Title: Possible genetic predisposition to lymphedema after breast cancer

Running Title: Genetic predisposition to secondary lymphedema

Authors:

Beth Newman PhD[1], Felicity Lose PhD[2], Mary-Anne Kedda PhD[1], Mathias Francois PhD[3], Kaltin Ferguson[2], Monika Janda PhD[1], Patsy Yates PhD[4], Amanda B. Spurdle PhD[2*], Sandra C. Hayes PhD[1*†]

Affiliations:

[1]Institute of Health and Biomedical Innovation and School of Public Health, Queensland University of Technology, Brisbane, Australia

[2]Molecular Cancer Epidemiology Laboratory, Genetics and Population Health Division, Queensland Institute of Medical Research

[3]Institute of Molecular Bioscience, University of Queensland, Brisbane, Australia

[4]Institute of Health and Biomedical Innovation and School of Nursing, Queensland University of Technology, Brisbane, Australia

[*] ABS lead laboratory-related activities, while SCH lead field-related activities.

[†]**Corresponding author:** Sandra C Hayes, Queensland University of Technology, School of Public Health, Victoria Park Road, Kelvin Grove, Brisbane, Queensland, Australia, 4059; email: sc.hayes@qut.edu.au; phone: +617 3138 9645 fax: +617 3138 3130

Abstract

Background: Known risk factors for secondary lymphedema only partially explain who develops lymphedema following cancer, suggesting that inherited genetic susceptibility may influence risk. Moreover, identification of molecular signatures could facilitate lymphedema risk prediction prior to surgery or lead to effective drug therapies for prevention or treatment. Recent advances in the molecular biology underlying development of the lymphatic system and related congenital disorders implicate a number of potential candidate genes to explore in relation to secondary lymphedema.

Methods and Results: We undertook a nested case-control study, with participants who had developed lymphedema after surgical intervention within the first 18 months of their breast cancer diagnosis serving as cases (n=22) and those without lymphedema serving as controls (n=98), identified from a prospective, population-based, cohort study in Queensland, Australia. TagSNPs that covered all known genetic variation in the genes *SOX18*, *VEGFC*, *VEGFD*, *VEGFR2, VEGFR3*, *RORC*, *FOXC2, LYVE1, ADM* and *PROX1* were selected for genotyping. Multiple SNPs within three receptor genes, *VEGFR2, VEGFR3* and *RORC*, were associated with lymphedema defined by statistical significance ($p<0.05$) or extreme risk estimates (OR<0.5 or >2.0).

Conclusions: These provocative, albeit preliminary, findings regarding possible genetic predisposition to secondary lymphedema following breast cancer treatment warrant further attention for potential replication using larger datasets.

Condensed Abstract

We undertook a nested case-control study, with participants who developed

lymphedema after breast cancer surgery serving as cases (n=22) and those without lymphedema serving as controls (n=98), identified from a prospective cohort study. TagSNPs that covered all known genetic variation in the genes *SOX18*, *VEGFC*, *VEGFD*, *VEGFR2, VEGFR3*, *RORC*, *FOXC2, LYVE1, ADM* and *PROX1* were selected for genotyping. Multiple SNPs within *VEGFR2, VEGFR3* and *RORC*, were associated with lymphedema defined by statistical significance ($p<0.05$) or extreme risk estimates (OR<0.5 or >2.0). These results provide provocative, albeit preliminary, findings in the identification of molecular signatures that could facilitate lymphedema risk prediction.

Keywords: breast cancer, lymphedema, genetic predisposition, risk factors

Introduction

Lymphedema is one of the most problematic complications following breast cancer treatment, experienced by approximately 30% of breast cancer survivors.[1-3] It represents failure of the lymphatic system to adequately drain fluid and proteins from the interstitial tissue and to circulate lymphocytes. Removal or damage to the lymph nodes or lymphatic vasculature during cancer treatment may impede proper physiological function of this network. Although lymphedema can occur in any part of the body, it generally refers to an accumulation of fluid and subsequent distortion of a limb.[4] Little is known about its prevention, and it is regarded as an incurable, progressive, disfiguring and disabling disorder that is difficult to manage, compromising function[5] and quality of life.[6]

Lymphedema may present immediately or years after breast cancer treatment,[7] although the majority of cases seem to appear within the first 12-18 months post-surgery.[3, 8-9] The published literature on risk factors is characterized by inconsistent relationships, but evidence is mounting for a few, including extent of surgery, extent of lymph node resection, radiation therapy, obesity, and surgical wound infection.[10] Nevertheless, it is clear that these characteristics only partially explain who develops lymphedema, and lymphedema can and does occur in women lacking these risk factors. It is therefore possible that inherited genetic susceptibility may play a role in the pathogenesis of secondary lymphedema.

There has been substantial progress in identifying genes that contribute to development of the lymphatic vascular system during embryogenesis and its subsequent regulation.[11-14] Some genes are now known to underlie primary

lymphedema,[15-18] a congenital or later-onset condition that occurs in the absence of any known injury or medical intervention, and findings from genetic studies of inherited lymphedema assist with the molecular dissection of lymphatic diseases.[19] We hypothesized that genes involved in familial lymphedema (*VEGFR3* (*flt4*), *FOXC2*[20] and *SOX18*[15]) and/or lymphangiogenesis in the embryo, such as *VEGFC*, *VEGFD (also known as FIGF)*, *VEGFR2 (KDR), RORC*, *LYVE1, ADM (Adrenomedullin)* and *PROX1,*[20-22] may also predispose to secondary lymphedema. We therefore undertook a comprehensive investigation of genetic variation in these 10 plausible candidate genes in a cohort of breast cancer survivors, to assess the role of inherited genetic susceptibility in the development of secondary lymphedema after breast cancer.

