# Genetic Epidemiology of Oral Clefts

Rolv T. Lie

Outline

1. Registry-based data parent-offspring recurrence
    – Birth defects have specific familial risks
    – Recurrence risks are high for oral clefts (30-fold)
    – Multiple genes are assumed to be important

2. Hunt for genes for oral clefts
    – GWAS have found 50+ associated SNPs
    – Estimates of associations

3. How much of the recurrence risk of oral clefts is explained by the effect of identified genetic variants?
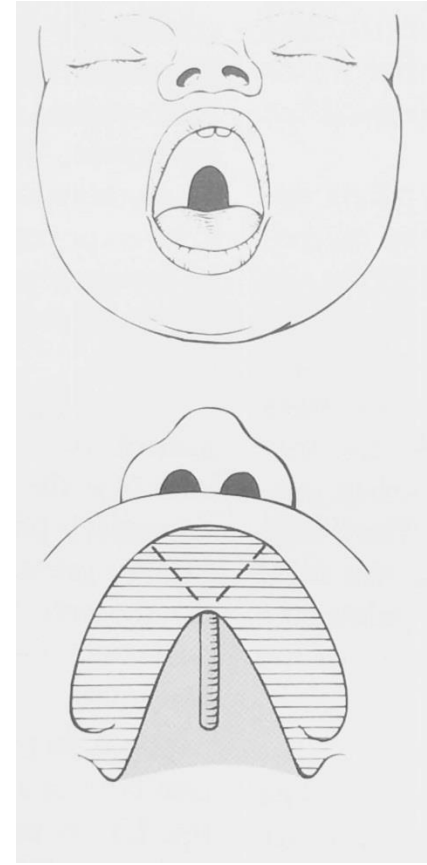
Oral clefts

Prolabium
Premaxilla
Nasal septum
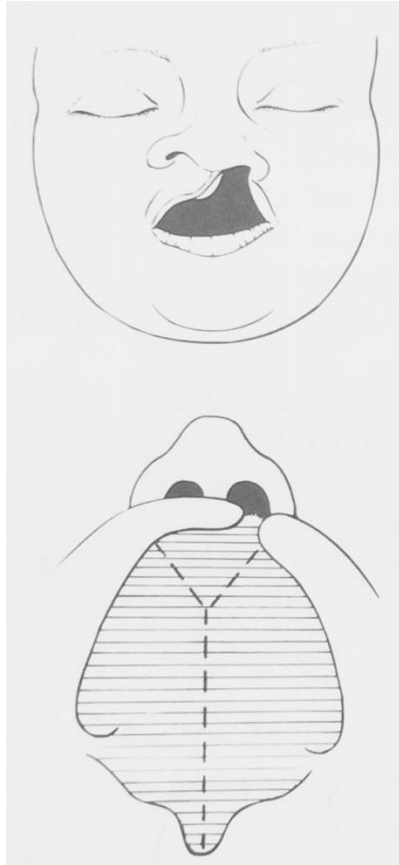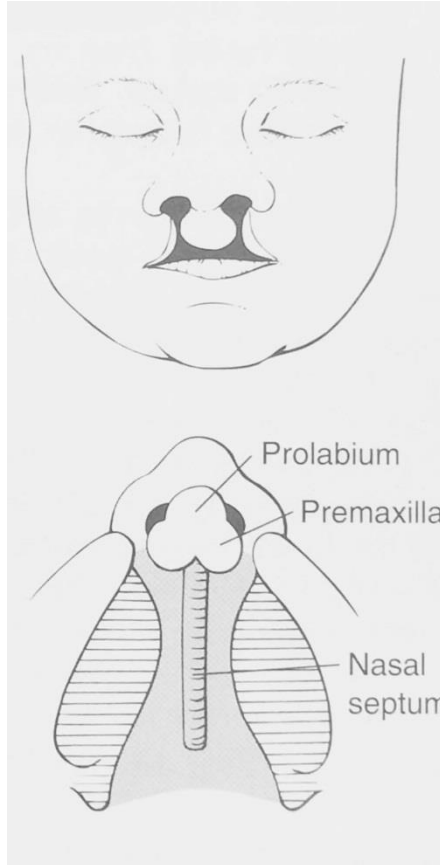
cleft lip only	cleft lip with cleft palate	cleft palate only
CLO	CLP	CPO

