Genetic Epidemiology 13:301–313 (1996)

ABSTRACTS FROM THE FIFTH ANNUAL MEETING OF THE INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY

RIO DE JANEIRO, BRAZIL
AUGUST 17 & 18, 1996

HOSTED BY:
Columbia University
College of Physicians and Surgeons
New York, NY, USA

Division of Biostatistics
Washington University School of Medicine
St. Louis, MO, USA

Instituto Oswaldo Cruz
Rio de Janeiro, Brazil

1

The Beta model for complex inheritance: IDDM and asthma

N.E. Morton

The β model for complex inheritance specifies a single logistic parameter. It is the only model under which paternal and maternal transmissions to sibs of specified phenotypes are independent. In available data on recurrence risks and identity by descent it is significantly more powerful than alternative methods. Additivity of loci on the logistic scale is not disproven.

IDDM typifies a clear dichotomy between normal and affected, with most of the information in affected x affected pairs. Asthma and correlated atopy represent more complicated phenotypes that are essentially quantitative, and extreme discordant pairs are most informative. We will present applications to IDDM as a benchmark (kindly provided by June Davies and John Todd) and to Wessex data on asthma, comparing the nonparametric approach with a two-locus disease model.

We have extended the β model to pairs of arbitrary relatives and to multipoint mapping and will compare the results with alternative methods.

2

Efficient strategies for genome scanning using maximum-likelihood affected sib-pair analysis

P.Holmans* & N.Craddock

(Dept. of Psychological Medicine, Univ. of Wales College of Medicine, Cardiff)

*Presenting Author

Detection of linkage using a systematic genome scan in nuclear families including an affected sibling pair is an important initial step on the path to cloning susceptibility genes for complex genetic disorders and it is desirable to optimise the efficiency of such studies. The aim is to maximise power whilst simultaneously minimising the total number of genotypings and probability of type I error. One approach to increase efficiency, which has been investigated by other workers, is grid tightening: a sample is typed initially using a coarse grid of markers and promising results are followed up using a finer grid. Another approach, not previously considered in the context of an affected sib-pair genome scan for linkage, is sample splitting: a portion of the sample is typed in the screening stage and promising results are followed up in the whole sample.
In the current study we have used computer simulation to investigate the relative efficiency of two stage strategies involving combinations of both grid tightening and sample splitting and found that the optimal strategy incorporates both approaches. In general, typing half the sample of affected pairs using a coarse (20cM) grid of markers in the screening stage is an efficient strategy under a variety of conditions. If the markers are informative, a 10cM grid is sufficient for the final stage, otherwise a 5cM grid is preferable. If Hardy-Weinberg equilibrium holds, it is most efficient not to type parents in the screening stage. If Hardy-Weinberg equilibrium does not hold (e.g. due to stratification) failure to type parents in the first stage increases the amount of genotyping required although the overall probability of type I error is not greatly increased, provided the parents are used in the final analysis.

3

Use of three Affected Sib-Pair methods to detect linkage between NIDDM and chromosome 20q

A. Philippi[1*], H. Zouali[2], E.H. Hani[3], N. Vionnet [3], M. Martinez[1], Ph. Froguel[3], F. Demenais[1].
[1] INSERM U. 358, Paris, [2] CEPH, Paris, [3] CNRS EP10, Lille, France.

Among four candidate regions investigated for linkage to non-insulin-dependent diabetes mellitus (NIDDM), the phosphoenolpyruvate carboxykinase (PCK1) gene, on chromosome 20q, showed suggestive evidence for linkage in 133 affected sib pairs from 55 French pedigrees with multiple NIDDM cases (p=0.006, using SIBPAL). A total of 738 family members (417 NIDDM cases) from 170 families (55 plus 115 families) were genotyped for eleven highly polymorphic markers, spanning a 64 cM region on 20q, which included the MODY1-linked-ADA locus. Linkage analysis was conducted with three affected sib-pair (ASP) methods which differ according to the test statistic used: (1) The Haseman-Elston method in SIBPAL; (2) The Terwilliger method in SIBPAIR; (3) The Risch-Holmans method in SPLINK. The two latter programs use the same weight to account for multiple affected sib pairs from the same sibship. The most significant results are obtained with the PCK1 and ADA loci in NIDDM pairs diagnosed before 46 years of age and with the ribophorin-II gene, RPN-II, in all NIDDM pairs:

Locus		no. of affected sib pairs	p values by ASP method (1)	(2)	(3)
PCK1	*(30 cM)**	55	0.001	0.008/0.010[‡]	0.001/0.002[‡]
ADA	*(11 cM)*	53	0.0002	0.0005/0.004	0.0002/0.002
RPN-II		270	0.001	0.002/0.009	0.0006/0.009

[‡]The p values are for unweighted and weighted sib pairs respectively
* Genetic distance between loci

These results show that weighting non-independent sib pairs increases substantially the p values, even in large samples. Linkage of at least two regions of 20q with NIDDM is suggested. A dense map is currently being constructed for multipoint mapping.

4

A META-ANALYSIS METHODOLOGY FOR COMBINING NON-PARAMETRIC SIBPAIR LINKAGE RESULTS

C Gu*, M Province, Z Li and DC Rao

Division of Biostatistics, Washington University School of Medicine, St. Louis, MO; Loyola University Med Center, Chicago, IL, USA

In parametric linkage studies, the lod scores are pooled across studies enabling pooled inferences, much in the spirit of meta-analysis of summary data. The need is great for similar methods for pooling non-parametric sibpair linkage results, especially in light of the increasing number of repeated linkage claims and disclaims, and the potentially inadequate power in many individual studies to detect oligogenic effects. Meta-analysis methodology is explored here accommodating different study designs, some employing quantitative traits (e.g. blood pressure) and some employing qualitative traits (e.g. clinical hypertension), under the assumption that the underlying (disease) trait loci are the same. The ability to pool across three commonly used sibpair methods are considered: the affected-sibpair method for dichotomous traits, the Haseman-Elston regression method, and the Risch-Zhang extreme discordant sibpair method for quantitative traits. Variation among individual studies, each based on a different sibpair method, is modeled using random effects models. A heterogeneity test is constructed to verify the "poolability" of the studies, followed by a pooled test for linkage. Practical issues of performing meta-analysis, such as publication bias, are discussed. Some guidelines are given for planning linkage studies and for reporting linkage results bearing potential future meta-analysis in mind.

5

Affected sib-pair strategies for late onset diseases: type I error and marker allele frequencies.

F. Tores[1], E. Barrillot[2] A. Philippi[1], G. Vaysseix[2] F. Demenais[1], M. Martinez[1]*.
[1]INSERM U.358, Paris, [2]INFOBIOGEN, Villejuif, France.

For diseases with unknown mode of inheritance, the affected sib-pair strategies offer robust tests. Linkage is inferred by testing whether the proportion of alleles at the genetic marker identical by descent (IBD), among affected sib-pairs (ASPs) is greater than random expectation. When parent genotype information is not available, large sample sizes are often required. Further decrease in power is expected when allele frequencies are misspecified. Various studies have investigated variation of power when incorporating linkage information from additional family members and/or when using tightly linked markers to permit inferring IBD status in the ASPs. The associated rate of false positives (FP) remains unknown. We investigated the effect of allele frequency specification on the FP rate through simulations of a qualitative trait in nuclear families with at least two affected sibs. Linkage analyses were conducted under three different ASP methods: Haseman-Elston't-test approach in SIBPAL; Terwilliger's chi-square test in SIBPAIR; Risch-Holmans' likelihood-ratio test in SPLINK [and in Mapmaker/SIB].Different situations were considered by varying; (1) s, the family sibship size (2) H, the marker(s) heterozygosity (3) the map density [number and distance, δ, between markers]; and (4) n, the family sample size. The simulated data were analyzed using no parental marker data under the true and false allele frequencies. The table shows, for different significance levels α, the FP rates when using a 4-allele marker, MA, when n=100 and s=4 (1 ASP and 2 unaffected sibs):

	True allele frequencies (H=75%)				False allele frequencies (H=44%)			
α =	5%	1%	0.5%	0.1%	5%	1%	0.5%	0.1%
SIBPAL	.0500	.0115	.0055	.0010	.4395	.2135	.1495	.0610
SIBPAIR	.0365	.0035	.0015	.0005	.4055	.1740	.1185	.0370
SPLINK	.053	.0125	.0065	.0025	*	*	*	*
SIB	.070	.0170	.0070	.0020	.4825	.2525	.1805	.0705

*does not apply since SPLINK estimates allele frequencies in the data.

While the estimated FP rates depend on the conditions (n, s, H), general trends can be outlined. The SIBPAIR approach is the most conservative. Under a given error on frequencies, higher FP rates are obtained with the SIBPAL than with the SIBPAIR approach. Adding a 4-allele marker (H=44%) tightly linked to MA leads to a slight decrease in the SIB FP rates (21% vs 5.5% for α=1% vs 0.5%). FP rates increase with the error on allele frequencies even for high informative marker maps.

6

Inflation of Type I Error when Lod Scores are Maximized over Genetic Parameters. Paula C. Abreu[1], Susan E. Hodge[1,2], David A. Greenberg[3]. [1]Biostatistics, Columbia U.; [2]NY State Psychiatric Institute; [3]Mt. Sinai Med. Center, New York, NY.