Materials and Methods

The study was a nested case-control design involving participants who had developed lymphedema within the first 18 months of their breast cancer diagnosis serving as cases and those without lymphedema serving as controls, identified from a prospective, population-based, cohort study called the 'Pulling Through Study' (PTS) conducted in Queensland, Australia.

Study design and sample recruitment of the original 'Pulling Through Study'

Recruitment and study design for the PTS have been described in detail elsewhere.[9] In brief, 417 women with primary, unilateral, invasive breast cancer, diagnosed in 2002, were randomly selected from the Queensland Cancer Registry and invited to participate in the PTS. Of these, informed consent was obtained for 68% (n=287). Starting at six months post-diagnosis, women were prospectively followed

for 12 months, with data collection procedures involving completion of a clinical assessment and/or self-administered questionnaire every three months. Lymphedema status was evaluated using the sum of arm circumferences (SOAC) method.[23] A woman who scored a difference of >5cm between the treated and untreated sides, during any of the five data collection sessions between 6 and 18 months post-diagnosis, was considered a lymphedema case. The remaining women were considered controls. Some of the women (26%) participated on a questionnaire-only basis; hence they lack objective assessments of lymphedema and therefore were not available for inclusion in these analyses.

Design and sample recruitment of the nested case-control study

The follow-up, nested case-control study reported here commenced approximately six years following the date of breast cancer diagnosis for those in the PTS. Of the 287 original participants, 11 withdrew from the study and 16 had died (identified through the Queensland Cancer Registry mortality database), leaving 260 women to be re-contacted. Address details were checked with the electronic White Pages, and a change of address search was carried out through Australia Post. When an address could not be confirmed from these sources, the last postal address recorded in our files was used. The 260 potential participants were mailed an introductory letter (reminding them of their involvement in the prior study and inviting their participation in the current project), a newsletter (detailing findings from the original study), and a project information sheet.

Following consent, collection of blood samples was arranged using the phlebotomist at the Queensland University of Technology (QUT) or from a local

commercial pathology collection center, as preferred by the participant. Approximately 8ml of blood was drawn from the unaffected arm into yellow-top, ACD tubes, and was transferred to QUT's laboratory for processing. Upon receipt, samples were centrifuged, and the buffy coat (white cells) removed, aliquoted, labeled with a unique code and stored at -80°C. Once samples were collected and processed from all participants, samples were transferred to the Queensland Institute of Medical Research, where genomic DNA was extracted from frozen buffy coat cells using the salting-out extraction method, quantified (Nanodrop Spectrophotometer ND-1000, Nanodrop Technologies, USA), diluted to a standardised concentration, and 12ng of each sample plated onto 384-well plates for genotype analysis.

Genetic analyses

Tag single nucleotide polymorphisms (tagSNPs) that covered all known genetic variation in *SOX18*, *VEGFC*, *VEGFD*, *VEGFR2, VEGFR3*, *RORC*, *FOXC2, LYVE1, ADM* and *PROX1* were selected for genotyping from HapMap data release 24/phase II, November 2008, NCBI build 36, dbSNP b126 (www.hapmap.org), using the Tagger program within Haploview version 4.1[24] on CEU samples only. To minimize the number of genotypes tested while optimizing the number of polymorphisms evaluated, TagSNPs were chosen with an $r^2 \geq 0.8$ using the pair-wise tagging approach. Several additional polymorphisms previously shown to have functional significance in primary lymphedema were also genotyped. To enhance coverage of each locus, we included SNPs within 5 kilobases (kb) of the 5' and 3' ends of each gene, based on the long splice variant where relevant.

SNPs were genotyped using iPLEX Gold assays on the Sequenom MassARRAY platform (Sequenom, San Diego, USA), as described previously.[25] There were four negative (H_2O) controls per 384-well plate, and quality control parameters included genotype call rates >95%, inclusion of 20 duplicate samples per 384-well plate (>5% of samples) with ≥98% concordance between duplicates, and Hardy-Weinberg Equilibrium p-values ≥0.05. For assays not found to be polymorphic, rare-allele frequencies were confirmed using SPSmart (http://spsmart.cesga.es/).[26] Across the 10 genes investigated, 152 SNPs were attempted, but 16 SNPs failed assay design or quality control standards and hence were excluded from further analysis. After genotyping, the Broad Institute SNAP (SNP Annotation and Proxy Search) proxy search tool was used to determine SNPs tagged by genotyped tagSNPs for bioinformatics analyses, using the 1000 genomes SNP data set (http://www.broadinstitute.org/mpg/snap/ldsearch.php).[27]

Statistical analyses

All statistical analyses were conducted by SPSS, version 18. The genotype frequency distributions among cases and controls were compared using unadjusted logistic regression analysis. Results are reported using odds ratios (ORs) with 95% confidence intervals (CI). Co-dominant mode of inheritance (i.e., rare-allele homozygote, heterozygote, and reference category of common-allele homozygote) was assumed in statistical analyses, using a trend test with 1 degree of freedom. Results were interpreted initially by means of p-value (<0.05). However, because of the limited statistical power available, it was determined *a priori* that ORs >2.0 (or equivalently, <0.5) would be acknowledged, irrespective of statistical significance, as long as results conformed to a pattern consistent with straightforward Mendelian

inheritance. This approach was considered appropriate for generation of hypotheses prior to seeking validation in larger studies, as is now common practice with breast cancer susceptibility genes.[28]

