

Linkage Analysis in Familial Non-Lynch Syndrome Colorectal Cancer Families from Sweden

Vinaykumar Kontham¹, Susanna von Holst¹, Annika Lindblom¹

Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

Abstract

Family history is a major risk factor for colorectal cancer and many families segregate the disease as a seemingly monogenic trait. A minority of familial colorectal cancer could be explained by known monogenic genes and genetic loci. Familial polyposis and Lynch syndrome are two syndromes where the predisposing genes are known but numerous families have been tested without finding the predisposing gene. We performed a genome wide linkage analysis in 121 colorectal families with an increased risk of colorectal cancer. The families were ascertained from the department of clinical genetics at the Karolinska University Hospital in Stockholm, Sweden and were considered negative for Familial Polyposis and Lynch syndrome. In total 600 subjects were genotyped using single nucleotide polymorphism array chips. Parametric- and non-parametric linkage analyses were computed using MERLIN in all and subsets of families. No statistically significant result was seen, however, there were suggestive positive HLODs above two in parametric linkage analysis. This was observed in a recessive model for high-risk families, at locus 9q31.1 (HLOD=2.2, rs1338121) and for moderate-risk families, at locus Xp22.33 (LOD=2.2 and HLOD=2.5, rs2306737). Using families with early-onset, recessive analysis suggested one locus on 4p16.3 (LOD=2.2, rs920683) and one on 17p13.2 (LOD/HLOD=2.0, rs884250). No NPL score above two was seen for any of the families. Our linkage study provided additional support for the previously suggested region on chromosome 9 and suggested additional loci to be involved in colorectal cancer risk. Sequencing of genes in the regions will be done in future studies.

Citation: Kontham V, von Holst S, Lindblom A (2013) Linkage Analysis in Familial Non-Lynch Syndrome Colorectal Cancer Families from Sweden. PLoS ONE 8(12): e83936. doi:10.1371/journal.pone.0083936

Editor: Nathan A. Ellis, University of Illinois at Chicago, United States of America

Received: October 8, 2013; **Accepted:** November 18, 2013; **Published:** December 11, 2013

Copyright: © 2013 Kontham et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Financial support was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet (20110483), the Swedish Cancer Society (110439), The Swedish research council (20103543), The Stockholm Cancer Foundation (111232) and Nilsson-Ehle Foundation. The SNP technology platform in Uppsala is supported by Knut & Alice Wallenberg Foundation via the Wallenberg Consortium and Uppsala University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: Annika.Lindblom@ki.se

☉ These authors contributed equally to this work.

Introduction

Colorectal cancer (CRC) is increasing in incidence and is ranked as the second and third most common cancer type in the western world and Sweden respectively. CRC has a lifetime risk of 5% and affects men and women equally. One major risk factor is a family history of the disease and 20–25% of all CRC cases have a close relative with the same disease [1]. Known syndromes such as familial adenomatous polyposis (FAP) and Lynch syndrome are responsible for less than 5% of colorectal cancer cases [2], which leaves a majority of the familial colorectal cancer cases unexplained. The inheritance often suggests a dominant transmission of the disease, but recessive inheritance and even a complex inheritance have been suggested [3]. A syndrome, familial CRC type X, has been suggested for families fulfilling criteria for Lynch

syndrome but without germline mutations [4]. However, families negative after Lynch syndrome diagnostics only rarely fulfil these strict criteria, mostly because a later onset or reduced penetrance. Recently Genome Wide Association Studies (GWAS) have been used to find genetic loci associated with a certain risk to develop CRC. These loci; 6p21, 8q23.3, 8q24.21, 9p24, 10p14, 11q13.4, 11q23.1, 14q22.2, 15q13.3, 16q22.1, 18q21.1, 19q13.1 and 20p12.3, 1q41, 3q26.2, 12q13.13, 20q13.33, Xp22.2 [5–14] could to some degree support the explanation of CRC as a complex disease. Historically, linkage analysis has been a successful tool finding monogenic disease causing colorectal cancer genes like APC [15], MSH2 [16] and MLH1 [17]. Also new candidate regions for additional existence of moderate to high-penetrant CRC loci have been reported from linkage studies, but no casual mutation has yet been found. The loci on chromosome 9q

Table 1. Description of families in linkage analysis.