O'Rahilly & Müller, 1992

1. Medical Birth Registry (MBR) of Norway

- All births in Norway since 1967 (~3 million)
- Medical information on delivery, child and mother
- Known ID-number of child and both parents in MBR
- Link a person's birth-record with birth-records of parents
- Ascertainment vary for birth defects (90% for cleft lip)

1. Recurence risks, mother-child

From: A population-based study of survival and childbearing among female subjects with birth defects and the risk of recurrence in their children.
N Engl J Med. 1999 Apr 8;340(14):1057-62. *Skjærven R, Wilcox AJ, Lie RT:*

(Study of a total of 187 544 children of mothers born in 1967-82)

TABLE 3. RISK OF SIMILAR AND DISSIMILAR BIRTH DEFECTS IN CHILDREN ACCORDING TO CATEGORY OF BIRTH DEFECT IN THE MOTHER.

DEFECT IN MOTHER*	NO. OF MOTHERS	NO. OF CHILDREN AT RISK	TOTAL NO. OF DEFECTS	SIMILAR DEFECT			DISSIMILAR DEFECT		
				NO. OF DEFECTS OBSERVED	NO. OF DEFECTS EXPECTED	ODDS RATIO (95% CI)†	NO. OF DEFECTS OBSERVED	NO. OF DEFECTS EXPECTED	ODDS RATIO (95% CI)†
Cardiac defect	54	75	3	0	0.13	—	3	1.7	1.8 (0.5–5.2)
Cleft palate	44	66	4	2	0.03	82 (13–290)	2	1.6	1.3 (0.2–4.4)
Cleft lip	104	149	11	7	0.20	38 (16–77)	4	3.4	1.2 (0.4–2.9)
Abdominal-wall defect	44	64	2	0	0.03	—	2	1.5	1.3 (0.2–4.6)
Clubfoot	453	666	32	15	2.80	5.5 (3.2–9.1)	17	13.1	1.3 (0.8–2.1)
Limb defect	159	236	7	2	0.37	5.6 (0.9–18.7)	5	5.3	0.9 (0.3–2.2)
All defects	1101	1613	62	26	3.94	6.8 (4.5–10.0)	36	34.6	1.0 (0.7–1.4)

*Defects are defined according to the *International Classification of Diseases, 8th Revision*. Not all categories of defects are listed, so numbers do not necessarily sum to totals shown.

†Mothers without the specific defect served as the reference group. CI denotes confidence interval.

(Study of a total of 110 427 children of fathers born in 1967-82)

Table 3. Risk of Similar and Dissimilar Birth Defects in Offspring, by Birth Defect Categories of Fathers

Categories*	Fathers, No.	Offspring at Risk, No.	Total Observed Defects in Offspring, No.	Similar Defects			Dissimilar Defects		
				Observed, No.	Expected, No.	Observed-Expected Ratio (95% CI)†	Observed, No.	Expected, No.	Observed-Expected Ratio (95% CI)†
Cleft lip	68	115	12	6	0.16	38 (14-93)	6	2.29	2.6 (1.0-6.1)
Abdominal wall defects	12	15	3	0	0.004	0 (0-1243)	3	0.32	9.5 (2.1-43)
Clubfoot	218	316	10	4	1.17	3.4 (0.9-9.0)	6	5.57	1.1 (0.4-2.4)
Limb defects	123	184	9	3	0.25	12 (2.5-37)	6	3.67	1.6 (0.6-3.7)
Genital defects	246	370	17	5	1.31	3.8 (1.2-9.2)	12	6.58	1.8 (0.9-3.3)
Anal defects	11	14	2	1	0.006	173 (4.3-1290)	1	0.29	3.4 (0.08-24)
Skin/hair/nail defects	29	42	2	1	0.029	35 (0.87-215)	1	0.87	1.2 (0.03-6.8)
Multiple defects	35	61	4	1	0.17	5.9 (0.2-35)	3	1.13	2.7 (0.5-8.4)
Total defects	850	1265	64	21	3.22	6.5 (4.0-10.4)	43	23.7	1.8 (1.3-2.5)

*Only categories with at least 2 occurrences in offspring are shown.
†Fathers without the specific defect were the reference. CI indicates confidence interval.