Results

Characteristics of the study participants

Of the 260 women invited to participate, a further 7 were found to be deceased, 22 did not respond to the invitation letter and could not be contacted, and 36 declined participation (primary reason: did not want to revisit illness). The remaining 195 women provided consent, 161 women provided a blood sample, and DNA was successfully extracted from 156 samples. Of these, 120 women had available data on the SOAC outcome measure used to define lymphedema status. Twenty-two women had evidence of lymphedema between 6-18 months post-diagnosis, while 98 had no evidence of lymphedema. Demographic and disease characteristics of participants in this case-control study were comparable to the initial research sample (Table 1). Of note, the original cohort was shown to be representative of the wider Queensland breast cancer population.[9]

Results of genetic analyses

Table 2 provides an overview of the 10 genes under investigation. Not surprisingly, the more tagSNPs tested within a particular genetic locus and its flanking regions, the more likely we were to find a statistically significant trend test and/or genotype-specific ORs of a magnitude greater than 2.0 or less than 0.5. Full results for all SNPs tested are detailed in Supplementary Table 1. Only three loci revealed trend tests with $p<0.05$, and these same genes contained larger numbers of

genotype-specific ORs with extreme magnitudes that conformed with Mendelian expectations: *VEGFR3*, *VEGFR2* and *RORC* (Table 2).

Table 3 presents detailed results for *VEGFR3*, *VEGFR2* and *RORC*. Multiple elevated (or reduced) genotype-specific ORs occurred for adjacent tagSNPs in the 5' flanking region and exon/intron 1 of the three genes, where four of the five statistically significant results were found. *VEGFR3* tagSNPs rs10464063, rs307814 and rs307811 (p_{trend}=0.039 and 0.040, respectively; in high linkage disequilibrium in our sample set), and rs11960332, as well as the polymorphisms covered by these SNPs, are all located in the 5' flanking region or intron 1 of the gene. Likewise, tagSNPs rs4284267, rs12128071 and rs11801866 (p_{trend}=0.037) are situated in these regions of the *RORC* gene, and *VEGFR2* tagSNPs rs2239702 (p_{trend}=0.010) and rs7667298 tag many SNPs located in that gene's 5' flanking region, exon 1 and intron 1. Similar clustering was observed, but to a lesser extent, in the 3' flanking region of *VEGFR3* (rs10055319 and rs11739214, p_{trend}=0.020) as well as tagSNPs rs6879285, rs1565818 and rs11747066 in intron 29 (long splice variant)/3' flanking region (short splice variant). Other regions of possible interest due to clustering of results occurred in *VEGFR2* at intron 2 (rs1531290 and rs4576072) and intron 7 (rs10020464, rs17711073, rs2034965 and rs17085326) (Table 3).

Three additional genes showed minimal evidence of clustering based on the presence of two adjacent tagSNPs, sometimes close to an additional tagSNP, with genotype-specific ORs >2.0 or <0.5. These include rs11947611 and rs1485766, near rs6828869, all located deep in intron 4 of *VEGFC*; rs12089523 and rs10494972 in intron 4 of *PROX1*; and rs17318858 and rs17403620, near rs17403795, in the 3'

region of *LYVE1* (Supplementary Table 1). The remaining tagSNPs with ORs beyond the *a priori* thresholds of interest are individually scattered across the length of the tested genes.

Conclusions

The Pulling Through Study was designed to investigate the development of lymphedema through prospective follow-up of a cohort of Australian women recently treated for breast cancer. Among the women in our original cohort who presented with lymphedema according to objective assessment (n=67), around 40% were not in any of the high-risk categories, which included receiving mastectomy, 20+ nodes excised, or treated on the non-dominant side,[9, 29] suggesting that these factors were partial causes at best. This follow-up study now provides suggestive evidence for the involvement of the genes *VEGFR2*, *VEGFR3* and *RORC* in the development of secondary lymphedema following breast cancer treatment. One or more tagSNPs in each of these genes showed a statistically significant association with lymphedema, and these individual results appeared within clusters of tagSNPs exhibiting odds ratios suggestive of altered lymphedema risk. All three genes code for receptor proteins, two of which come from the same gene family, and the clusters of noteworthy findings occur in analogous gene regions with potential biological function predicted on the basis of bioinformatic analysis.

VEGF (Vascular Endothelial Growth Factor) is a key player in angiogenesis and lymphangiogenesis, and interacts with numerous proteins, including VEGFC and VEGFD and the receptors VEGFR2 and VEGFR3.[30] The VEGFC and VEGFD ligands stimulate lymphatic vessel growth[31-32] and can ameliorate secondary

lymphedema in mice.[33] VEGFC/D-induced lymphangiogenesis is mediated by *VEGFR3*, and *VEGFR3* inhibition correlates with inhibition of lymphatic development and lymphedema;[34] their increased expression is associated with metastatic disease.[35] VEGFR3 has also been shown to cooperate with VEGFR2 in lymphatic vessel sprouting.[32] In addition, a hereditary form of lymphedema called Milroy disease (also known as familial primary congenital lymphedema (PCL)) has been attributed to non-functional forms of *VEGFR3*.[36-37] It is therefore reasonable to hypothesize that more modest functional variations in these genes may predispose to secondary lymphedema.