	Total	>3 CRC	mean <50	Sibs only
No. families	121	27	8	49
Mean age	64	62	48	66
No. families with any <50	27	8	8	6
No. families with any <60	72	21	8	22

doi: 10.1371/journal.pone.0083936.t001

[18-20], 3q [21,22] and 14q [23,24] have been reported more than once. A previous linkage study in familial serrated neoplasia suggested a locus on chromosome 2q [25]. Recently, 4 different loci with a significant HLOD above 3; on chromosomes 12q24 in all CRC families, 4q21 in early onset-, 15q22.31 in high-risk- and 8q13.2 in moderate-risk families was suggested [26]. The current study performed a genome wide linkage (GWL) scan using 600 individuals from 121 Swedish CRC families and analysed all families, early onset-, high-risk and moderate risk families separately.

Materials and Methods

Ethics statement

The study was undertaken in agreement with the Swedish legislation of ethical permission and according to the decision in the Stockholm regional ethical committee (2008/125-31.2). All participants gave written informed consent to participate in the study.

Patients

The families were ascertained through the clinical genetics department at the Karolinska University Hospital in Stockholm, Sweden between 1990 and 2005. FAP was excluded using medical records from affected individuals and Lynch syndrome was excluded using our current clinical protocol [27]. Families were included in the study if there was at least two affected relatives informative for linkage analysis, thus at least a sib-pair. Details on the families are shown in table 1. Early onset families were defined as families with a mean-age of diagnosis less than 50, high-risk families were defined as families with three or more affected individuals in close relatives. Moderate risk families were defined as families with two or more sibs affected. Eight families fulfilled criteria for CRC type X [4] (two overlapped with early onset families) but were not analysed separately.

Genotyping

Genomic DNA was extracted from peripheral blood using standard procedures. Genotyping was done separately in two different sets of family material. Genotyping of 548 patients with 6090 markers was performed by the Illumina Infinium assay [28,29] using the Illumina HumanLinkage-12 DNA analysis bead chip. The overall reproducibility in the genotype data was 99.996% based on 6.25% of duplicated genotypings. The average call rate per SNP was 99.57%. Additionally, 52 subjects were genotyped using the Illumina Golden Gate assay

[30] and the Illumina Linkage Panel IVb (6008 markers). The overall reproducibility in the genotype data was 100% based on 2.2% of duplicated genotypings. The average call rate per SNP was 97.27%. Arrays were processed according to manufactures protocol at the SNP Technology Platform in Uppsala and available upon request.

Linkage analysis

Pedcheck [31] was used to check for the initial Mendelian inheritance analysis among the families. The family based genetic model was used for parametric linkage analysis for all chromosomes, including chromosome X. As a supplement non-parametric analysis using Whittemore and Halpern NPL statistics was made [32]. LOD scores as well as heterogeneity LOD scores were computed using MERLIN (version 1.1.2) [33] and was given for all genotyped positions. Analyses were done assuming both dominant and recessive traits. For autosomal dominant and recessive mode of inheritance the disease allele frequency was set to 0.0001. The penetrance rates for the dominant and recessive mode of inheritance for homozygous normal, heterozygous, and homozygous affected were set to 0.05, 0.80, 0.80 and 0.001, 0.001, 1.0 respectively.

Individuals with colorectal cancer or a polyp with high degree dysplasia were coded as affected. Family members with unclear status were coded as unknown. Families were ascertained assuming a dominant trait and spouses were therefore coded as unaffected. Four different analyses were performed using different sets of families and patients; all families, all families with at least three cases (high-risk), all families with CRC among sibs (moderate-risk) and families with a mean age below 50 (early-onset) to be comparable with the results from the recent linkage study by Cicek et al.

All 121 families, including all subjects from both genotyping sessions, were used for linkage analysis. Thus, two marker files were merged and 7256 markers were used in the analysis. Merlin by default allows a maximum of 24 bits for each family, why four large families had to be split. The families were split so that each sub-family used one common ancestor and fitted into the limit as defined while running the program. The original 121 families were analysed as 126 (one family had to be split into three).