Parent-offspring recurrence of oral clefts

Familial risk of oral clefts by morphological type and severity: population based cohort study of first degree relatives. BMJ. 2008. *Sivertsen A et al.*

Table 1 | Relative risk* of recurrence across generations of cleft lip only (CLO), cleft lip and palate (CLP), and cleft palate only (CPO)

Index cases† (parent)	Recurrent cases† (offspring)				Relative risk* (95% confidence interval)		
Main categories of cleft	At risk	CLO	CLP	CPO	CLO	CLP	CPO
Cleft lip only (n=154)	293	3	7	2	19 (6.1 to 57.5)	29 (13.6 to 59.8)	9 (2.3 to 37.4)
Cleft lip and palate (n=182)	340	5	8	1	27 (11.2 to 64.9)	28 (14.0 to 56.7)	4 (0.6 to 28.5)
Cleft palate only (n=150)	288	0	1	11	–	4 (0.6 to 29.1)	54 (29.7 to 98.0)
No clefts (reference) (n=366 815)	702 210	388	601	516	1.0	1.0	1.0

*Relative risks (estimated as odds ratios in logistic regression models) are ratios of risk of recurrence and risk in reference group.
†Index cases are clinically verified cleft cases without non-cleft birth defects. Recurrent cases include all recorded cases among stillborn or live born babies, cases with or without other defects, and cases that were registered either in clinical data or in medical birth registry.

		Recurrence of cleft lip (with or without cleft palate)		
Familial relationship	At risk	Recurrences of cleft lip	Relative risk* (95% confidence interval)	P difference
Mother-offspring	295	11	27 (14.9 to 49.2)	0.97
Father-offspring	338	12	27 (15.0 to 47.2)	
Parent-offspring total	633	23	27 (17.7 to 40.3)	0.31
Subsequent full sibling	879	40	35 (25.5 to 48.4)	

Meta-analysis of published data from Denmark and Norway

Estimates of recurrence risk from parent to child of isolated oral clefts among Scandinavians:

	RR (95% CI)	_Absolute risks_	
		Recurrence	_Reference_
CLO - CLO:	**42** (31-56)	_2.3 %_	_0.055 %_
CLP - CLP:	**29** (22-37)	_2.5 %_	_0.086 %_
CPO -CPO:	**32** (24-42)	_2.3 %_	_0.073 %_

(A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark… J Med Genet. 2010 _Grosen D et al.)_

nature genetics

Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate

Elisabeth Mangold[1,26], Kerstin U Ludwig[1,2,26], Stefanie Birnbaum[1,26], Carlotta Baluardo[3], Melissa Ferrian[3], Stefan Herms[1,2], Heiko Reutter[1], Nilma Almeida de Assis[1], Taofik Al Chawa[1], M..........., Michael St........, Sandra Barth[1,2], Na............ Carola Lauster[6], G.......... Rudolf H Reich[9], A.......... Bettina Blaumeiser.......... Stefan Schreiber[14], ... Regine P Steegers-.......... Sven Cichon[1,2], Pet.......... Michael Knapp[4], M.......... Per Hoffmann[1,2] &...........

We conducted a g..........
nonsyndromic clef..........
in 401 affected ind..........
in an independent..........
new loci associate..........
$P = 1.07 \times 10^{-8}$, re..........
1.34–2.53) and 10..........

Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci

Kerstin U Ludwig[1,2,30], Elisabeth Mangold[1,30], Stefan Herms[1,2], Stefanie Nowak[1,3], Heiko Reutter[1,4], Anna Paul[5,6], Jessica Becker[1,2], Ruth Herberz[1,2], Taofik AlChawa[1,2], Entessar Nasser[1,2], Anne C Böhmer[1,2], Manuel Mattheisen[1,7,8], Margrieta A Alblas[1,2], Sandra Barth[1,2], Nadine Kluck[1,2], Carola Lauster[9], Bert Braumann[10],

locus. 20 SNPs with $P < 10^{-5}$ remained. Five chromosomal loci (8q12.3,

studies h..........
environm..........