It is well-known that calculating lod scores under multiple genetic models will inflate type I error (α), yet the precise magnitude of this increase remains to be determined. We simulated 1,000 datasets of 20 nuclear families each, under specified genetic parameters. The disease and marker were unlinked. We maximized the maximum lod score (Zmax) over 10 penetrance values, always assuming the correct genetic model (dominant, D, or recessive, R). We determined how many datasets yielded a maximized Zmax greater than 0.59, 0.83, and 1.17 (size = .05, .025, and .01, respectively, for a 1-sided χ^2 with 1 d.f.). The table shows α values resulting from our simulations:

Cutoff Z	Size for χ^2 (1 d.f)	Generating model D80%	D50%	R80%	R50%
.59	.05	.086	.122	.084	.107
.83	.025	.047	.063	.044	.063
1.17	.01	.024	.038	.023	.032

(D80% indicates dominant with 80% penetrance, etc.) Thus, actual α is inflated by varying amounts, apparently depending heavily on genetic model and penetrance. We did not observe any maximized Zmax values over 3.0 in our simulations, and no more than 4/1000 over 2.0. We will also present results under other models and penetrances, as well as what happens when we maximize Zmax over more than one genetic model.

7

Information provided by genealogical extensions to search for genes involved in heterogeneous rare dominant diseases.

GENIN E*, CLERGET-DARPOUX F
INSERM U155 Paris, France.

To study heterogeneous diseases in which different dominant mutations of genes located in different regions of the genome are involved, a strategy consists in performing genealogical extensions to detect affected relatives that have inherited the same mutation. The strategy is useful for diseases with late age of onset or short survival time since for these diseases, only poor information is obtained from close relatives of probands. If a genealogical extension is performed, we can expect to identify a living affected relative of the proband and thus to dispose of a pair of distantly affected relatives. Such a pair can provide information on the region of the genome where the mutated gene maps. When we already know, from previous linkage studies, that a gene located in a region R of the genome is involved in the disease, the pair of affected relatives permit to determine if it is indeed a mutation on this gene that is involved in the disease. Once the implication of the gene in region R has been confirmed, some information may be obtained to reduce the length of the region R. We have derived the expected lod score analytically and the expected length of the region where the mutation is suspected to map as a function of the number n of meioses between the two affecteds of the pair and of the polymorphism of the markers available in the region.

8

Contribution of consanguinity to the genetic study of multifactorial diseases.

GENIN E*, CLERGET-DARPOUX F
INSERM U155 Paris, France.

To study genetic factors involved in multifactorial diseases, methods have been developed in panmictic populations. However, for some diseases, it may be of interest to perform studies in consanguineous populations. In this context, the role of genetic factors may be easier to evidence.

We have studied the statistical properties, in presence of inbreeding, of two methods that permit to evidence the role of a candidate gene in a multifactorial disease using genetic markers close to the gene. These two methods are 1) association test on marker genotypes and 2) linkage test on affected sib pairs. We have investigated their robustness and power in presence of inbreeding. Ignoring inbreeding when it exists does not significantly increase the rate of false positives for both methods. However, the presence of inbreeding may increase the power to detect the role of a disease susceptibility factor. When cases and controls are assumed to be drawn from a population with a mean inbreeding coefficient F, the power of tests of association is increased, especially if the susceptibility factor has a recessive or quasi recessive mode of action. For the sib pair test of linkage, we have proposed a new statistic (na test) that jointly takes into account inter and intra-individual identity of marker alleles. Considering inbred sib pairs, we have shown that accounting for the intra-individual identity increases the power to detect the role of a susceptibility factor. The gain in power depends on the relative penetrances and allele frequencies at the disease susceptibility locus.

It was already known that it may be efficient to 'take advantage of inbreeding' to map genes involved in rare recessive diseases. We show here that it may also be efficient to evidence genetic risk factors in multifactorial diseases.

9

RANDOMIZATION TESTS TO CONTROL FALSE-POSITIVES IN GENOMIC SCANS FOR QTLs.

A.A Todorov* & D.C. Rao

Division of Biostatistics, Washington University School of Medicine, St. Louis, USA.

We have evaluated the role of randomization tests as robust alternatives to asymptotic statistical tests in genomic scans for QTLs using sibpair data and the Haseman-Elston method. The test is performed by repeatedly simulating vectors of identity-by-descent at the marker loci, computing the regression statistic at each, and comparing the resulting simulated profiles to the observed one. Our results indicate that this procedure is conservative, as would be expected.

Randomization tests can be used to reduce the proportion of spurious peaks within a single profile. Profiles are simulated conditional upon the marker information at the locus with the highest peak. This allows an assessment of the probability of observing the other peaks solely due to chance. A particular peak is retained when it cannot be ruled out by any conditional profile. This procedure results in a significant decrease in the number of regions that have to be considered further without much compromise on power.

In a preliminary study, we simulated profiles for 250 sibpairs with one sib in the upper tail, assuming one common recessive gene. In a 100cM region with markers every 10cM, we detected 821 positives in 1,500 samples at $p < 0.0001$ which included 15 false positives. Taking the highest peak whenever $p < 0.01$ and rejecting others using conditional profiles yielded 1,378 true and 98 false positives (with much higher power).

10

Homozygosity mapping : lod score and posterior probability of linkage.

GENIN E*, CLERGET-DARPOUX F
INSERM U155 Paris, France.

The strategy of homozygosity mapping proposed by Lander & Botstein (Science 236, 1987) has proven to be very useful to map genes involved in rare recessive diseases. It consists in searching for a region of the genome where independent inbred affecteds are homozygous. The disease locus is then likely to map in this region. To determine if the observed homozygosity confers good evidence of linkage, a lod score may be computed. In each family, this lod score depends on the inbreeding coefficient F of the affected, on the frequency p_1 of the marker allele for which the affected is homozygous and on the frequency q of the mutated allele. For a given set of (p_1,q), the lod score is not a monotonous function of F but it reaches a maximum for $F^*=(p_1q)^{1/2}/(1+(p_1q)^{1/2})$. If $p_1=0.25$ and $q=10^{-3}$, F^* is 0.016; it is approximately the inbreeding coefficient of a progeny of second cousins. Thus, for this set (p_1,q) of parameters values, more remote consanguinities will not provide more information in term of linkage evidence.

The computation of lod scores on families with one inbred affected sib leads to a problem of statistical significance. We have computed the posterior probability of no linkage after a lod score of +3 is obtained, i.e. the probability that linkage does not exist given linkage is claimed. It depends on p_1 and on the familial structure. For a sample of affected progenies of first-cousins that gives a maximum lod

score of +3, the posterior probability of no linkage is 16% if p_1=0.1 and 21% if p_1=0.4. Thus, when the frequencies p_1 of marker alleles for which affecteds are homozygous increase, the reliability of the test decreases. To ensure a posterior probability of no linkage smaller than 5%, a threshold value of the lod score greater than the conventional +3 should then be used to conclude linkage.

11

Evidence for Interaction between *IDDM3* and *IDDM11*. L Field, J Nagatomi, G Thomson*, R Tobias, Z Zhang. University of Calgary, Calgary, Canada and *University of California at Berkeley, Berkeley, USA.

We recently reported a novel susceptibility locus for insulin-dependent diabetes, designated *IDDM11*, detected by highly significant linkage to D14S67 on chromosome 14q (affected sib pair P = 10^{-5}) (Genomics 33:1-8,1996). Almost all the evidence for linkage came from families in which HLA sharing in affected siblings was ≤50%, suggesting that *IDDM11* acts biologically independently of HLA. We have now analyzed joint sharing of genes at D14S67 and D15S107, the latter near the *IDDM3* locus on chromosome 15q which we reported also showed stronger linkage in families with less HLA predisposition (Nat.Genet.8:189-94,1994). The objective was to determine if there was evidence for interaction between *IDDM11* (D14S67 marker) and *IDDM3* (D15S107 marker). 237 families were classified as having ≤50% (neutral/decreased) or >50% (increased) mean gene sharing at D14S67 and D15S107 in their affected sib pair(s). Results showed that families with increased sharing at D15S107 had a higher frequency of increased sharing at D14S67 than families with neutral/decreased sharing at D15S107 (P = 0.12). In other words, there was a tendency for increased sharing to occur at both loci simultaneously, suggesting possible biological interaction between *IDDM11* and *IDDM3*. In the total dataset, there was no association between diabetes and D14S67 using the family-based AFBAC test (AJHG 57:487-98, 1995). However, in families with increased sharing at D15S107, there was significant association of diabetes with D14S67 (P = 0.015). These findings suggest that although *IDDM11* and *IDDM3* may each act independently of HLA, they may themselves interact epistatically.

Supported by the Medical Research Council of Canada.

12

VARIANCE COMPONENTS MODELLING OF QUANTITATIVE TRAITS RELATED TO THE ASTHMA PHENOTYPE IN A POPULATION-BASED SAMPLE OF NUCLEAR FAMILIES.

LJ Palmer*, PR Burton, LC Gurrin, JA Faux, PN LeSouëf, AL James, G Ryan, WOCM Cookson, AW Musk.