Bioinformatic analysis of tagSNPs located in the 5' flanking region and exon/intron 1 of *VEGFR2* and *VEGFR3* that are associated with risk of lymphedema in our study revealed that these SNPs, or SNPs that are tagged by them, are predicted to have an effect on transcription factor binding sites.[38] An example is the common SNP rs10464063, located upstream of *VEGFR3*, which displayed substantially elevated ORs up to 8.29 (p_{trend}=0.053) for the rare-allele homozygote genotype. Of the nine SNPs tagged by this SNP (rs10464063 included), eight are predicted to alter transcription factor binding sites,[38] and hence may affect expression of *VEGFR3*. *VEGFR2* tagSNP rs2239702 displayed the most significant p_{trend} -value (0.010) of all SNPs tested in this study, with an increased risk of secondary lymphedema up to 6.72 for the rare-allele homozygote genotype. Rs2239702 tags over 15 SNPs and 4 of these are predicted to alter transcription factor binding sites.[38] To our knowledge, the functional effects of these SNPs have not been investigated experimentally.

The other interesting region of the *VEGFR3* gene involves three tagSNPs in intron 29 of the long splice variant or the 3' flanking region in the short version. TagSNPs rs6879285 and rs11747066 (which are highly correlated with each other, r^2=0.83) both tag rs1049095, located in the 3' untranslated region of the short splice variant of *VEGFR3*. This SNP is predicted to occur within 2 potential miRNA binding sites,[39] alteration of which could affect protein production. The third tagSNP, rs1565818, also tags the non-synonymous SNP p.R1146H (rs1130379), a common polymorphism reported not to be associated with PCL.[18] However, the amino acid substitution is located in the cytoplasmic tyrosine kinase domain of *VEGFR3,*[40] where various mutations have been found in PCL families.[18, 36-37] The functional effects of p.R1146H have not been reported, although bioinformatic analysis implies that this residue is not highly evolutionarily conserved and therefore is most likely a benign alteration;[41] nevertheless, the results from our study support further evaluation of this tagSNP and the polymorphisms it tags.

The only other statistically significant association we observed between a tagSNP and lymphedema occurred in the *RORC* gene. The functional significance of this member of the Retinoid-related Orphan Receptor family remains unexplored in human secondary lymphedema;[42] however, in mice, this gene is essential for lymphoid organogenesis.[21] Also, one of its ligands, retinoic acid, recently has been shown to modulate lymphangiogenesis *in vivo* in the mouse embryo.[43] Similar to the two *VEGF* receptor genes, *RORC* tagSNP rs11801866 is located at the 5' end of *RORC* (in intron 1 of the commonly expressed splice variant[44]). Both rs11801866 and another interesting tagSNP in this region, rs12128071, are predicted to affect transcription factor binding sites.[38]

Two other genes had more limited evidence from this study for involvement in secondary lymphedema. *PROX1* is a human homologue of the *Drosophila* homeobox gene *prospero*, expressed in lymphatic vessels of adults, and essential to maintain lymphatic endothelial cell identity.[45] However, the two tagSNPs with provocative findings located in intron 4 of *PROX1* are now recognized to be in high linkage disequilibrium with each other, and although they tag a large block of many SNPs, none are predicted to have any functional effect.[38] *LYVE1* is a marker for commencement of lymphatic development, but the cluster of tagSNPS in the *LYVE1* 3' region (rs17318858 and rs17403620, near rs17403795) also had no predicted functional effects based on current knowledge.[38]

To our knowledge, this is the first report evaluating potential genetic predisposition to secondary lymphedema following breast cancer diagnosis and treatment. It is based on rigorous follow-up of a population-based cohort of women with breast cancer to detect lymphedema based on objective measurements, followed by a systematic and thorough exploration of polymorphisms in 10 genes of potential interest to the etiology of secondary lymphedema. Bonferroni correction was not applied, as this is considered overly conservative for a hypothesis-generating study. The most significant limitation of the study relates to its small sample size. Although 67 cases had been identified via clinical assessment in the Pulling Through Study at the time of funding, attrition was much higher than anticipated between the 18-month examination and the 6-year follow-up, in part due to more difficulty obtaining blood samples from all previous participants and a higher mortality rate than expected among the women with lymphedema.[46] Also, it is plausible that some of the woman

classified as controls had developed lymphedema between the 18-month post-diagnosis assessment and 6-year blood draw, making it more difficult to find clinically and statistically significant differences between our cases and controls. Nevertheless, despite the limited statistical power, we identified two genetic loci from the same biological family, *VEGFR2* and *VEGFR3*, and a third gene, *RORC*, to have clusters of interesting results for SNPs located in analogous parts of the genes. Moreover, a number of these SNPs could potentially influence transcription factor binding sites and protein production. Many of these polymorphisms are sufficiently common to be of potential public health importance if these associations are genuine, e.g., minor allele frequencies between 20-40% for SNPs in the two *VEGFR* genes.