Four g..........
been pub..........
thus add..........
(ref. 2)...........
we comb..........
Baltimo..........
Phenoty..........
After qu..........
Europea..........
Asian tr..........
1,318 co..........

A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near *MAFB* and *ABCA4*

Terri H Beaty[1], Jeffrey C Murray[2], Mary L Marazita[3], Ronald G Munger[4], Ingo Ruczinski[1], Jacqueline B Hetmanski[1], Kung Yee Liang[1], Tao Wu[1,5], Tanda Murray[1], M Daniele Fallin[1], Richard A Redett[6], Gerald Raymond[6], Holger Schwender[1], Sheng-Chih Jin[1], Margaret E Cooper[3], Martine Dunnwald[2], Maria A Mansilla[2], Elizabeth Leslie[2], Stephen Bullard[7], Andrew C Lidral[7], Lina M Moreno[7], Renato Menezes[3], Alexandre R Vieira[3], Aline Petrin[2], Allen J Wilcox[8], Rolv T Lie[9], Ethylin W Jabs[6,10], Yah Huei Wu-Chou[11], Philip K Chen[11], Hong Wang[6], Xiaoqian Ye[10,12], Shangzhi Huang[13], Vincent Yeow[14], Samuel S Chong[15], Sun Ha Jee[16], Bing Shi[17], Kaare Christensen[18], Mads Melbye[19], Kimberly F Doheny[20], Elizabeth W Pugh[20], Hua Ling[20], Eduardo E Castilla[21], Andrew E Czeizel[22], Lian Ma[23], L Leigh Field[24], Lawrence Brody[25], Faith Pangilinan[25], James L Mills[26], Anne M Molloy[27], Peadar N Kirke[28], John M Scott[27], Mauricio Arcos-Burgos[29] & Alan F Scott[6]

Case-parent trios were used in a genome-wide association study of cleft lip with and without cleft palate. SNPs near two genes not previously associated with cleft lip with and without cleft palate (*MAFB*, most significant SNP rs13041247, with odds ratio (OR) per

Although CL/P can occur in many Mendelian malformation syndromes, the isolated, non-syndromic form constitutes 70% of all cases[2]. Evidence for genetic control of CL/P is compelling: recurrence risks are 20–30 times greater than population prevalences[3,4], and both

Study of individuals of European descent:

A Population-Based Study of Effects of Genetic Loci on Orofacial Clefts

L.M. Moreno Uribe[1], T. Fomina[2], R.G. Munger[3], P.A. Romitti[4], M.M. Jenkins[5], H.K. Gjessing[2,6], M. Gjerdevik[2,6], K. Christensen[7], A.J. Wilcox[8], J.C. Murray[9], R.T. Lie[2,6*], and G.L. Wehby[10*]
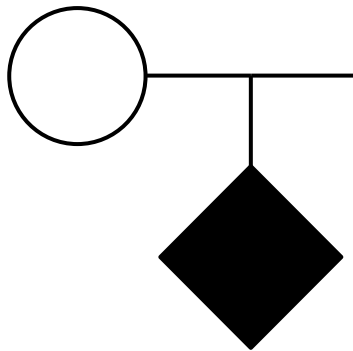
2011 case-dyads
(mother and child)
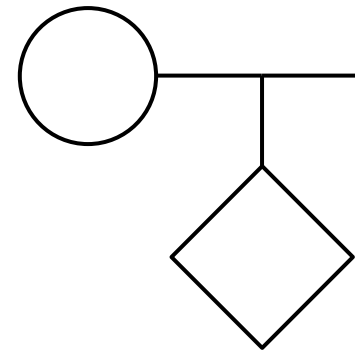
4521 control-dyads
(mother and child)

Haplin -
Estimation of effects of child's alleles, parent of origin effects and maternal alleles

Frequency of 14 SNPs that had effect for one category of clefts in our analyses

Gene	SNP	MAF
PAX7	rs742071	0.40
ABCA4_AR	rs560426	0.46
ABCA4_AR	rs215184	0.42
IRF6	rs642961	0.20
THADA	rs759026	0.23
COL8A	rs793464	0.41
8q21.3	rs125433	0.34
8q24	rs987525	0.22
FOXE1	rs375824	0.39
KIAA-VAX1	rs707816	0.17
SPRY2	rs800164	0.49
TPM1	rs187314	0.27
NOG1	rs227731	0.45
MAFB	rs130412	0.40