TVW Telethon Institute for Child Health Research, PO Box 855, West Perth, Perth, W.A. 6872 Australia.

Asthma is associated with 3 primary pathophysiological factors: atopy, airway structure and airway responsiveness. The familial determinants and inter-relationships of quantitative traits reflecting these factors were investigated in a general population sample of families from Busselton, Western Australia. **Methods:** A cohort of 1020 individuals comprising 234 nuclear families was comprehensively evaluated. The quantitative traits assessed included total serum IgE levels, specific IgE levels, eosinophil levels, FEV_1, FVC and PD_{20} to methacholine. Maximum likelihood variance components modelling using *Fisher* was used to partition phenotypic variation into genetic and non-genetic components. **Results:** Log_e total serum IgE level exhibited the greatest narrow-band heritability (h^2_N) (h^2_N=49.2%±5.6% [±SE]: σ^2_A=1.38, σ^2_{TOTAL}=2.80). Log_e PD_{20} to methacholine (h^2_N=28.5%±5.5%: σ^2_A=0.42, σ^2_{TOTAL}=1.46), eosinophil levels (h^2_N=18.4%±4.2%: σ^2_A=2.12, σ^2_{TOTAL}=11.52) and FVC (h^2_N=11.4±1.6%: σ^2_A=0.031, σ^2_{TOTAL}=0.272) demonstrated moderate levels of heritability. FEV_1 exhibited relatively low heritability (h^2_N=3.4±10.4%: σ^2_A=0.031, σ^2_{TOTAL}=0.91). All estimates were adjusted for age and gender. Extended modelling suggested that the data were consistent with only a small overlap in the additive genetic determinants of these quantitative traits. However, there was a significant overlap in the additive genetic determinants of total and specific serum IgE levels. **Conclusions:** Atopy (IgE levels), airway structure (FEV_1 and FVC levels), and airway responsiveness (PD_{20}) vary in the extent of their determination by additive genetic effects. Furthermore, the additive genetic determinants of these factors are mainly unshared. These data are consistent with the existence of at least 3 distinct genetic pathways to asthma.

13

Genotype and the long-term course of rheumatoid arthritis. Lindsey Criswell*, Carol Such, Hua Mu, Mary-Claire King. Department of Medicine, UCSF; and Departments of Medicine and Medical Genetics, University of Washington, Seattle, USA.

Susceptibility and severity of rheumatoid arthritis (RA) have been linked to the inheritance of HLA-DRB1 alleles encoding a "shared epitope" (SE). We examined the relationship between the # and type of SE copies inherited and long-term outcomes of 149 community-based Caucasian females with RA followed annually for up to 12 years. Outcomes examined were physician assessment of RA course (0-2); annual measures of pain (0-100), function (0-3), # of painful joints (0-12), history of joint surgery (TJR), and resource utilization. Models accounted for correlation among serial observations for the same patient, and differences in patient age and disease stage. Results for our basic shared epitope model reveal strong associations between longitudinal outcomes and the # of SE copies inherited. An additive model of risk is apparent for some, but not all outcomes. In other models, outcomes varied according to the specific SE genotype inherited.

Outcome	β (P), SE=1	β (P), SE=2
Function	0.35 (0.0000)	0.20 (0.038)
Joint count	0.77 (0.04)	0.15 (0.74)
Pain	8.80 (0.002)	4.64 (0.16)
# visits	2.91 (0.004)	-0.20 (0.87)
	OR (P), SE=1	OR (P), SE=2
RA course	4.31 (0.0003)	4.79 (0.0010)
TJR surgery	2.13 (0.11)	3.47 (0.02)
RA hosp.	2.60 (0.02)	3.36 (0.013)

14

Mapping a susceptibility locus controlling infection intensities by *Schistosoma mansoni*.
L. Abel[1]*, S. Marquet[2], E.M. Carvalho[3], D. Hillaire[2], A. Prata[4], A. Dessein[2]. [1]INSERM U436, Paris, [2]INSERM U399, Marseille, France, [3]Lab. Immunol., Salvador, Brazil, [4]Fac. Med. Triangulo Mineiro, Uberaba, Brazil.

Schistosomiasis causes great suffering for millions of people in subtropical countries. The intensity and the pathological consequences of schistosome infections in subjects living in hyperendemic areas have been shown to depend on their intrinsic resistance to infection. In particular, our segregation analysis performed in a Brazilian population gave evidence for a major gene controlling infection intensities by *Schistosoma mansoni* (Abel L et al, Am J Hum Genet, 48, 959, 1991).

The present study was conducted in the same Brazilian families. Infection intensities were measured by the individual fecal egg counts used in our first analysis. The first step was to refine the segregation analysis by taking into account the covariates (contact with infested water, age, sex) simultaneously with the major gene effect. Results showed again the presence of a codominant major gene accounting for 66% of the infection intensity variance residual from covariates. Further lod-score analyses were performed with the major gene parameters estimated from segregation analysis.

Genotyping was carried out on 160 subjects belonging to 11 informative families of the original sample. The

primary map consisted in 245 microsatellite markers from the GENETHON panel. All markers gave lod-scores lower than 1.9, the threshold described as suggestive linkage in a genome-wide scan, except 2 adjacent markers which provided values greater than 3. Among additional markers presently analyzed in this region, one yielded a maximum lod-score of 4.52 at θ=0.04. This region contains several candidate genes encoding immunological molecules that have been demonstrated to be critical for protection against schistosomes. To our knowledge, this is the first lod-score analysis reporting the genetic localization of a major locus controlling human susceptibility/resistance to an infectious pathogen.

15

Genetic Analysis of Indicators of Cholesterol Synthesis and Absorption: Lathosterol and Phytosterols in Dutch Twins and Their Parents

D.I. Boomsma[1], H.J.M. Kempen[2,3] and J.A. Gevers Leuven[2]

Significant familial aggregation was observed for plasma levels of lathosterol (an indicator of whole-body cholesterol synthesis) and plant sterols campesterol and β-sitosterol (indicators of cholesterol absorption) in 160 Dutch families consisting of adolescent twins and their parents. For lathosterol a moderate genetic heritability in parents and offspring was found. In addition, shared environment also contributed significantly to plasma lathosterol concentrations in siblings. However, a model with different genetic heritabilities in the two generations fitted the data almost as well. For plasma plant sterol concentrations high genetic heritabilities were found. For campesterol heritability was 80% and for β-sitosterol it was 73%, without evidence for differences in heritability between generations. No influence of common environmental influences shared by family members was seen for either campesterol or β-sitosterol.

[1] **Department of Physiological Psychology, Vrije Universiteit, Amsterdam, The Netherlands.**
[2] **Gaubius Laboratory TNO-PG, Leiden, The Netherlands.**
[3] **Present address: Hoffman-LaRoche, CH-4070 Basel, Switzerland.**

Funded by the Netherlands Heart Foundation (86.083 and 88.042).

16

AN OLIGOGENIC MODEL FOR RESTING METABOLIC RATE

T Rice, A Tremblay, O Dériaz, L Pérusse, DC Rao*, and C Bouchard

Division of Biostatistics, Washington University School of Medicine, St. Louis, MO, U.S.A.

Evidence for a putative major gene was examined for resting metabolic rate (RMR) measured in 80 nuclear families participating in phase 2 of the Québec Family Study. Segregation analysis (POINTER) was conducted both prior to and after adjusting for several covariates, primarily fat mass (FM) and fat-free mass (FFM). Prior to covariate adjustment, the major gene evidence was unclear--i.e., there was either a major (Mendelian) or a multifactorial (polygenic and/or familial environment) effect. However, after adjusting the RMR for FM and FFM, a major gene hypothesis was clear. The major locus accounted for 57% of the variance, affected 7% of the sample, and led to high values of RMR. From previous studies we know that there is a putative major gene for FM and a major non-Mendelian effect for FFM (Rice T., Am J Hum Genet, 52, 967, 1993). Based on these combined results, a plausible hypothesis is that (1) the gene(s) affecting body size and body composition may have an effect on RMR (i.e., genetic pleiotropy), and (2) there is an additional major locus affecting only RMR (i.e., RMR appears to be an oligogenic trait).

17

Heritability of body mass index in a black population in a low risk environmental setting: Rotimi CN, Li Z, Ogunbiyi JO, McGee D, Cooper RS. Loyola University Medical Center, Department of Preventive Medicine, 2160 South First Ave, Maywood, IL 60153.

The problem of obesity has reached epidemic proportions in most technologically developed nations. In contrast however, most developing nations including those of Sub-Saharan Africa are still dealing with problems of inadequate food supply and the health consequence of under nutrition. Within the framework of an ongoing international collaborative study we examined familial patterns of the degree of heaviness as measured by body mass index (BMI) in a lean (mean BMI: male=21.7; females=22.6 kg/m^2) population sample of 223 nuclear families constituting 689 persons from Ibadan, Nigeria. BMI was adjusted for the effects of age by regressing it on up to a cubic polynomial in age separately for males and females. Age terms significant at the 5% levels were retained in the final models. The phenotype was then created from the best regression model by standardizing the age adjusted BMI to zero mean and unit variance. The phenotype was then ranked and normalized. The computer program SEGPATH (Province and Rao, 1995) which uses maximum likelihood method was used to estimate eight familial corrections: Father-mother (fm), father-son(fs), father-daughter(fd), mother-son(ms), mother-duaghter(md), son-son(ss), daughter-daughter(dd) and son-daughter(sd). Ten different model were fitted to test specific hypotheses using the likelihood ratio test. The estimated correlations are: fm=0.13, fs=0.22, fd=0.14, ms=0.19, md=0.43, ss=0.45, dd=0.57 and sd=0.34. Familial effect (ie., heritability) was estimated from the most parsimonious model as 45%; most of which may be due to genes since the spouse correlations were not significantly different from zero.