In summary, this research extends findings from primary congenital lymphedema and animal models to secondary lymphoedema following cancer diagnosis. The possibility that the three receptor genes identified confer genetic predisposition to secondary lymphedema following breast cancer treatment warrants further attention for potential replication using larger datasets. If confirmed, understanding the role of inherited genetic variation in lymphedema pathogenesis could lead to improvements in clinical management of breast cancer patients. First, with constant improvements in genome-wide sequencing technology and lowering of genotyping costs for clinical use, it is possible to envision identification of lymphedema molecular signature/s that could aid in the prediction and modified management of women at risk. This information could be used to target women for monitoring of lymphedema status and rapid referral to specialized care. It has been suggested that early detection of lymphedema may facilitate more effective management, resulting in reduced severity and associated disability.[47] The emergence

of new technologies in drug design and development, combined with the identification of novel molecular targets specific to the onset of lymphedema, also may enable further development of tailored therapies, both for treatment and prevention of the condition. Finally, although the findings in this study are specifically relevant to breast cancer, if confirmed, there are implications for other patients at risk of secondary lymphedema following injury, melanoma, or gynecological, prostate, or head and neck cancers.

Acknowledgements:

The original Pulling Through Study was funded by the National Breast Cancer Foundation (NBCF), while the 6-year follow-up was supported by funds from the Australian National Health and Medical Research Council (NHMRC) and Cancer Australia (Grant #497235). ABS was supported by a NHMRC Senior Research Fellowship, SCH was supported by a NBCF Early Career Development Fellowship, and MJ was supported by a NHMRC Career Development Award. This work was made possible by the women who generously gave their time to participate, the clinical support provided by Drs. Chris Pyke, John Bashford and Christobel Saunders, and the research contributions of Dr Tracey DiSipio, Sheree Rye and Tracy O'Mara.

Author Disclosure Statement:

No competing financial interests exist.

References

1. Petrek JA, Heelan MC. Incidence of breast carcinoma-related lymphedema. Cancer 1998;83:2776-2781.
2. Erickson VS, Pearson ML, Ganz PA, Adams J, Kahn KL. Arm edema in breast cancer patients. JNCI 2001;93:96-111.
3. Clark B, Sitzia J, Harlow W. Incidence and risk of arm oedema following treatment for breast cancer: a three-year follow-up study. Q J Med 2005;98:343-348.
4. Foldi E. The treatment of lymphoedema. Cancer 1998;83:2833-2834.
5. Hayes S, Rye S, Battistuta D, Newman B. Prevalence of upper-body symptoms following breast cancer and its relationship with upper-body function and lymphoedema. Lymphology 2010;43:178-187.
6. Finnane A, Hayes S, Obermair A, Janda M. Quality of life of women with lower-limb lymphoedema following gynaecological cancer. Expert Review of Pharmacoecon Outcomes Res 2011;11:287-297.
7. Petrek JA. Lymphedema in a cohort of breast carcinoma survivors 20 years after diagnosis. Cancer 2001;92:1368-1377.
8. Edwards TL. Prevalence and aetiology of lymphoedema after breast cancer treatment in southern Tasmania. Aust N Z J Surg 2000;70:412-418.
9. Hayes SC, Janda M, Cornish B, Battistutta D, Newman B. Lymphoedema following breast cancer: incidence, risk factors and effect on upper body function. J Clin Oncol 2008;26:3536-3542.
10. Williams AF, Franks PJ, Moffatt CJ. Lymphoedema: estimating the size of the problem. Palliat Med 2005;19:300-313.

11. Wigle JT, Oliver G. Prox1 function is required for the development of the murine lymphatic system. Cell 1999;98:769-778.

12. Srinivasan R, Geng X, Yang Y, Wang Y, Mukatira S, Studer M, Porto M, Lagutin O, Oliver G. The nuclear hormone receptor Coup-TFII is required for the initiation and early maintenance of Prox1 expression in lymphatic endothelial cells. Genes Dev 2010;24:696-707.

13. Karkkainen MJ, Haiko P, Sainio K, Partanen J, Taipale J, Petrova TV, Jeltsch M, Jackson D, Talikka M, Rauvala H, Betsholtz C, Alitalo K. Vascular endothelial growth factor C is required for sprouting of the first lymphatic vessels from embryonic veins. Nat Immunol 2004;5:74-80.

14. Francois M, Caprini A, Hosking B, Orsenigo F, Wilhelm D, Browne C, Paavonen K, Karnezis T, Shayan R, Downes M, Davidson T, Tutt D, Cheah KS, Stacker SA, Muscat GE, Achen MG, Dejana E, Koopman P. Sox18 induces development of the lymphatic vasculature in mice. Nature 2008;456:643-647.

15. Irrthum A, Devriendt K, Chitayat D, Matthijs G, Glade C, Steijlen PM, Fryns JP, Van Steensel MA, Vikkula M. Mutations in the transcription factor gene SOX18 underlie recessive and dominant forms of hypotrichosis-lymphedema-telangiectasia. Am J Hum Genet 2003;72:1470-1478.

16. Ferrell RE, Levinson KL, Esman JH, Kimak MA, Lawrence EC, Barmada MM, Finegold DN. Hereditary lymphedema: evidence for linkage and genetic heterogeneity. Hum Mol Genet 1998;7:2073-2078.

17. Fang J, Dagenais SL, Erickson RP, Arlt MF, Glynn MW, Gorski JL, Seaver LH, Glover TW. Mutations in FOXC2 (MFH-1), a forkhead family

transcription factor, are responsible for the hereditary lymphedema-distichiasis syndrome. Am J Hum Genet 2000;67:1382-1388.

18. Evans A, Bell R, Brice G, Comeglio P, Lipede C, Jeffery S, Mortimer P, Sarfarazi M, Child A. Identification of eight novel VEGFR-3 mutations in families with primary congenital lymphoedema J Med Genet 2003;40:697-703.