Meta Analysis by 4 study sites (post Haplin Strat)

R-syntax:

Run Haplin Strat

```
# Log-transformation of RRs
log.RR <- log(esti$RR.est)
# Pull approximate SE for log RR from confidence intervals
log.SE <- (log(esti$RR.upper) - log(esti$RR.lower))/(2*1.96)
# Weights for the meta-analysis are the inverse of the variances
w <- 1/log.SE^2
# Common log RR
tot.log.RR <- sum(w*log.RR)/sum(w)
# SE for common log RR
tot.log.SE <- 1/sqrt(sum(w))
# CI for common RR
tot.RR <- exp(c(est = tot.log.RR, lower = tot.log.RR - 1.96*tot.log.SE, upper = tot.log.RR + 1.96*tot.log.SE))
```
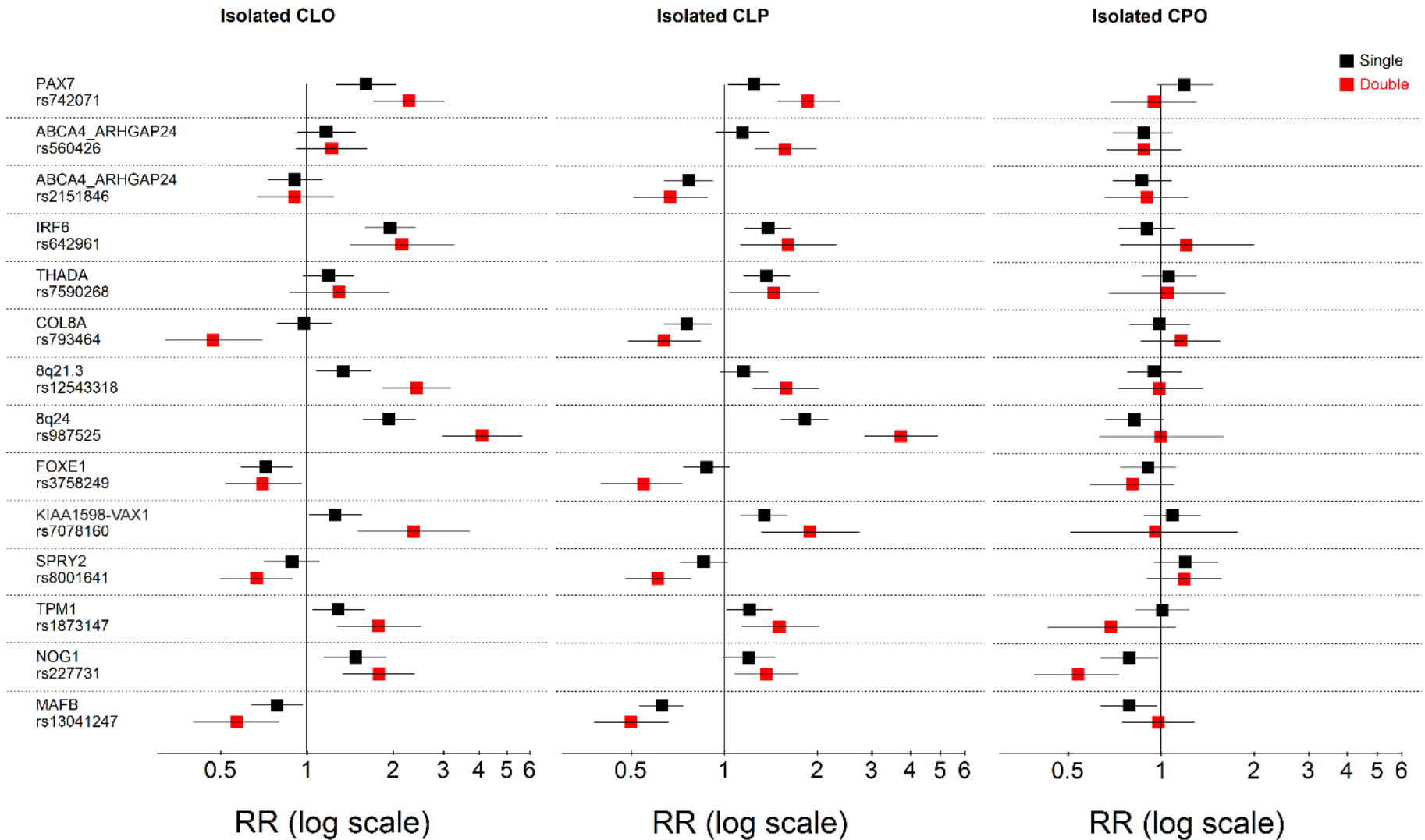
(heterogeneity test of RRs is a part of Haplin Strat)