Since presence of linkage disequilibrium (LD) may inflate multipoint linkage statistics, a threshold of $r^2 = 0.1$ has been used to avoid that false positive results inflate the statistics [34]. LD among SNPs with $r^2 > 0.1$ was accounted for, by MERLIN organizing the markers into clusters. MERLIN makes use of the population haplotype frequencies to assume LD within each cluster. To maintain uniformity in our study subsets, the same clusters were continuously used in all analysis.

Results

A total of 600 individuals from 121 families were successfully genotyped. The analyses were conducted for all and each of three different subgroups; high-risk, moderate-risk and early-onset families (table 1). There was no individual statistically significant (above three) LOD score or HLOD in any of the

analysis (Figure 1). However, there were positive HLODs above two (table 2).

For high-risk families (in total 27 families) one locus on chromosome 9q31.1, showed an HLOD above two assuming recessive inheritance and an estimated 66% of families linked (table 2). Maximum was for the marker rs1338121 with a HLOD of 2.2. LOD on the same locus in recessive model was 0.7. For dominant disease this locus showed a LOD score of 1.4 and an HLOD of 1.6.

For moderate-risk families (in total 49 families, analysed as 50) one locus on the tip on chromosome Xp had LOD and HLOD above two with maximum 2.2 and 2.5 respectively, for the marker rs2306737 in recessive analysis (table 2). LOD score and HLOD for dominant analysis were both 1.8.

Finally, for the group of families with early-onset (only 8 families) two loci showed positive LOD scores and HLODs above two in recessive analysis (table 2). One was distal on chromosome 4p with a maximum LOD and HLOD of 2.2 for rs920683 and the other was on chromosome 17p13.2 with a maximum LOD and HLOD of 2.0 for rs884250. The LOD and HLOD for dominant analysis was 0.976 on chromosome 4p and 1.25 on chromosome 17p13.2. The LODs were similar in parametric and nonparametric analysis and 100% linked families were assumed for both loci.

No NPL score above two was seen for any of the families. HLODs above 2 are presented in table 2. All loci with an HLOD>1.0 in parametric analysis are shown in table 3.

Discussion

We used SNP genotyping to perform a linkage analysis in 121 CRC families and did not find any overall statistically significant results with a LOD or HLOD over 3. A few large families were split to be able to use MERLIN, and lost a bit of its power in the analysis. The effect of this was of little importance.

However, we did find LODs and HLODs above 2, suggestive of linkage. A previous linkage study used 356 families and showed one locus with HLOD > 3 (12q) and 4 with HLOD > 2 (on chromosomes 4q, 15q, 17q and 12q), all in dominant analysis [26]. We found no support for any of these regions in our analysis.

In our substudy of large, high-risk families (more than three affecteds), a locus on chromosome 9q31 was found. The same region has been suggested before, although in dominant models, and was here again identified by us using linkage analysis in a recessive model. This locus was previously suggested by one sib-pair study and one study using linkage in many kindred and also previously by us in one large family with rectal cancer and adenomas [18-20]. Those three studies used microsatellites for genotyping while previous linkage studies using SNPs were unable to replicate the locus [18,26,35]. One study has also shown support for the region using a five-SNP haplotype in the region [36]. Two genes have so far been suggested as the predisposing genes in this region. First, it was suggested that germline allele-specific expression resulted in reduced expression of the gene altering SMAD-mediated TGF-beta signaling [37]. Secondly, a study of the GALNT12

gene demonstrated truncating somatic and germline mutations in CRC patients but none in controls and genetic defects in the O-glycosylation pathway in part underlie aberrant glycosylation, and thereby contribute to the development in a subset of CRC [38]. The region suggested by our present study overlaps well with the region suggested by one family before [19]. The region spans over almost 9 Mb and includes the above mentioned genes and numerous of others. Four families contribute mostly to the positive, dominant and recessive HLOD score. Only one family had early-onset disease. The Cicek study defined for a similar group of 67 families one locus with HLOD >3 (15q) and 4 with HLOD > 2 (chromosomes 12q, 14q, 17q and Xp) some using dominant or recessive (chromosome 14) model. We found no support for any of those regions in our study. However, we had an HLOD>1 close to the chromosome X region.