19. Ferrell R Finegold DN. Research perspectives in inherited lymphedema: An update. Ann N Y Acad Sci 2008;1131:134-139.

20. Butler MG, Isogai S, Weinstein BM. Lymphatic development. Birth Defects Res C Embryo Today 2009;87:222-231.

21. Kurebayashi S, Ueda E, Sakaue M, Patel DD, Medvedev A, Zhang F, Jetten AM. Retinoid-related orphan receptor gamma (RORgamma) is essential for lymphoid organogenesis and controls apoptosis during thymopoiesis. Proc Natl Acad Sci U S A 2000;97:10132-10137.

22. Jin D, Harada K, Ohnishi S, Yamahara K, Kangawa K, Nagaya N. Adrenomedullin induces lymphangiogenesis and ameliorates secondary lymphoedema. Cardiovasc Res 2008;80:339-345.

23. Hayes S, Cornish B, Newman B. Comparison of methods to diagnose lymphoedema among breast cancer survivors: 6-month follow-up. Breast Cancer Res Treat 2005;89:221-226.

24. Barrett J, Fry B, Maller J, Daly M. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-265.

25. Lose F, Nagle C, O'Mara T, Batra J, Bolton K, Song H, Ramus S, Gentry-Mahara A, Menon U, Gayther S, Pharoah P, Kedda M, Spurdle A. Vascular

endothelial growth factor gene polymorphisms and ovarian cancer survival. Gynecol Oncol 2010;119:479-483.

26. Amigo J, Salas A, Phillips C, Carracedo A. SPSmart: adapting population based SNP genotype databases for fast and comprehensive web access. BMC Bioinformatics 2008;9:428.

27. Johnson A, Handsaker R, Pulit S, Nizzari M, O'Donnell C, de Bakker P. SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008;24:2938-2939.

28. Ahmed S, Thomas G, Ghoussaini M, Healey CS, Humphreys MK, Platte R, Morrison J, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 2009;41:585-590.

29. Hayes SC, Janda M, Cornish B, Battistutta D, Newman B. Lymphoedema secondary to breast cancer: how choice of measure influences diagnosis, prevalence and identifiable risk factors. Lymphology 2008;41:18-28.

30. Roy H, Bhardwaj S, Yla-Herttuala S. Biology of vascular endothelial growth factors. FEBS Lett 2006;580:2879-2887.

31. Jeltsch M, Kaipainen A, Joukov V, Meng X, Lakso M, Rauvala H, Swartz M, Fukumura D, Jain RK, Alitalo K. Hyperplasia of lymphatic vessels in VEGF-C transgenic mice. Science 1997;276:1423-1425.

32. Veikkola T, Jussila L, Makinen T, Karpanen T, Jeltsch M, Petrova TV, Kubo H, Thurston G, McDonald DM, Achen MG, Stacker SA, Alitalo K. Signalling via vascular endothelial growth factor receptor-3 is sufficient for lymphangiogenesis in transgenic mice. Embo Journal 2001;20:1223-1231.

33. Rutkowski JM, Moya M, Johannes J, Goldman J, Swartz MA. Secondary lymphedema in the mouse tail: Lymphatic hyperplasia, VEGF-C upregulation, and the protective role of MMP-9. Microvasc Res 2006;72:161-171.

34. Kim H Dumont DJ. Molecular mechanisms in lymphangiogenesis: model systems and implications in human disease. Clin Genet 2003;64:282-292.

35. Tammela T Alitalo K. Lymphangiogenesis: Molecular mechanisms and future promise. Cell 2010;140:460-476.

36. Karkkainen MJ, Ferrell RE, Lawrence EC, Kimak MA, Levinson KL, McTigue MA, Alitalo K, Finegold DN. Missense mutations interfere with VEGFR-3 signalling in primary lymphoedema. Nat Genet 2000;25:153-159.

37. Irrthum A, Karkkainen M, Devriendt K, Alitalo K, Vikkula M. Congenital Hereditary Lymphedema Caused by a Mutation That Inactivates VEGFR3 Tyrosine Kinase. Am J Hum Genet 2000;67:295-301.

38. Zongli Xu Jack A. Taylor. SNPinfo: Integrating GWAS and Candidate Gene Information into Functional SNP Selection for Genetic Association Studies. Nucleic Acids Research. http://manticore.niehs.nih.gov/snpfunc.htm. 2009.

39. Hiard S, Charlier C, Coppieters W, Georges M, Baurain D. Patrocles: a database of polymorphic miRNA-mediated gene regulation in vertebrates. Nucleic Acids Res 2010;38:D640-651.

40. Galland F, Karamysheva A, Pebusque MJ, Borg JP, Rottapel R, Dubreuil P, Rosnet O, Birnbaum D. The FLT4 gene encodes a transmembrane tyrosine kinase related to the vascular endothelial growth factor receptor. Oncogene 1993;8:1233-1240.

41. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-249.

42. Medvedev A, Chistokhina A, Hirose T, Jetten AM. Genomic structure and chromosomal mapping of the nuclear orphan receptor ROR gamma (RORC) gene. Genomics 1997;46:93-102.

43. Marino D, Dabouras V, Brandli A, Detmar M. A Role for All-Trans-Retinoic Acid in the Early Steps of Lymphatic Vasculature Development. Journal of Vascul Res 2011;48:236-251.