18

IDENTIFICATION OF NEW POLYMORPHISMS OF THE ANGIOTENSIN I-CONVERTING ENZYME (ACE) GENE, AND STUDY OF THEIR RELATIONSHIP TO PLASMA ACE LEVELS BY TWO-QTL SEGREGATION-LINKAGE ANALYSIS

L. Tiret,[1*] E. Villard,[2] S. Visvikis,[3] R. Rakotovao,[1] F. Cambien,[4] F. Soubrier [2]
INSERM [1] U258, [2] U358, [4] SC7, Paris, [3] CMP, Nancy, France

Plasma ACE levels are highly genetically determined. A previous segregation-linkage analysis suggested the existence of a functional mutation located within or close to the ACE locus, in almost complete linkage desequilibrium (LD) with the ACE I/D polymorphism and accounting for half the ACE variance. In order to identify the functional variant at the molecular level, we compared ACE gene sequences between 4 subjects selected for having contrasted ACE levels and I/D genotypes. We identified 10 new polymorphisms, among which 8 were genotyped in 95 healthy nuclear families, in addition to the I/D polymorphism. These polymorphisms could be divided into two groups: 5 polymorphisms in the 5' region and 3 in the coding sequence and the 3' UTR. Within each group, polymorphisms were in nearly complete association, whereas polymorphisms from the two groups were in strong negative LD. After adjustment for the I/D polymorphism, all polymorphisms of the 5' group remained significantly associated with ACE levels, suggesting the

existence of two quantitative trait loci (QTL) acting additively on ACE levels. Segregation-linkage analyses including one or two ACE-linked QTLs in LD with two ACE markers were performed to test this hypothesis. The two QTLs and the two markers were assumed to be in complete LD. Results supported the existence of two ACE-linked QTLs which would explain together 38% and 49% of the ACE variance in parents and offspring, respectively. One of these QTLs might be the I/D polymorphism itself, or the newly characterized $4656(CT)_{n}$ polymorphism. The second QTL would have a frequency around 0.20 incompatible with any of the yet identified polymorphisms.

19

MOLINESS, MELANOMA AND CDKN2: A TWIN STUDY

Martin NG*, Green AC, Ward PJ, Queensland Institute of Medical Research, Brisbane, Australia

The incidence of melanoma is rising rapidly throughout the western world, and nowhere more than Queensland which has a particularly susceptible, largely Celtic population living with very high UV exposure. Whether the heritability of melanoma and its risk factors will be higher or lower in the face of such high prevalence is an empirical question. Recently a major gene for familial melanoma (CDKN2,MTS1,MLM,p16) has been mapped to chromosome 9p but it is not yet clear what proportion of melanoma risk is attributable to it. We do know that having a high density of common moles (melanocytic nevi) confers a greatly increased risk of melanoma. We have mapped and counted moles on the skin of 350 random twin pairs in SE Queensland, first when they are 12 and later when they are 14 years old. A multivariate genetic analysis of mole counts at different body sites and their covariation with measures of sun exposure will be presented. Duplicate counts by different observers and separate counts for left and right sides of the body can be used to differentiate measurement error, developmental instability, and other measured and unmeasured sources of individual variation. Polymorphic markers have been typed at CDKN2 and sib pair linkage analysis performed to estimate the proportion of variance in moliness due to variation at this locus.

20

HIGLY PENETRANT GENES FOR EARLY ONSET BREAST CANCER PROSPECTIVELY VERIFIED

P Møller*, K Heimdal, A Dørum, L Mæhle, S Tretlie.
*Unit of Medical Genetics, The Norwegian Radium Hosptial, N-0310 Oslo, Norway.

The penetrance figures for breast cancer genes are derived from retrospective studies subjected to ascertainment errors. We have defined 1124 healthy women at risk for breast cancer according to clinical criteria, and are following these women prospectively. The cohort has a mean age of 42.9 +-11.3 years, 611 have been examined once, 513 have been followed for a mean of 1.8 years.

By March 8th 1996 our files shows that we have found 18 infiltrating cancers (see accompanying abstract to ICHG 96). In addition, we found 16 precancers (ca in situ/atypia).

Expectated number of infiltrating cancers were calculated. For each woman the Swedish figures for age-specific time in preclinical detectable stadium were applied, time for actual follow up were added, and the sum multiplied by the Norwegian age-specific incidences for breast cancer. Expected number of sporadic cancers were 3.3.

Our cohort had an assumed a priori risk of 25% for being gene carriers at birth. Applying 2% penetrance per adult year (Easton et al 1995), and correcting for reduced carrier probability in older women, expected numbers of inherited cancers to be found were 13.1. The exact age-related penetrance figures from Easton et al with a peak penetrance of 3.2% for age group 40-49 years, gave 12.7 expected inherited cancers. The best fit to the given model was obtained by assuming 4% penetrance in the age group 35-54 years.

The high incidence of precancers, as well as arguments for assumed a priori risk, possible errors and alternative interpretations will be discussed.

21

A Population-Based Study of Risk Factors in Male Breast Cancer. Hoda Anton-Culver, Daniela Seminara*, Tom Kurosaki, Thomas H. Taylor, Epidemiology Division, University of California Irvine.

A case-control study of male breast cancer was undertaken in Orange, San Diego, and Imperial Counties, California. Comparisons were made between 163 male breast cancer cases diagnosed in these counties from 1984 to 1994 and male population controls group-matched by age. Factors such as family history of cancer in first-degree relatives, personal medical history and use of medications, history of x-ray and radiation exposures, smoking history, and body mass index were compared between cases and controls. Family history of cancer of all types and of breast/prostate cancer are significantly increased in male breast cancer cases (both $p<0.001$), and the increase is consistent in all age groups. Family history of breast, ovarian, prostate, or endometrial cancers is seen in 27% of cases vs 8% of population controls ($p<0.001$). X-ray and radiation exposures were significantly increased in male breast cancer patients ($p< 0.001$). Smoking history is significantly higher in male breast cancer cases ($p<0.001$), as well as personal history of diabetes ($p<0.025$) and hypertension ($p<0.01$), while Klinefelter's syndrome is not a significant risk factor. Cox regression analyses to identify factors related to survival in male breast cancer show stage of disease, but not family history of cancer or multiple primary cancers, as an independent factor related to survival. These results confirm the association between family history of cancer and male breast cancer while providing preliminary data on the association between personal and family cancer history and survival.

22

Quantifying Familial Aggregation of Brain Tumors. F. Seillier-Moiseiwitsch[1], B. Newman[1], H. Pinheiro[1], M. Daly[1], S. Carozza[1], G. Bonney[2], M. Wrensch[3]. [1]University of North Carolina Chapel Hill, NC, [2]Fox Chase Cancer Center, Philadelphis, PA, and [3]University of California, San Francisco, CA, USA.

One of the few recognized risk factors for brain neoplasms is a family history, although the importance of genetic influences is not known. A population-based, case-control study was conducted among adults 20 years and older residing in 6 counties of the San Franciso Bay Area in Northern California. Patients diagnosed with an incident, primary malignant glioma were identified using rapid case ascertainment (n=462); comparison subjects were frequency-matched on age, gender, and ethnicity and identified through random-digit dialing (n=443). Structured interviews with subjects or their proxies obtained detailed information on parents, siblings, and children. Comparison between cases and controls using

logistic regression revealed a modestly increased risk of brain tumor in a first-degree relative (odds ratio=1.7; 95% confidence interval: 1.0-3.2). When analyzed by regressive models, strong evidence for familial aggregation was obtained; the best-fitting model contained parameters for the affection status of the mother, father, and siblings. When segregation analyses were attempted using REGD models in SAGE, evidence for a recessive mode of inheritance was obtained, but problems with convergence and plausibility were encountered. Analyses currently are underway with compound regressive models to determine whether they will allow better estimation of the genetic mechanism underlying the observed familial aggregation. Discovering important environmental risk factors also may benefit from understanding the role played by inherited susceptibility in brain cancer risk.

these were diagnostic of the MJD/SCA3 chromosome 14q24.3 expansion and one was diagnostic of SCA1 chromosome 6p22-p23. Among the three (3)remaining probands one had a positive family history consistent with SCA,and two (2) had no prior family history. The age of onset in these two (22 years and 48 years) suggests an autosomal recessive phenocopy, an as yet unidentified sporadic mutation or an acquired condition. These data urge caution in providing predictive testing to at-risk individuals of Azorean descent as not all cases of SCA are MJD/SCA3.