Effects of child's alleles, 14 SNPs

We found no maternal gene effects, as expected

PoO-analysis for CLP (similar for CLO)

Gene/Locus	SNP	Maternal allele		Paternal allele		P-difference
		RR	95% CI	RR	95% CI	
PAX7	rs742071	1.21	0.94 - 1.56	1.34	1.06 - 1.71	0.58
ABCA4-ARHGAP29	rs560426	1.21	0.95 - 1.55	1.11	0.86 - 1.44	0.66
IRF6	rs642961	1.41	1.14 - 1.76	1.39	1.10 - 1.76	0.84
THADA	rs7590268	1.34	1.07 - 1.68	1.43	1.14 - 1.79	0.76
8q21.3	rs12543318	1.32	1.07 - 1.63	1.07	0.83 - 1.37	0.20
8q24	rs987525	1.65	1.30 - 2.10	2.09	1.68 - 2.61	0.09
FOXE1	rs3758249	0.88	0.71 - 1.09	0.90	0.71 - 1.13	0.89
VAX1	rs7078160	1.33	1.05 - 1.69	1.48	1.17 - 1.87	0.28
KIAA1598	rs4752028	1.27	0.99 - 1.62	1.55	1.23 - 1.95	0.15
SPRY2	rs8001641	0.94	0.76 - 1.17	0.78	0.61 - 1.00	0.21
TPM1	rs1873147	1.28	1.03 - 1.59	1.20	0.94 - 1.53	0.52
NOG1	rs227731	1.12	0.87 - 1.44	1.32	1.03 - 1.70	0.35
MAFB	rs13041247	0.59	0.46 - 0.75	0.68	0.53 - 0.86	0.62

Notes: P-difference is the p-value of the difference in effects between maternal and paternal alleles.

3. How much of the recurrence risk is explained by the effect of these 14 SNPs?

Assumptions and calculations - RR of recurrence produced by SNPs

- Hardy-Weinberg equilibrium
- Random mating
- Multiplicative effects (no interactions between SNPs)

For each SNP:

$$\text{RR}_{\text{recurrence}} = \left. \text{P}(\text{D}_{\text{child}}|\text{D}_{\text{parent}}) \middle/ \text{P}(\text{D}_{\text{child}}|not\ D_{parent}) \right.$$

$$\text{P}(\text{D}_{\text{child}}|\text{D}_{\text{parent}}) = \left. \text{P}(\text{D}_{\text{child}} \cap \text{D}_{\text{parent}}) \middle/ \text{P}(\text{D}_{\text{parent}}) \right.$$

$$\text{P}(\text{D}_{\text{child}} \cap \text{D}_{\text{parent}}) = \sum \underbrace{\text{P}(G_{child} \cap G_{parent})}_{\substack{\text{Frequency of} \\ \text{genotype combination}}} \underbrace{\text{P}(\text{D}_{\text{child}}|\text{G}_{\text{child}})\text{P}(\text{D}_{\text{parent}}|\text{G}_{\text{parent}})}_{\substack{\text{Risk estimate} \\ \text{for each genotype}}}$$

Joint risk of mother and child for each SNP

Allele frequency $P(a)=p$, $p(A)=1-p$

r=baseline risk for AA genotype, a is risk-allele

RR_s=relative risk single dose (aA), RR_d=relative risk double dose (aa)