44. He YW, Deftos ML, Ojala EW, Bevan MJ. RORgamma t, a novel isoform of an orphan receptor, negatively regulates Fas ligand expression and IL-2 production in T cells. Immunity 1998;9:797-806.

45. Johnson NC, Dillard ME, Baluk P, McDonald DM, Harvey NL, Frase SL, Oliver G. Lymphatic endothelial cell identity is reversible and its maintenance requires Prox1 activity. Genes Dev 2008;22:3282-3291.

46. Hayes S, Rye S, Round T, Lopez J, Saunders C, Pyke C, Bashford J, Battistutta D, Newman B. Prevalence and prognostic significance of secondary lymphedema following breast cancer. Lymphatic Research and Biology 2011 (in press).

47. Stout Gergich NL, Pfalzer LA, McGarvey C, Springer B, Gerber LH, Soballe P. Preoperative assessment enables the early diagnosis and successful treatment of lymphedema. Cancer 2008;112:2809-2819.

Table 1. Demographic and disease characteristics of women from the Pulling Through Study and its genetic follow-up study.

	Original Pulling Through Study cohort (n=287)	Participants in the genetics study[a] (n=120)
	n (%)[b]	n (%)[b]
Age (years)		
<50	105 (31.4)	38 (27.2)
≥50	182 (68.6)	82 (72.8)
Most extensive surgery		
CLE[c]	185 (64.9)	83 (69.6)
Mastectomy	102 (35.1)	37 (30.4)
Largest tumor size		
<16mm	171 (60.3)	78 (65.8)
16+mm	116 (39.7)	42 (34.2)
Number of nodes positive		
None removed	38 (13.1)	17 (14.1)
None positive	158 (55.9)	66 (56.6)
1-3	59 (20.1)	28 (22.2)
4+	29 (9.8)	9 (7.2)
Unavailable	3 (1.1)	0 (0.0)
Overall histologic grade		

One	76 (26.7)	35 (29.4)
Two	90 (31.7)	33 (27.5)
Three	91 (30.7)	41 (33.5)
Unavailable	30 (10.8)	11 (9.6)
Histologic type		
Infiltrating ductal	210 (72.6)	88 (73.0)
Infiltrating Lobular	44 (15.6)	14 (12.0)
Other	33 (11.7)	18 (15.0)

[a] those from the original cohort with sufficient data to calculate cumulative burden of lymphedema between 6-18 months post-diagnosis and who provided a blood sample; [b] results have been appropriately weighted (<50 years, 1.0; ≥50 years, 1.3) for oversampling of younger women; [c] CLE, complete local excision.

Table 2. Overview of genetic analyses and results for the 10 candidate genes hypothesized to influence risk of secondary lymphedema following breast cancer treatment

Gene	Gene region[a]	Size[b] (kilobases)	# tagSNPs to cover HapMap SNPs in gene region[c]	# tagSNPs tested[d]	Trend test p<0.05 (# of tagSNPs)	Genotype-specific OR[e] >2.0 or <0.5 (# of tagSNPs)	Mendelian pattern[f] of 'extreme' ORs (# tagSNPs)
VEGFC	chr4:177718895..177599691	119.2 kb	17 tagSNPs to capture 126 SNPs	17	0	5	4
VEGFR2	chr4:55939427..55996762	57.3 kb	29 tagSNPs to capture 55 SNPs	24	1	12	11
VEGFR3	chr5:180023507..180081624	58.1 kb	33 tagSNPs to capture 52 SNPs	32	3	14	13
RORC	chr1:151773548..151809348	35.8 kb	23 tagSNPs to	22	1	9	8

			capture 39 SNPs				
VEGFD	chrX:15358718..15407577	48.9 kb	7 tagSNPs to capture 36 SNPs	7	0	4	2
PROX1	chr1:214156860..214214761	57.9 kb	13 tagSNPs to capture 35 SNPs	11	0	5	2
LYVE1	chr11:10574413..10595365	20.9 kb	11 tagSNPs to capture 25 SNPs	8	0	4	3
ADM	chr11:10321642..10333924	12.3 kb	8 tagSNPs to capture 11 SNPs	7	0	1	1
SOX18	chr20:62674081..62685979	11.9 kb	3 tagSNPs to capture 3 SNPs	3 + 3SNPs[d]	0	3	0
FOXC2	chr16:86595857..86607536	11.7 kb	2 tagSNPs to capture 2 SNPs	2	0	0	0

[a] Co-ordinates from NCBI37 assembly; [b] An extra 5kb on each end of the genetic locus was included for TagSNP selection; [c] As determined by Haploview version 4.2 using HapMap data release 24/phaseII, November 2008, on NCBI B36 assembly, dbSNP 126; [d] Differences in # tagSNPs

from # SNPs tested are due to failures in design or genotyping (16 of 152 total SNPs attempted). Additional SNPs were chosen to capture genetic variation in the *SOX18* gene more completely; [e] Odds ratio (OR) reveals at least a doubling or halving of lymphedema risk for women who are heterozygous and/or homozygous for the rare allele; [f] Pattern of odds ratios across the three genotypes conforms with co-dominant, dominant or recessive mode of inheritance.