25

Non-syndromic cleft lip with or without cleft palate (NSCL/P) is a common and complex birth defect characterized by heterogeneous etiology. In 1995, Stein et al. reported that BCL3, or a nearby gene, was implicated in the origin of NSCL/P in some multiplex families. We ascertained 45 US and 22 Mexican multiplex NSCL/P families through a multi-center collaborative project. We performed the transmission disequilibrium test (TDT) and linkage analyses (using both parametric and non-parametric techniques) on these families. Markers on chromosome 19q were examined, including D19S178, ApoC2[acl], ApoC2[007-008], and BCL3. For the likelihood-based linkage analyses, NSCL/P was assumed to be an autosomal dominant disorder, with different levels of incomplete penetrance (i.e. 80%, 60% or 30%) and gene frequency = 0.001.

The TDT yielded a significant association between NSCL/P and allele 3 of the BCL3 marker (US data, 37 trios, p=0.002; Mexican data, 21 trios, p=0.049; combined data, p<0.001) and between NSCL/P and allele 13 of the D19S178 marker in the Mexican data (21 trios, p<0.001). Only 30 US and 11 Mexican families were informative for these markers, and so they failed to yield evidence for linkage using likelihood methods, even after adjusting for heterogeneity.

Although these findings, obtained from two ethnically different populations, partially confirm Stein's report, more research needs to be conducted to better define the role of markers on 19q in the etiology of NSCL/P.

24

GENETIC HETEROGENEITY OF MUTATIONS AMONG PATIENTS OF PORTUGUESE ANCESTRY MANIFESTING SPINOCEREBELLAR ATAXIA.

MacLeod P,[1] Silveira,I[4] Langlois S,[1]Lima M,[2] RadvanyJ[5]Farrer L,[3] Maciel P,[4] Gaspar C,[4] Rouleau G,[4] Department of Medical Genetics, University of British Columbia [1], Department of Anthropology, Universidade Dos Acores [2],Department of Neurology, Boston University [3], Centre for Research in Neuroscience, McGill University [4] and the Department of Neurology,Hospital Israelita Albert Einstein, Sao Paulo Brazil.[5]

As part of an ongoing international effort to map genes which produce the various forms of spinocerebellar ataxia (SCA) we have ascertained a subset of 10 affected probands of Azorean Portuguese origin studied to confirm the diagnosis prior to predictive testing for other at-risk family members, we identified an expansion of unstable triplet repeats in 7. Six (6) of

The APOE gene is a known risk factor for Alzheimer's disease (AD), with the ε4 allele being associated with an earlier age of onset and possibly more rapid progression of the disease. The estimates of the proportion of AD cases carrying one or more copies of the ε4 allele vary dramatically between studies. We have examined APOE allele frequencies and genotypes in a population-based sample from Florida, chosen using census tract age and ethnicity data for residence in the Dade County area of Florida. Locations, mostly retirement communities, were selected based on these data. All residents within these communities were then asked to participate. Unaffected individuals in this sample were those which scored normally on screening instruments or were pronounced normal after a full memory disorder evaluation. Affected individuals fulfilled NINCDS criteria for probable AD. Unknown diagnoses were omitted from the study. The results from this sample were compared to those from a sample drawn from self-referring families with multiple cases of AD.

The unaffected individuals in the population based sample displayed an ε4 allele frequency of around 15% - consistent with previously published population frequencies. By contrast, the affected individuals in the population based group show the expected distortion of APOE alleles (increased frequencies of ε4), although this was not as striking as the distortion in the affected familial group. However, in the familial sample the unaffected controls (age-matched siblings) showed an increased ε4 frequency similar to that of the affected members in the population sample. Analysis of variance of "survival time" defined as age of onset or age at latest evaluation show that compared to family membership neither APOE alleles nor genotype is a strong predictor of onset age in these cases. The significant differences between results from the two samples point to the unwitting selection of familial factors other than the APOE locus in the family history positive samples. We are currently exploring the interaction between APOE genotype and family history.

In the population based sample, the expected trend in age dependent dose effect of ε4 is observed (0, 1 or 2 copies of ε4 giving rise to mean ages of onset of 76, 72 and 67 years respectively), and serial assessment of the cognitive status in 60 cases suggests that ε4 is associated with an initially greater rate of decline.

26

A genetic epidemiological study of Alzheimer's Disease in France.

CLERGET-DARPOUX F*,CAMPION D, MARTINEZ M, BICKEBÖLLER H, BABRON MC, HANNEQUIN D, FREBOURG T, BRICE A, AGID Y and the French Network on Alzheimer's Disease.
INSERM U155 Paris, France.

The aim of the French collaborative study on Alzheimer's Disease (AD) was to delineate the different etiological subgroups and to quantify their proportions.
A segregation analysis was first performed on 92 families of early onset probands (≤ 60 years) including first and second degree relatives. A clear cut off in the data was obtained, evidencing that a small subgroup (18%) of patients had inherited a dominant mutation conferring a risk close to 100% at age 60 for carriers.
This subgroup of patients (type I) was screened for the APP and Presinilin 1 and 2 genes. In 8 families, the causing mutation is on Presinilin I, in 3 on APP. In 4 families, no mutation was detected suggesting the

existence of another AD gene.
Besides, the role of APOE4 was studied in 417 non type I patients and compared to 1130 controls matched on 4 age classes (<60, 60-69, 70-79, 80) and sex. The relative risk associated to APOE4 (RRε4) was significantly greater than in each age class with a maximum in the intermediate classes. The RRε4 was not different between sex (whatever the age class). This result, which may seem contradictory with the higher incidence of AD after 60 in females than in males, is in fact well explained by a strong difference in the allele frequency of APOE4 in female and male controls (18% versus 13%). The study confirms a differential risk associated with APOE2 and APOE3 as well as a dose effect of APOE4.

27

FAMILIAL AND NON-FAMILIAL FACTORS IN THE PREDICTION OF DISRUPTIVE BEHAVIORS IN BOYS AT RISK FOR SUBSTANCE ABUSE

Partha P. Majumder (Indian Statistical Institute, Calcutta, India), Howard B. Moss (University of Pittsburgh, U.S.A.), Lenn Murrelle (University of Pittsburgh, U.S.A.)

This study aimed to identify (i) a core disruptive behavior disorder postulated to presage a substance use disorder, and (ii) the relative importance of parental behavioral phenotype, familial and non-familial factors in the determination of this core disorder in children. Data were collected from members of 118 SA+ families ascertained through substance dependent fathers. Comparable data were also collected from 186 SA- control families. The first principal component of symptom counts of behavior disorders which explained about 65% of the variance was strikingly similar across family members within and between SA+ and SA- families. This component was identified with the core disorder of disruptive behaviors that was best predicted by magnitudes of parental disruptive behaviors and by familial environmental factors in sons of SA+ parent(s). However, only familial environmental factors were significant predictors in the control (SA-) families. Parental behavioral phenotype was found to be the most important in the determination of the magnitude of a core behavior disorder in children in SA+ families, while familial environmental factors are most important in control families. These findings suggest that strong genotype-environment interactions may determine disruptive behavior disorders in children.

28

SEGREGATION ANALYSIS OF ENERGY INTAKE AND ENERGY EXPENDITURE IN FAMILIES FROM ANDHRA PRADESH, INDIA

M.F. Feitosa*, T. Rice , A. Nirmala Reddy and DC Rao

Dept. Genética, Instituto Oswaldo Cruz, Fiocruz, Brazil, Division of Biostatistics, Washington University, St. Louis, MO USA, Anthropometry & Human Genetics Unit, Indian Statistical Institute, India

Possible genetic component was assessed for energy intake and energy expenditure using segregation analyses. Data were collected on 1,691 individuals in 432 pedigrees from Andhra Pradesh, India, and adjusted for age effects within sex.

For energy expenditure, there is a significant multifactorial component with generational differences, and a significant major effect not consistent with Mendelian transmission. Corrections for skewness appear to have little effect on the results, as they did not change the results obtained for untransformed data. For energy intake, we were unable to get clean final solutions (due to convergence problems). However, for energy intake adjusted for weight, the segregation results do not support a major gene effect, either before or after transformation to remove skewness. The most parsimonious model suggests a significant multifactorial component with generational differences, and a significant non-Mendelian major effect. Consequently, either for energy expenditure or for energy intake there is little evidence to support a major gene effect. However, numerous environmental, behavioral, and socioeconomic factors could have obscured any genetic effect, if present. (Supported by CNPq and NIH grant GM28719).

29

Apolipoprotein E genotype ε4/ε2 in the STANISLAS Cohort Study. Dominance of the ε2 allele ?