Father:

Mother: genotype	mother's frequency	mother's risk	Father aa (p^2, RR_d*r) child's genotype	child's frequency	child's risk	Father aA ($2p(1-p)$, RR_s*r) child's genotype	child's frequency	child's risk	Father AA ($(1-p)^2$, r) child's genotype	child's frequency	child's risk
aa	p^2	RR_d*r	aa	p^4	RR_d*r	aa	$p^3(1-p)$	RR_d*r	aA	$p^2(1-p)^2$	RR_s*r
						aA	$p^3(1-p)$	RR_s*r			
aA	$2p(1-p)$	RR_s*r	aa	$p^3(1-p)$	RR_d*r	aa	$p^2(1-p)^2$	RR_d*r	aA	$p(1-p)^3$	RR_s*r
			aA	$p^3(1-p)$	RR_s*r	aA	$2p^2(1-p)^2$	RR_s*r	AA	$p(1-p)^3$	r
						AA	$p^2(1-p)^2$	r			
AA	$(1-p)^2$	r	aA	$p^2(1-p)^2$	RR_s*r	aA	$p(1-p)^3$	RR_s*r	AA	$(1-p)^4$	r
						AA	$p(1-p)^3$	r			

Linkage Strategies for Genetically Complex Traits. I. Multilocus Models

Neil Risch

Departments of Epidemiology and Public Health, and of Human Genetics, Yale University School of Medicine, New Haven, CT

Summary

In order to investigate linkage detection strategies for genetically complex traits, multilocus models of inheritance need to be specified. Here, two types of multilocus model are described: (1) a multiplicative model, representing epistasis (interaction) among loci, and (2) an additive model, which is shown to closely approximate genetic heterogeneity, which is characterized by no interlocus interaction. A ratio λ_R of risk for type R relatives that is compared with population prevalence is defined. For a single-locus model, $\lambda_R - 1$ decreases by a factor of two with each degree of relationship. The same holds true for an additive multilocus model. For a multiplicative (epistasis) model, $\lambda_R - 1$ decreases more rapidly than by a factor of two with degree of relationship. Examination of λ_R values for various classes of relatives can potentially suggest the presence of multiple loci and epistasis. For example, data for schizophrenia suggest multiple loci in interaction. It is shown in the second paper of this series that λ_R is the critical parameter in determining power to detect linkage by using affected relative pairs.

If risk loci can be combined in a multiplicative model, then their total effect on the recurrence risk ratio (RRR) is the product of the RRR from each locus

Tests of pairwise multiplicative interactions, case-only

Test of pairwise interactions with ordered logistic regression showing unadjusted p-values*. Case-only analysis shown in upper right triangle and corresponding check of association among controls shown in italics in lower left triangle.

	PAX7 rs742071	ABCA4 rs560426	IRF6 rs642961	THADA rs7590268	8Q21_3 rs12543318	8Q24 rs987525	FOXE1 rs3758249	KIAA1598 rs4752028	SPRY2 rs8001641	TPM1 rs1873147	NOG1 rs227731	MAFB rs13041247
PAX7 rs742071		.066	.89	.16	.89	.86	.48	.90	.32	.33	.78	1.0
ABCA4 rs560426	*.073*		.49	.68	.43	.97	.35	.18	.044	.29	.29	.49
IRF6 rs642961	*.71*	*.09*		.36	.48	.021	.731	.78	.30	.84	.52	.35
THADA rs7590268	*.48*	*.45*	*.58*			.45	.51	.95	.43	.21	.34	.15
8Q21_3 rs12543318	*.44*	*.80*	*.40*	*.74*		.22	.43	.61	.69	.65	.59	.36
8Q24 rs987525	*.98*	*.47*	*.98*	*.035*	*.79*		.47	.42	.77	.73	.92	.67
FOXE1 rs3758249	*.46*	*.18*	*.30*	*.44*	*.56*	*.49*		.17	.75	.52	.13	.36
KIAA1598 rs4752028	*.63*	*.78*	*.78*	*.099*	*.79*	*.65*	*.16*		.26	.49	.82	.76
SPRY2 rs8001641	*.10*	*.31*	*.055*	*.020*	*.16*	*.21*	*.74*	*.73*		.97	.84	.66
TPM1 rs1873147	*.51*	*.89*	*.14*	*.11*	*.49*	*.56*	*.091*	*.038*	*.55*		.73	.48
NOG1 rs227731	*.18*	*.13*	*.32*	*.93*	*.67*	*.27*	*.11*	*.14*	*.40*	*.32*		.064
MAFB rs13041247	*.61*	*.012*	*.46*	*.65*	*.39*	*.24*	*.56*	*.55*	*.031*	*.89*	*.67*	

