially gene regulatory networks based on Boolean and Bayesian approaches. The next two chapters treat visualization and compression of gene data, and supercomputer implementation of genomic signal processing systems. Part II concludes with two chapters on systems biology and medical implications of genomic research. Finally, part III discusses the future trends in genomic signal processing and statistics research.

*EURASIP Book Series on SP&C, Volume 2, ISBN 977-5945-07-0.*

Please visit <http://www.hindawi.com/spc.2.html> for more information about the book. To place an order while taking advantage of our current promotional offer, please contact [books.orders@hindawi.com](mailto:books.orders@hindawi.com)