S. Visvikis*, K. Bohnet, A. Regis-Bailly, M. Vincent-Viry, R. Gueguen, G. Siest

Centre de Médecine Préventive, URA-CNRS 597, 2, Avenue du Doyen Jacques Parisot, F-54501 VANDOEUVRE-LES-NANCY Cédex

Apolipoprotein (apo) E has been discussed as a potential marker for cardiovascular risk, but information about lipid traits in healthy individuals having one of the rare apoE genotypes is scarce. Our work was designed to answer the following questions: 1. Are the allelic effects of ε2 and ε4 on the lipid traits total, LDL- and HDL-cholesterol, triglycerides, apo AI, apo B and apo E additive or dominant ? 2. If there is additivity, do the allelic effects of ε2 and ε4 on these lipid traits have the same magnitude ? 3. Are the allelic effects neutralised in ε4/ε2 individuals who are under the influence of both rare alleles ? Two linear regression models were used for statistical analysis, a genotypic model, that considers all six apoE genotypes separately and an allelic model, that is based on the hypothesis of additivity of allelic effects. We observed that the allelic effects of ε2 on apo B and apoE serum levels are dominant over that of ε4. Allelic models are thus not adequate to study the influence of apoE polymorphism on apoB and apoE serum levels. Allelic effects are additive for total cholesterol, LDL-C, HDL-C and apoAI, with ε2 having a greater impact than ε4. Genotype ε2/ε2 causes the most important deviation of serum lipid and apolipoprotein levels from the ε3/ε3 reference. In ε4/ε2 individuals, allelic effects of ε2 and ε4 are neutralised for all variables studied except for apoE. ApoE is the only variable for which serum level differs significantly between ε4/ε2 and ε3/ε3 individuals ($p < 0.001$). The clinical relevance of the elevated apoE serum levels in carriers of the ε2 allele remains to be determined.

30

Polymorphisms at the *NAT2 Locus Associated with Increased Cancer Risk in First-Degree Relatives of Lung Cancer Patients.** P Yang*, M Romkes, AG Schwartz, L Rittmeyer, A Adedoyin, R Landreneau, K Mauro, R Branch. University of Pittsburgh Medical Center (UPMC), Pittsburgh, PA 15261

The purpose of this study was to test the feasibility of using genetic polymorphisms of a specific metabolic enzyme, N-acetyltransferase, as a biomarker to identify high risk cancer families. The mutant *NAT2** genotype (*mut*) has been associated with lung cancer risk. Fifty six lung cancer patients were recruited in a study at UPMC. The mean age at diagnosis of the patients was 62 years; 92% of the male vs. 78% of the female patients were cigarette smokers. Through in-person interviews we have obtained

information including history of cancer on all first-degree relatives of each patient. The mean age of these relatives was 51 years. In a total of 443 relatives, 11 (2.5%) had lung cancer and 60 (14%) had any cancer, 34 and 42 times higher, respectively, than the expected numbers based on NCI SEER data. The overall trend was that the younger the age at diagnosis, the higher the ratio of the observed to the expected number of cancers. The association between patients' genotype at *NAT2** locus and cancer occurrence in relatives was examined. For the *mut* and *wt* (*NAT2*4/*4*, homozygous wild type) group, the percentage of patients having at least one first-degree relative with cancer was 76% and 38%, respectively (OR=5.24, p=0.05). These results indicate that patient's *mut* genotype at the *NAT2** locus may be used in conjunction with other predicting factors to identify high risk cancer families. Larger sample sizes and the addition of a control group are needed to confirm these preliminary results.

31

AGE OF ONSET IN FAP: VARIATION BETWEEN TWO HOT-SPOT MUTATIONS OF THE APC GENE

S. Presciuttini*, V. Gismondi, L. Bertario, P. Sala, C. Rossetti, C. Mareni, M. Ponz de Leon, C. Marchese, P. Radice and L. Varesco

**Italian Polyposis Registry, Milan, and Univ. of Pisa, Italy*

Familial adenomatous polyposis (FAP), a Mendelian condition highly predisposing to colorectal cancer is caused, in the vast majority of cases, by chain-terminating mutations of the APC gene. Two different hot-spot mutation sites have been identified, at codons 1061 and 1309 respectively, which together account for about 25% of the total germ-line mutational load. In previous studies we have shown that mutation 1061 gives rise to a "late FAP" phenotype, whereas mutation 1309 is associated with the "early FAP" phenotype. Here we extend the analysis to 11 families with mutation 1061 (54 verified FAP patients and 7 assumed cases) and 13 families with mutation 1309 (63 patients and 34 assumed cases). This work is carried out within the scope of AIRC's special project *Hereditary colorectal tumors*.

Age of onset was subdivided into two variables: age at FAP diagnosis and age at death from colorectal cancer (without any further stratification of cases). The following table shows the results of the comparison between mutations (means in years).

	1061		1309		
	N	Mean±sd	N	Mean±sd	P
Age at diagnosis	54	35.7±14.8	60	25.3±13.4	<.001
Age at death	19	48.0±13.2	39	40.4±11.2	<.05

This result definitely proves that variation of cancer susceptibility in FAP is in part mutation-specific (mutation 1309 is, on average, more cancer prone than mutation 1061). However, there was a huge variation among families within mutation, and this phenomenon implies other unknown pedigree factors affecting the phenotypic expression of the mutant gene.

32

INCIDENCE OF HPVs IN THE GASTROINTESTINAL TUMORS.

Sobti, R.C*., K.Parashar., R. Kalra and N. Capalash, Department of Bio-technology, Panjab University, Chandigarh. 160014 (India).

It has been documented that a number of factors including the viruses lead to alterations of genetic code, thus, initiating malignant transformation. One such transforming virus is the human papilloma virus (HPV). Each of the nearly 60 different type of HPVs reported has aproximately 7,900 bp long genome with around 8-10 open reading frames (ORFs) on the same DNA strand; ORFs E6 and E7 being associated with the transforming property. Though HPVs were amongst the first to be associated with human neoplasia, their exact role in establishment of cancer is still not fully known. Studies in our laboratory on cervical cancers under low stringency conditions have revealed the presence of HPV-16 DNA in 88 and 80% of cases with invasive carcinoma and tissues with abnormal cytology respectively. The association was 40 and 20% in these cases under high stringency conditions. The present study is being conducted to find out association, if any, between HPV types 11, 16, 18 and 41 with cancers of gastrointestinal origin. Preliminary studies on oesophageal carcinomas have doubted the association with HPV 11 and 41, while HPV 16 and 18 gave weakly positive signals in 10% cases each respectively under low stringency conditions. To substantiate these observations the association has to be further studied under high stringency conditions and these experiments are in progress. Besides status of p53 protein in cancerous cells is also being evaluated. The incidence of chromosomal aberration, SCEs and alterations in the pattern of heterochromatin in the lymphocytes of patients with gastrointestinal cancer are being worked out.

33

EVALUATION OF PRENATAL DIAGNOSIS OF CONGENITAL HEART DISEASE.

C. STOLL, Y. ALEMBIK, B. DOTT, B. DE GEETER, M.P. MEYER, A. PENNERATH, M.P. PETER.
Centre Hospitalo-Universitaire, Strasbourg, France.

Prenatal diagnosis performed by fetal ultrasound scan is now a routine part of antenatal care in many countries. That an increasing number of fetal anomalies may be detected on prenatal ultrasound is beyond doubt. What is possible is not, however, always practical, especially when congenital heart diseases (CHDs) are concerned and when whole antenatal populations are screened rather than high-risk groups. Thanks to our registry of congenital anomalies, a retrospective study was undertaken to evaluate the prenatal detection of CHDs by ultrasound scan in 92062 consecutive pregnancies of known outcome from 1990 to 1993. Only 116 out of 781 malformed fetuses with CHDs without chromosomal anomalies were detected (11.0 per cent). The sensitivity of detection varied from around 45 per cent for malformations such as hypoplastic left heart and transposition of great arteries to around 5 per cent for ventricular septal defect and atrial septal defect. Thirty out of 99 cases (29.9%) of chromosomal anomalies with CHD were detected. The effectiveness of the detection of some forms of major congenital heart disease has increased dramatically since our last evaluation in 1988 (Prenatal Diagnosis 1993, 13:453-461) by including routine examination of the four-chamber view and of the inflow and outflow tracts of the fetal heart. Our results stress the need to obtain a definite clear four-chamber view, to perform scans at ≥18 weeks of gestation, and to train sonographers in order to improve the prenatal detection of CHDs.

34

Epidemiology of twinning rate in Japan.

Y. Imaizumi.
Institute of Population Problems, Ministry of Health and Welfare, Tokyo, Japan

Twinning rates in all of Japan for 1951-1968 and 1974-1994 were analyzed using data from vital statistics of Japan. Mean twinning rates (per 1,000 births) by zygosity were 4.12 for MZ and 2.26 for DZ twins during the period 1955-1967. The corresponding rates were 3.90 and 1.86 in 1974, respectively.

Overall twinning rate was 6.43 in 1951 and remained nearly constant until 1968, then decreased to 5.79 in 1974 and gradually increased to 6.72 in 1988, and rapidly increased to 8.32 in 1994. In Japan, ovulation-inducing hormones have been used since 1966, and the first *in vitro* fertilized baby was born in 1983. According to the survey of ovulation-inducing hormones through questionnaires in 1993, the twinning rate following treatment with HMG therapy was 21% (Aono, 1994). Similarly, from a survey of *in vitro* fertilization through questionnaires, the rates of multiple births were 19% in 1988, 23% in 1989, 28% in 1990, 22% in 1991, and 24% in 1992. The higher twinning rates since 1988 have been attributed to the higher proportion of mothers treated with ovulation-inducing hormones and partially attributed to *in-vitro* fertilization.

35

Do Perinatal Factors influence Later Phenotypes and their Heritability Estimates ?