For moderate-risk families, with affected sibs only, one region on the tip on chromosome X was suggested in recessive model (max HLOD=2.2) and similar to the chromosome 9 locus there was also a positive but lower LOD using a dominant model (0.7). This region is almost 6 Mb, has not been suggested before and contains numerous candidate genes. The Cicek study included 200 moderate risk families and also using recessive model they identified one locus with an HLOD>3 (8q) and 4 loci with an HLOD>2 (chromosomes 1q, 6p, 8q and 22q). We could not find support for any of those loci.

Finally we observed using a recessive model two loci with high HLODs (4p and 17p) for the early-onset families. However, this group consisted of only 8 families. Family 8 was also among the high-risk families linked to chromosome 9 locus and for this size of family it is expected that a whole genome study should generate linkage to many regions and thus most will be false positives. We could not replicate any of the loci from Cicek et al. using dominant or recessive model (chromosomes 4q, 14q, 15q and 22q).

A previous study focusing on adenoma and colorectal adenoma and carcinoma found linkage to chromosomal regions on chromosomes 18q21 and 2p22 in 69 families while a sub-analysis of 55 families with cancer only showed linkage to chromosome 3q21-24[21]. None of these regions were confirmed by the Cicek study or by this study. It is surprising that 4 linkage studies including this, all using SNP-markers do not generate any overlapping loci [21,26,35]. There are several differences between the studies though, which are possible explanations for this discrepancy. The ethnicity of the subjects in the studies is different, one study is from USA, one is from the Netherland's, one from UK and our study is based on the Swedish population. The sample sizes are also different, the US study have in total 356 families, the Dutch only seven large families, the UK study 69 and our study 121 families. The recruitment process also differed between all four studies, but in general most of the different results could be explained also by biology and different predisposing elements in each sample sets. Future experiments need to consider this heterogeneity in their design.

Slightly to our surprise we could see also positive parametric LOD-scores in this study compared to our previous ones