* Threshold for significance with Bonferroni adjustment (66 tests) is 0.0008.

R-program SNPrec.R, CLP datafile

(Thanks to Julia)

```
snp_name,maf,r,eff_homozyg,CI_homo_l,CI_homo_u,eff_heterozyg,CI_het_l,CI_het_u
PAX7,0.4,0.00086,1.88,1.5,2.36,1.25,1.04,1.52
ABCA41,0.46,0.00086,1.59,1.26,2,1.15,0.95,1.41
ABCA42,0.42,0.00086,0.67,0.51,0.88,0.77,0.64,0.92
IRF6,0.2,0.00086,3.75,2.86,4.92,1.83,1.54,2.18
THADA,0.23,0.00086,1.46,1.04,2.03,1.38,1.17,1.64
COL8A,0.41,0.00086,0.65,0.49,0.85,0.76,0.64,0.91
8q21,0.34,0.00086,1.58,1.24,2.03,1.15,0.97,1.38
8q24,0.22,0.00086,3.75,2.86,4.92,1.83,1.54,2.18
FOXE1,0.39,0.00086,0.54,0.4,0.73,0.87,0.74,1.04
KIAAVAX,0.17,0.00086,2.05,1.43,2.93,1.36,1.14,1.62
SPRY2,0.49,0.00086,0.61,0.48,0.78,0.86,0.71,1.03
TPM1,0.27,0.00086,1.5,1.13,2,1.2,1.01,1.43
NOG1,0.45,0.00086,1.37,1.08,1.73,1.2,0.99,1.47
MAFB,0.4,0.00086,0.5,0.38,0.66,0.63,0.53,0.75
```

Soon available in Haplin ☺

How much of the recurrence risk is explained by the effect of these 14 SNPs?

	Parent-child RR of recurrence	RR by 14 SNPs	% of excess risk explained
CLO-CLO	42	1.47(1.35-1.60)	1.1%
CLP-CLP	29	1.45 (1.36-1.55)	1.6%
CPO-CPO	32	1.05 (1.02-1.07)	0.2%

$$\% \text{ explained} = \frac{\text{recurrence risk(SNPs) - population risk}}{\text{recurrence risk(family based) - population risk}}$$

$$= \frac{RR_{SNPs} - 1}{RR_{family} - 1}$$

Transformation of parent-offspring recurrence to heritability of liability

	RR (95% CI)	Absolute risks		h^2(family)*
		Recurrence	Reference	
CLO - CLO:	42 (31-56)	2.3 %	0.055 %	72%
CLP - CLP:	29 (22-37)	2.5 %	0.086 %	68%

Recurrence "predicted" by 14 SNPs:

	RR (95% CI)	Absolute risks		h^2("GWAS")*
		Recurrence	Reference	
CLO - CLO:	1.47 (1.35-1.60)	0.081%	0.055%	6%
CLP - CLP:	1.45 (1.36-1.55)	0.125%	0.086%	6%

Missing heritability of liability: h^2(family) - h^2("GWAS")

CLO: 72%-6%=66%

CLP: 68%-6%=62%

* Falconer DS. Inheritance of liability to certain diseases, estimated from the incidence among relatives. Ann Hum Genet 1965

Conclusions

- Categories of birth defects appear to have distinct (genetic) causes

- Fetal genes are likely to contribute and CLO and CLP have common etiology

- A hand-full of fetal SNPs have moderate effect for cleft lip

- Effects of these on CLO and CLP are similar

- The SNPs explain only very little of the recurrence risk

- Few SNPs are identified for cleft palate only

- Future: Huge samples, rare genetic variants, interactions ++