R.Vlietinck*,C.Derom,R.Loos,G.Beunen,E.Thiery,M.Danckaerts,R.Derom
Center for Human Genetics, Catholic University of Leuven, Belgium

We tested the influence of prenatal factors on postnatal phenotypes.
Since 1964 the East Flanders Prospective Twin Study has followed prospectively all multiple births in that area. We determined their zygosity by bloodgroups and DNA markers and, for the monozygotics (MZ), the moment of splitting by examination of the fetal membranes: dichorionics split before and monochorionics split after the fourth day after conception, while monoamniotics split even after the eighth day. On all twins the placenta was thoroughly examined, DNA collected and perinatal examination performed.
The prenatal factors included cord insertion, fetal membranes, placental vascular anastomoses, weight and surface. The postnatal ascertainable factors included maternal age, parity, gestational age, birthweight, fetal presentation, mode of delivery, birthrank and age at examination.

We examined the influence of these prenatal factors on phenotypes at later age. Anthropometric, motor, cardiovascular variables were measured in adolescents; craniodentofacial and neurological maturation was assessed in preadolescents; intelligence, cognitive performance, school performance, cerebral and somatic lateralisation were measured in schoolchildren; child behavior and family environment were determined in 1,200 twin pairs.

The influence of the prenatal factors was evaluated by comparing the means per group, the intrapair differences, the variances and covariances and the influence on the heritability.

Placental vascular anastomoses were found in none of the dichorionic and in most of the monochorionic twins. Monochorionics had significantly more marginal insertion of the cord while in dichorionics the insertion was more central. The mean performance IQ of the monochorionic twins was significantly 7 points lower than that of dichorionics. The means of nearly half of the 38 biometric variables differed significantly according to the moment of splitting. Only in late splitting MZ twins the within-pair difference differed significantly from null in 27 variables. More than one third of the variance was explained by the perinatal factors: 23% by prenatal factors, mainly due to cord insertion and 13% by postnatal factors, mainly by maternal age and gestational age.

The results indicate that somatic and craniofacial growth, IQ, school and learning problems and behavior were significantly influenced by perinatal factors, up to 18 years of age. Without accounting for these prenatal factors, the heritability estimates were often severely biased.

Funding: Belgian Scientific Research Fund 3.0125.94 and M.M.Delacroix Fund.

36

PREVENTION OF BIRTH DEFECTS THROUGH PRIMARY GENETICS CARE. Moreno Fuenmayor H, Valera V, Socorro-Candanoza L, Martinez X, Herrera M, Godoy P, Hernández-Pirela D, Rodriguez Z, Concho E. Servicio de Medicina Genética Perinatal, Hospital Chiquinquirá, Maracaibo, Venezuela. email: hmoreno@luz.ve

The experience in setting up a Birth Defects Prevention Program through primary genetics care is presented. One medical geneticist with training in ultrasound and invasive prenatal diagnosis is helped by two laboratory technicians, two nurses, and two obstetricians-ecosonographists. Two pediatricians and one obstetrician carry the birth defect registry and the ascertainment of mothers with high risk pregnancies for fetal defects (HRPFD). Patients attended are those detected because of a positive family or obstetric history, women with increased risks because of medical conditions or ultrasound findings, puerperal women with abnormal reproductive outcome, or self-referred patients. 576 patients have been attended adding up to 15 consultations a month, and a potential population of 60 new patients a month is expected when the program is extended to other child-maternal facilities. Appropriate ascertainment seems related to decrement in the number of consultations due to increased maternal age. A self-ranking questionnaire appears to be an efficient but still not ideal ascertainment strategy but early ultrasound evaluation is a powerful approach. The program is also one of genetics epidemiology surveillance since basic epidemiological figures for the HRPFD are being developed and clustering of the Pena-Shokeir syndrome has been observed.

37

PERSISTENCE OF INDUCED BEHAVIORAL MUTATIONS THROUGH SUCCESSIVE GENERATIONS.

J.H. Schröder* and R. Siegmund**
*) GSF-Research Center Munich, D-85758 Oberschleissheim, Germany
**) Institute of Anthropology, Humboldt University, D-10098 Berlin, Germany

Extrapolations from mutation experiments of vertebrates to humans indicate a decrease in learning ability and an increase in intraspecific aggressiveness. Thus, offspring from X-irradiated rat populations decreased in maze learning ability on a percentile basis in mean I.Q. from 100 to 94.65. Environmentally induced mutations affecting innate behavior often persist through many successive generations thus impairing the gene-pool of the populations in question. Male offspring of X-irradiated mice, in particular translocation heterozygotes of two defined lines, were more often winners in standardized agonistic encounters and also more often non-learners in lever boxes than their chromosomally normal controls. The comparison with humans also seems to be justified because not only rodents but also different species of teleostean fish respond to induced mutations in a similar manner. In long-term population-genetic experiments, the progeny of ancestrally irradiated guppies revealed an impairment of fitness characters (increase of intermale aggressiveness combined with a reduction of sexual activity) which led to a reduction of reproductive success.

38

Idiopathic epilepsy with generalized tonic clonic seizures (IETCS) in Antioquia, Colombia: Is the joint Amerindian and Negroid racial admixture the cause of its high prevalence?

I.Jiménez[1],O.Mora[1],M.E.Jiménez[1],L.Zuluaga[1],R.Isaza[1], J.L.Sánchez[1], S.Uribe[1], C.Valenzuela[2], R.Blanco[2], M.Arcos[2]*.[1]Antioquian League Against Epilepsy, Medellín,Colombia. [2]Dept. Cell Biol.& Gen.Medical, School,U.Chile,Santiago,Chile.

Most Colombian populations stem from the admixture of Caucasians, Amerindians and Negroids. These two latter ethnical groups show a significantly higher prevalence of epilepsy than the former one. We tested the hypothesis that the high prevalence of IETCS found in the Antioquian population, is due to their possible joint Negroid and Amerindian ethnic components. We had previously demonstrated that inheritance is the principal factor in the development of epilepsy in this community. Analyses of racial admixture, heterogeneity between populations, genetic distance, and phyletic relationships were performed among Epileptic and Non Epileptic Antioquians, and with Caucasians, Spaniards, Basques, Jews, Chileans, Negroids, Amerindians and Mongoloids. RH, MNS, ABO and FY blood systems were used as genetic markers. In the Epileptics the estimated Negroid and Amerindian rates of admixture were low (3% and 14% respectively); However, their estimated proportion of Amerindian admixture was significatively higher than the estimated for the Non Epileptic population. The latter finding was consistent with the analysis of heterogeneity between populations that discriminated Epileptic population from Non Epileptic Antioquian population ($p<0.05$). Epileptic and Non Epileptic Paisas clustered in topology with Caucasians, very close to Spaniards and Basques and highly distant from Negroids and Amerindians. Therefore, the origin of the high prevalence of IETCS in the Antioquian (Paisa) population cannot be explained by the hypothetical joint Negroid and Amerindian ethnical admixture. But, using additional genetic markers and other methods of racial estimation of admixture it is necessary to corroborate if the Amerindian admixture component is significatively higher in the Epileptic population than in the Non Epileptic Paisa population.

M.Arcos has financial support from COLCIENCIAS

39

Effect of family selection on the detection of a candidate gene in a complex quantitative trait.

N. Chatelain[1], M. Rosenberg-Bourgin[2], F. Demenais[1]*
[1]INSERM U. 358, Paris, [2]INSERM U. 155, Paris, France.

Combined segregation-linkage (CSL) analysis can be a powerful method to detect a candidate gene contributing to the variation of a complex quantitative trait. Assessing that the genetic variant investigated is the causal variant has important implication for further studies of gene function. In many instances, family samples are selected through subjects with extreme values of the trait for gene mapping. Since the mode of ascertainment is often ill-defined, simulations were conducted to investigate the robustness of the CSL strategy to a selection bias. Quantitative and marker data were simulated on 100 replicates of eight-member nuclear family samples. The generated models included a candidate gene, dominant or recessive, responsible for 13-16% of the total variance, in complete or incomplete linkage disequilibrium (LD) with a tightly linked marker (no recombination), plus residual familial correlations (ρ's). Sample sizes were chosen to yield 80% power to detect the true model in random samples. Families were selected through at least two sibs with values within the top 5% of the distribution. Analyses were conducted with the regressive models without correcting for ascertainment. When data are generated under a complete LD model, the gene effect is detected in all replicates and residual correlations in 60% to 95% of them. While absence of LD is always rejected, the proportion of replicates fitting the generated complete LD model varies according to the mode of inheritance and whether the ρ's are ignored or accounted for in the analysis. This proportion increases from 40% (ρ's=0) to 75% (ρ's$\neq$0) when the gene is dominant, and from 65% (ρ's=0) to 85% (ρ's$\neq$0) when the gene is recessive. When the generated model includes an incomplete LD, the gene effect is detected in 80% of replicates with the ρ's being significant in 30% to 70% of them, depending on the parameter values. Absence of LD is rejected in at least 75% of replicates while the rejection rate of complete LD is about 50% for a dominant gene and 80% for a recessive gene, with or without the ρ's. The CSL strategy appears robust to a selection bias, especially when there is complete LD. Ascertainment-assumption free methods need also to be explored.