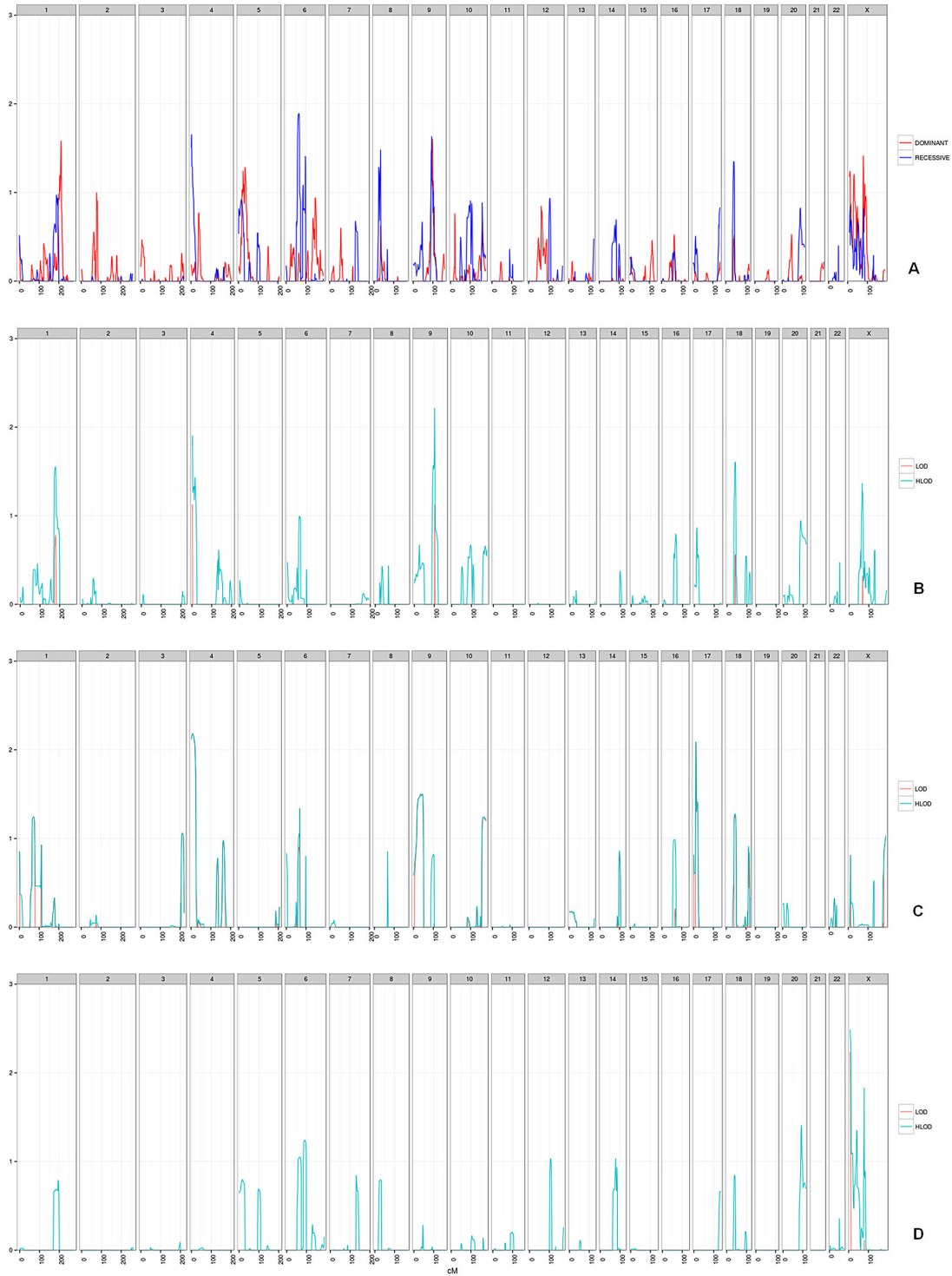


Figure 1. Plot of LOD / HLOD*. a) Plot of HLODs for all families in the study, dominant (in Red) and recessive (in Blue) models (n=121).

b) LOD/HLOD plot of study group with more than 3 affected individuals (n=27).

c) LOD/HLOD plot of study group with mean age of diagnosis < 50 (n=8).

d) LOD/HLOD plot of study group with affected sibs (n=49).

* -negative values of scores are not plotted shown.

For b,c,d - LODs are represented in Red, HLODs are represented in Cyan.

doi: 10.1371/journal.pone.0083936.g001

Table 2. Summary of Colorectal cancer linkage results with maximum observed HLODs greater than 2.

Study Group	No. Of Families	Linked Region	cM, SNP	Model	HLOD (α)
All families	121	-	-	-	-
More than 3 affected	27	9q31.1	102.68, RS1338121	Recessive	2.212 (0.66)
Mean age at Diagnosis < 50	8	4p16.3	7.17, RS920683	Recessive	2.184 (1.00)
		17p13.2	11.51, RS884250	Recessive	2.086 (1.00)
Families with Affected Sibs	49	Xp22.33	7.42, RS2306737	Recessive	2.486 (0.79)

doi: 10.1371/journal.pone.0083936.t002

Table 3. Summary of Colorectal cancer linkage results with maximum observed HLODs between 1 and 2.