40

Algorithmic Improvements to MAPMAKER/SIBS

Ramana M. Idury
Sequana Therapeutics, Inc.
11099 N. Torrey Pines Rd, STE 160
La Jolla, CA 92037, U.S.A.
ramana@sequana.com

MAPMAKER/SIBS [2] is an interactive program for performing sib-pair analysis, which is one of the most important statistical techniques for genetic analysis of complex traits. SIBS uses the Hidden Markov Model to compute the distribution of the inheritance vector for each nuclear family. The heart of the program is an algorithm for computing the product of two specific matrices of size $2^n \times 2^n$, where n is twice the number of children in a nuclear family, with $(n+1)2^n$ multiplications and $\frac{(n+1)(n+2)}{4}2^n$ additions using the properties of linear algebra [1].

We present an algorithm for the same problem using the divide-and-conquer paradigm from computer science. Our algorithm uses the same number of multiplications as the above method but requires only $n2^n$ additions. Moreover, our method uses only 2^{n+1} words of auxiliary space whereas the other method needs $(n+1)2^n$ words.

Another aspect of our algorithm is that one can incorporate sex-specific recombination fractions without increasing the time or space complexity. On the other hand, the method in [1] slows down the algorithm by a factor of n to achieve this. In fact, it is possible to vary recombination fraction from one individual to another so other variations such as age-specific recombination are possible at no additional cost. We have tested our algorithm on only SIBS but we expect our improvements to hold in other MAPMAKER based programs such as HOMOZ as well since they are based on the Hidden Markov Model.

References

[1] L. Kruglyak *et al. Am. J. Hum. Genet.*, 56:519–527, 1995.

[2] L. Kruglyak and E. S. Lander. *Am. J. Hum. Genet.*, 57:439–454, 1995.

41

Population Based Family Study Design: An Interdisciplinary Research Framework for Genetic Epid.

Zhao LP, Hsu L, Davidov O, Potter J, Prentice RL

Dissecting the genetic and environmental etiology of complex traits such as cancer and coronary heart disease requires an interdisciplinary research strategy. Traditionally, genetic studies involve selected families and investigate genetic etiology of diseases while epidemiologic studies sample population based cases and controls and assess environmental risk factors. Integrating population based and family study designs, we propose a class of population based family study designs as an interdisciplinary research framework for genetic epidemiology. One important design within this framework is the case family control study,where one samples population based cases and controls at the first stage, a prespecified set of case-relatives at the second stage, and collects their environmental risk factors and disease status. At the third stage, blood and tissue samples are collected from families that have excessive numbers of cases, and are used to generate genetic markers. Data from all three stages are then combined to simultaneously assess environmental risk factors, major disease genes in the general population and localize the disease genes. Methods for combined linkage segregation and aggregation analysis are currently being developed We believe that population based family study designs will retain the efficiency of genetic studies for localizing genes while being population based. Therefore, they will allow the assessment of independent and interactive contribution of genetic and environmental factors to complex traits.

42

QTL analysis of hypertension in rats: sensitivity to statistical assumptions

M. A. Spence*, P. Flodman, M. G. McLaughlin, M. Printz, H. J. Jacob, and D. Panestil
Dept. of Pediatrics, UC Irvine, CA, USA; Massachusetts General Hospital, Boston and Medical College of Wisconsin, USA; and Depts. of Pharmacology and Medicine, UC San Diego, CA, USA

From an SHR x WKY rat cross, F1 (n=85) and F2 (n=399) individuals were measured for 31 variables including diastolic/systolic blood pressure, plasma and urine metabolites. 210 markers were typed in F2 males.

For a "quick & dirty" preliminary genomic screen raw data, ratios, and linear combinations were analyzed. McLaughlin et al. (1995) detected 37 autosomal linkages (zmax > 3.0) choosing to report 3 regions with significant influence on blood pressure (bp), encompassing 8 variables.

Our statistical evaluation of 45 traits (raw data and a limited number of physiologically relevant functions of the data) included only 2 of the 8 bp variables. Most variables deviated from normality and were transformed. Gender, age and weight effects were accounted for by regression analyses. The correlations among traits and the presence of parental strain differences were assessed. BCROSS (SAGE) was used to fit and evaluate genetic models.

For the 45 variables in our residual analyses, the genomic screen detected 18 linkages. We confirmed 8 linkages, including the bp region, found 7 additional linkages, and for 10 linkages zmax dropped below 3.0. This variety of changes is expected when comparing the two approaches.

43

Genetic Epidemioloy of Highland Isolates and Migrants to the Lowlands

K. B. BULAYEVA

Vavilov Institute of the General Genetics (Moscow)

Genetic epidemiology has been studied in "split" isolates, part of which was forcibly moved from traditional highlands into radically new lowland areas in Daghestan (Northern Caucasus, Russia). To compare with these "split" isolates, populations of the aboriginal lowlanders have been studied. It has been found that during adaptation to new environments, a statistically significant differentiation between the highlanders and the migrants from these "split" isolates has occurred, involving both marriage and genetic structure and neurophysiological sensitivity. The differences were determined by a selective death of about 30% of the migrants with certain genotypes. In the survivors the morbidity and the prenatal mortality had sharply increased, while the life span significantly diminished. The data obtained shows that among the migrants the congenital and hereditary defects drasticly increased from 5% in 1978 to 15,5% in 1995. Fertility and prenatal mortality revealed, respectively, positive and negative correlation with inbreeding. Crow's index of potential selection and its components due to differential mortality and differential fertility were estimated for the highlanders, migrants and lowlanders.

44

GENETIC HOMOZYGOSITY AND NORMAL vs. PATHOLOGICAL HUMAN VARIATION

Dragoslav M. Marinković

Faculty of Biology, University of Belgrade, 11000 Belgrade, Yugoslavia

The degree of individual genetic homozygosity has been estimated according to the presence of 20-30 qualitative homo-recessive characters ("HRC-TEST"), in samples of healthy and diseased individuals. In the majority of such studies, including more than 3000 individuals, an average increase of HRCs was found among patients with urogenital and neuropsychiatric diseases, myotonic distrophy, diabetes mellitus, chronic lymphocytic leukemia, congenital hip dislocation, Ca uteri. A similar or lower mean homozygosity was found in patients with allergic bronchial asthma, tuberculosis, lung and Ca mammae, with a significant difference in the frequency distribution of the studied morphophysiological characters.

Further applications in HRC-testing may have practical use in genetic counselling, since homozygosity seems to be correlated with an individual's succeptibility to different diseases.

45

Comparative Study of Differential Fitness in the Highland, Migrant and Lowland Populations of the Caucasian Aboriginal Ethnics (Northern Caucasus, Russia)

Dr. Tatijana Pavlova

Vavilov Institute of General Genetics (Moscow)

The Crow's index of total selection has been calculated in the populations of highlanders, migrants from the highlands to the lowlands and lowlanders. It was based on the assumption, that if all the variations in mortality and fertility are due solely to differences in the value of the associated genotypes, then its value is a function of the effectivences of natural selection. Among the highlanders the mortality components may be seen to be 0.049 for prenatal deaths and 0.531 for postnatal deaths; the fertility component is equal to 0.627, the total index is - 1.207; in the group of migrants these values are 0.124, 0.227, 0.435 and

0.786 respectively; in the lowlanders - 0.106, 0.417, 0.178 and 0.701 respectively. The proportion of the total index due embryonic mortality in the group of highlanders is 4.1%, migrants - 15.8%, and lowlanders - 15.5% ; due to postnatal mortality is - 44%, 29% and 59% respectively; due to differential fertility - 52%, 55% and 25% respectively.

46

GENETIC AND BIOMEDICAL CORRELATIONS AMONG FOUR SOUTHERN CHILEAN POPULATIONS.

Moreno RS*, Llop E, Harb Z, Arcos M.
Medicine Faculty. University of Chile. PO 70061 Santiago-7. Santiago-CHILE, and Hospital Exequiel Gonzalez Cortes.

The populations studied are located on 41°45` S (CAR), 41°50^S (QUE), 42°40` S (DET), and 43°24`(LAI) around 73° W at Chiloé island. They are studied for eight biophysiological variables, namely systolic and dyastolic arterial pressure (SAP & DAP), cardiac and respiratory frequency (CF & RF), weight (W), size (S), body mass index (BMI), and glicemy. Two blood groups are identificated, ABO and RH, in order to discriminate Amerindian and Spaniard admixture components. The traits are described as quantitative variables, and their correlational studies with geographic and genetical distances are realizated, according to a genetic epidemiological view.

The gene frequencies and geographical distances are correlated with caucasian admixture that decrease in north-south cline. The allele O frequency (AOF) is directly correlated to latitudinal distances (r : 0.56). However, they are inversely correlated to allele Rh negative frequency (r : -0.61 and r : -0.95, respectively). The variables PAS, RF, and G correlated significatively with AOF (r : -0.64, r :-0.55, and r : -0.72, respectively) by whole populations sampled. In female samples, they are r : -0.75, r : -0.83, and r : -0.73, respectively, in addition, BMI has r : -0.78. On the other hand, male samples only BMI (r : 0.56) and G (r : -0.79) are correlated to AOF.

The above findings support the thesis that these traits are influenced by sex. Moreover, the high correlations between female samples and AOF, greater than the correlations the male samples and AOF, suggest that the females could have a major Amerindian componen than males sampled.

FONDECYT grants 1930884 and 1950594.