Study Group	No. Of Families	Linked Region	cM, SNP	Model	HLOD (α)
All Families	121	1q32.1	204.36, RS2032018	Dominant	1.191 (0.46)
		5p15.2	33.99, RS879253	Dominant	1.258 (0.48)
		9q22.31	97.96, RS4534181	Dominant	1.602 (0.59)
			94.37, RS7037744	Recessive	1.632 (0.32)
		Xp11.21	79.25, RS2015312	Dominant	1.414 (0.72)
		4p16.3	2.97, RS736455	Recessive	1.653 (0.38)
		6p21.1	64.36, RS722269	Recessive	1.892 (0.28)
		8p22	33.01, RS334206	Recessive	1.479 (0.26)
		18p11.21	40.13, RS1043925	Recessive	1.351 (0.27)
More than 3 affected	27	1q32.2	211.46, RS1507765	Dominant	1.414 (0.61)
		2p16.2	77.83, RS1483869	Dominant	1.438 (0.60)
		4q28.3	134.94, RS426029	Dominant	1.045 (0.54)
		5p15.1	34.80, RS1505034	Dominant	1.678 (0.71)
		9q31.1	102.68, RS1338121	Dominant	1.672 (0.79)
		12q13.12	63.98, RS7532	Dominant	1.086 (0.51)
		16q12.2	65.68, RS1990637	Dominant	1.556 (0.77)
		Xp11.21	79.24, RS1560514	Dominant	1.788 (1.00)
		1q25.2	178.47, RS227530	Recessive	1.554 (0.55)
		4p16.3	2.97, RS736455	Recessive	1.903 (0.68)
		18p11.21	40.13, RS1043925	Recessive	1.605 (0.48)
		Xp11.3	67.42, RS1137070	Recessive	1.267 (0.56)
Mean age at Diagnosis < 50	8	2p16.3	74.90, RS1394207	Dominant	1.145 (1.00)
		9p21.3	45.61, RS10757309	Dominant	1.614 (1.00)
		10q22.1	86.06, RS1227938	Dominant	1.200 (1.00)
		16q21	80.78, RS17822576	Dominant	1.181 (1.00)
		17p13.1	24.87, RS1391766	Dominant	1.258 (1.00)
		1p33	72.10, RS1934405	Recessive	1.244 (1.00)
		6p21.1	67.58, RS4714772	Recessive	1.335 (1.00)
		9p21.3	45.61, RS10757309	Recessive	1.501 (1.00)
		10q26.3	166.39, RS7072831	Recessive	1.239 (0.90)
		18q11.2	42.92, RS12959039	Recessive	1.279 (1.00)
Families with Affected Sibs	49	4p15.2	38.72, RS216113	Dominant	1.112 (0.96)
		6q23.3	136.59, RS975676	Dominant	1.848 (1.00)
		Xp22.33	11.69, RS749706	Dominant	1.860 (1.00)
		6q14.1	91.67, RS885582	Recessive	1.241 (0.33)
		12q23.1	107.90, RS17290272	Recessive	1.034 (0.26)
		14q24.3	78.12, RS888412	Recessive	1.032 (0.29)
		20q13.31	93.25, RS186659	Recessive	1.410 (0.37)

doi: 10.1371/journal.pone.0083936.t003

[22,23] where only HLODs were obtained. This might be related to the fact that we used SNPs instead of microsatellites

in this study. Microsatellites are much more informative and thus often very low negative LODs are seen when a family is

not linked. In this SNP oriented experiment few families had LODs below -2 and thus the power to exclude linkage in our experiment was low. However, power to detect linkage was retained (which we tested using our family 24 with a LOD of 3 for the chromosome 9 locus) [19]. The result could also relate to the fact that families in this study were smaller compared to our previous studies using 20 and 30 large pedigrees [22,23]. Here we used a different strategy with 121 smaller families and in contrast to our previous linkage studies we could find support to the candidate region on chromosome 9. We previously used SimWalk2 for linkage analysis with microsatellite markers but it was not conducive for analysis of SNP data. To justify using MERLIN, we analysed family 24 and chromosome 9 using SimWalk2 (data not shown), which took three weeks to complete but obtained identical results.

In conclusion, our linkage study provided additional support for the region on chromosome 9, some evidence for new loci to be involved in colorectal cancer risk and no support for other

previously reported loci. This suggested heterogeneity if familial CRC should influence the design of future association and linkage studies.

Acknowledgements

Genotyping was performed by the SNP Technology Platform in Uppsala, (www.genotyping.se). Support by Bioinformatics Infrastructure for Life Sciences (BILS) is gratefully acknowledged.

Author Contributions

Conceived and designed the experiments: AL VK SvH. Performed the experiments: VK SvH. Analyzed the data: VK SvH. Contributed reagents/materials/analysis tools: AL VK SvH. Wrote the manuscript: AL VK SvH.

References

- Olsson L, Lindblom A (2003) Family history of colorectal cancer in a Sweden county. *Fam Cancer* 2: 87-93. doi:10.1023/A:1025734200635. PubMed: 14574157.
- Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF et al. (2009) Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin Genet* 76: 1-18. doi:10.1111/j.1399-0004.2009.01230.x. PubMed: 19659756.
- Penegar S, Wood W, Lubbe S, Chandler I, Broderick P et al. (2007) National study of colorectal cancer genetics. *Br J Cancer* 97: 1305-1309. doi:10.1038/sj.bjc.6603997. PubMed: 17895893.
- Lindor NM, Rabe K, Petersen GM, Haile R, Casey G et al. (2005) Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency - Familial colorectal cancer type X. *JAMA: Journal of the American Medical Association* 293: 1979-1985. doi:10.1001/jama.293.16.1979.
- Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A et al. (2007) Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 39: 989-994. doi: 10.1038/ng2089. PubMed: 17618283.
- Tomlinson IP, Webb E, Carvajal-Carmona L, Broderick P, Howarth K et al. (2008) A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet* 40: 623-630. doi:10.1038/ng.111. PubMed: 18372905.
- Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z et al. (2007) A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet* 39: 984-988. doi:10.1038/ng2085. PubMed: 17618284.
- Tenesa A, Farrington SM, Prendergast JG, Porteous ME, Walker M et al. (2008) Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet* 40: 631-637. doi:10.1038/ng.133. PubMed: 18372901.
- Jaeger E, Webb E, Howarth K, Carvajal-Carmona L, Rowan A et al. (2008) Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat Genet* 40: 26-28. doi:10.1038/ng.2007.41. PubMed: 18084292.
- Houlston RS, Webb E, Broderick P, Pittman AM, Di Bernardo MC et al. (2008) Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet* 40: 1426-1435. doi:10.1038/ng.262. PubMed: 19011631.
- Houlston RS, Cheadle J, Dobbins SE, Tenesa A, Jones AM et al. (2010) Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat Genet* 42: 973-977. doi:10.1038/ng.670. PubMed: 20972440.
- Haiman CA, Le Marchand L, Yamamoto J, Stram DO, Sheng X et al. (2007) A common genetic risk factor for colorectal and prostate cancer. *Nat Genet* 39: 954-956. doi:10.1038/ng2098. PubMed: 17618282.
- Broderick P, Carvajal-Carmona L, Pittman AM, Webb E, Howarth K et al. (2007) A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat Genet* 39: 1315-1317. doi:10.1038/ng.2007.18. PubMed: 17934461.
- Dunlop MG, Dobbins SE, Farrington SM, Jones AM, Palles C et al. (2012) Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nat Genet* 44: 770-776. doi: 10.1038/ng.2293. PubMed: 22634755.
- Bodmer WF, Bailey CJ, Bodmer J, Bussey HJ, Ellis A et al. (1987) Localization of the gene for familial adenomatous polyposis on chromosome 5. *Nature* 328: 614-616. doi:10.1038/328614a0. PubMed: 3039373.
- Peltomäki P, Aaltonen LA, Sistonen P, Pylkkänen L, Mecklin JP et al. (1993) Genetic mapping of a locus predisposing to human colorectal cancer. *Science* 260: 810-812. doi:10.1126/science.8484120. PubMed: 8484120.
- Lindblom A, Tannergård P, Werelius B, Nordenskjöld M (1993) Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nat Genet* 5: 279-282. doi:10.1038/ng1193-279. PubMed: 7903889.
- Kemp ZE, Carvajal-Carmona LG, Barclay E, Gorman M, Martin L et al. (2006) Evidence of linkage to chromosome 9q22.33 in colorectal cancer kindreds from the United Kingdom. *Cancer Res* 66: 5003-5006. doi:10.1158/0008-5472.CAN-05-4074. PubMed: 16707420.
- Skoglund J, Djureinovic T, Zhou XL, Vandrovцова J, Renkonen E et al. (2006) Linkage analysis in a large Swedish family supports the presence of a susceptibility locus for adenoma and colorectal cancer on chromosome 9q22.32-31.1. *J Med Genet* 43: e7. PubMed: 16467217.
- Wiesner GL, Daley D, Lewis S, Ticknor C, Platzer P et al. (2003) A subset of familial colorectal neoplasia kindreds linked to chromosome 9q22.2-31.2. *Proc Natl Acad Sci U S A* 100: 12961-12965. doi:10.1073/pnas.2132286100. PubMed: 14566058.
- Kemp Z, Carvajal-Carmona L, Spain S, Barclay E, Gorman M et al. (2006) Evidence for a colorectal cancer susceptibility locus on chromosome 3q21-q24 from a high-density SNP genome-wide linkage scan. *Hum Mol Genet* 15: 2903-2910. doi:10.1093/hmg/ddl231. PubMed: 16923799.
- Picelli S, Vandrovцова J, Jones S, Djureinovic T, Skoglund J et al. (2008) Genome-wide linkage scan for colorectal cancer susceptibility genes supports linkage to chromosome 3q. *BMC Cancer* 8: 87. doi: 10.1186/1471-2407-8-87. PubMed: 18380902.
- Djureinovic T, Skoglund J, Vandrovцова J, Zhou XL, Kalushkova A et al. (2006) A genome wide linkage analysis in Swedish families with hereditary non-familial adenomatous polyposis/non-hereditary non-polyposis colorectal cancer. *Gut* 55: 362-366. doi:10.1136/gut.2005.075333. PubMed: 16150854.
- Saunders IW, Ross J, Macrae F, Young GP, Blanco I et al. (2012) Evidence of linkage to chromosomes 10p15.3-p15.1, 14q24.3-q31.1 and 9q33.3-q34.3 in non-syndromic colorectal cancer families. *Eur J Hum Genet* 20: 91-96. doi:10.1038/ejhg.2011.149. PubMed: 21829229.
- Roberts A, Nancarrow D, Clendenning M, Buchanan DD, Jenkins MA et al. (2011) Linkage to chromosome 2q32.2-q33.3 in familial serrated

- neoplasia (Jass syndrome). *Fam Cancer* 10: 245-254. doi:10.1007/s10689-010-9408-8. PubMed: 21165777.
26. Cicek MS, Cunningham JM, Fridley BL, Serie DJ, Bamlet WR et al. (2012) Colorectal cancer linkage on chromosomes 4q21, 8q13, 12q24, and 15q22. *PLOS ONE* 7: e38175. doi:10.1371/journal.pone.0038175. PubMed: 22675446.
 27. Lagerstedt Robinson K, Liu T, Vandrovцова J, Halvarsson B, Clendenning M et al. (2007) Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. *J Natl Cancer Inst* 99: 291-299. doi: 10.1093/jnci/djk051. PubMed: 17312306.
 28. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS (2005) A genome-wide scalable SNP genotyping assay using microarray technology. *Nat Genet* 37: 549-554. doi:10.1038/ng1547. PubMed: 15838508.
 29. Steemers FJ, Chang W, Lee G, Barker DL, Shen R et al. (2006) Whole-genome genotyping with the single-base extension assay. *Nat Methods* 3: 31-33. doi:10.1038/nmeth842. PubMed: 16369550.
 30. Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F et al. (2003) Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol* 68: 69-78. doi:10.1101/sqb.2003.68.69. PubMed: 15338605.
 31. O'Connell JR, Weeks DE (1998) PedCheck: A program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63: 259-266. doi:10.1086/301904. PubMed: 9634505.
 32. Whittemore AS, Halpern J (1994) A Class of Tests for Linkage Using Affected Pedigree Members. *Biometrics* 50: 118-127. doi: 10.2307/2533202. PubMed: 8086596.
 33. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30: 97-101. doi:10.1038/ng786. PubMed: 11731797.
 34. Goode EL, Badzioch MD, Jarvik GP (2005) Bias of allele-sharing linkage statistics in the presence of intermarker linkage disequilibrium. *BMC Genet* 6 Suppl 1: S82. doi:10.1186/1471-2156-6-S1-S82. PubMed: 16451697.
 35. Middeldorp A, Jagmohan-Changur SC, van der Klift HM, van Puijtenbroek M, Houwing-Duistermaat JJ et al. (2010) Comprehensive genetic analysis of seven large families with mismatch repair proficient colorectal cancer. *Genes Chromosomes Cancer* 49: 539-548. PubMed: 20222047.
 36. Gray-McGuire C, Guda K, Adrianto I, Lin CP, Natale L et al. (2010) Confirmation of linkage to and localization of familial colon cancer risk haplotype on chromosome 9q22. *Cancer Res* 70: 5409-5418. doi: 10.1158/0008-5472.CAN-10-0188. PubMed: 20551049.
 37. Valle L, Serena-Acedo T, Liyanarachchi S, Hampel H, Comeras I et al. (2008) Germline allele-specific expression of TGFBR1 confers an increased risk of colorectal cancer. *Science* 321: 1361-1365. doi: 10.1126/science.1159397. PubMed: 18703712.
 38. Guda K, Moinova H, He J, Jamison O, Ravi L et al. (2009) Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers. *Proc Natl Acad Sci U S A* 106: 12921-12925. doi:10.1073/pnas.0901454106. PubMed: 19617566.