Table 3 Associations between tagSNPs[a] within *VEGFR3*, *VEGFR2* and *RORC* (odds ratio (OR) with 95% confidence intervals (CI)) and presence of lymphedema between 6- and 18-months following breast cancer surgery

Genotype	N[b]	OR	95% CI	*p*-value	Genotype	N[b]	OR	95% CI	*p*-value	Genotype	N[b]	OR	95% CI	*p*-value
VEGFR3					***VEGFR2***					***RORC***				
rs10464063 A>G					rs2239702 G>A					rs11204897 T>C				
AA	30	1.00	-		GG	61	1.00	-		TT	90	1.00	-	
AG	49	7.44	0.90,61.40	*0.06*	GA	37	3.60	1.10,11.76	0.03	TC	26	1.19	0.39,3.66	0.76
GG	27	8.29	0.93,74.05	*0.06*	AA	8	6.72	1.23,36.74	0.03	CC	1	c	c	1.00
MAF[d] = 0.49					MAF = 0.25					MAF = 0.12				
Trend test				***0.05***	*Trend test*				***0.01***	*Trend test*				***0.29***
rs307805 A>G					rs7667298 G>A					rs4845366 C>G				
AA	99	1.00	-		GG	35	1.00	-		CC	71	1.00	-	
AG	18	0.53	0.11,2.49	*0.42*	GA	59	1.78	0.52,6.08	0.36	CG	36	0.81	0.28,2.34	0.70
GG	1	c	c	*1.00*	AA	24	2.85	0.64,10.39	0.18	GG	9	0.51	0.06,4.41	0.54

Genotype	N[b]	OR	95% CI	*p*-value		N[b]	OR	95% CI	*p*-value		N[b]	OR	95% CI	*p*-value
MAF = 0.08					MAF = 0.45					MAF = 0.23				
Trend test				***0.35***	*Trend test*				***0.18***	*Trend test*				***0.50***
rs307814 G>A*					rs12502008 C>A					rs4284267 A>T				
GG	38	1.00	-		CC	51	1.00	-		AA	108	1.00	-	
GA	54	1.93	0.56,6.70	*0.30*	CA	47	1.11	0.40,3.07	0.85	AT	10	0.49	0.06,4.08	0.51
AA	17	4.64	1.10,19.50	*0.04*	AA	20	0.82	0.20,3.42	0.79	TT	0	-	-	
MAF = 0.40					MAF = 0.37					MAF = 0.04				
Trend test				***0.04***	*Trend test*				***0.87***	*Trend test*				***0.51***
rs307811 C>T*					rs1531290 C>T					rs12128071 C>T				
CC	41	1.00	-		CC	30	1.00	-		CC	97	1.00	-	
CT	47	1.62	0.44,5.98	*0.47*	CT	50	2.54	0.65,9.97	0.18	CT	11	1.05	0.21,5.28	0.96
TT	18	4.63	1.12,19.19	*0.04*	TT	30	2.25	0.51,9.99	0.29	TT	2	4.71	0.28,79.01	0.28
MAF = 0.39	14				MAF = 0.50					MAF = 0.07				

Genotype	N[b]	OR	95% CI	*p*-value	N[b]	OR	95% CI	*p*-value	N[b]	OR	95% CI	*p*-value
Trend test				***0.04***	*Trend test*			***0.32***	*Trend test*			***0.43***

rs11960332 C>T						rs4576072 A>G						rs3811417 A>G				
	CC	87	1.00	-			AA	83	1.00	-			AA	87	1.00	-	
	CT	25	0.31	0.07,1.44	*0.14*		AG	32	0.37	0.10,1.37	0.14		AG	29	0.66	0.20,2.15	0.49
	TT	5	[c]	[c]	*0.99*		GG	3	[c]	[c]	0.99		GG	2	c	c	0.99
MAF = 0.15						MAF = 0.16						MAF = 0.14					
Trend test					***0.08***	*Trend test*					***0.09***	*Trend test*					***0.36***
rs10479476 C>A						rs6837735 G>A						rs11801866 A>T					
	CC	107	1.00	-			GG	81	1.00	-			AA	94	1.00	-	
	CA	2	c	c	0.99		GA	32	1.02	0.36,2.90	0.98		AT	16	3.43	1.07,10.94	0.04
	AA	1	c	c	1.00		AA	3	c	c	0.99		TT	0		-	
MAF = 0.02						MAF = 0.16						MAF = 0.07					
Trend test					***0.99***	*Trend test*					***0.68***	*Trend test*					***0.04***
rs11748431 C>T						rs2305949 G>A						rs6685811 G>A					
	CC	59	1.00	-			GG	67	1.00	-			GG	118	1.00	-	
	CT	46	1.06	0.40,2.83	0.90		GA	43	0.67	0.24,1.93	0.46		GA	1	c	c	1.00

	TT	11	0.44	0.05,3.77	0.45		AA	8	1.39	0.25,7.66	0.71		AA	0	-	-	
MAF = 0.29						MAF = 0.25						MAF = 0.004					
Trend test					***0.63***	*Trend test*					***0.85***	*Trend test*					***1.00***
rs307806 C>T						rs7692791 A>G						rs4845604 C>T					
	CC	92	1.00	-			AA	31	1.00	-			CC	89	1.00	-	
	CT	24	0.88	0.27,2.92	0.84		AG	56	0.35	0.12,1.06	0.06		CT	29	1.71	0.61,4.75	0.31
	TT	1	c	c	1.00		GG	23	0.52	0.14,1.94	0.33		TT	0	-	-	
MAF = 0.11						MAF = 0.46						MAF = 0.12					
Trend test					***0.71***	*Trend test*					***0.21***	*Trend test*					***0.31***
rs4700966 A>G						rs2305948 G>A						rs12030974 C>T					
	AA	96	1.00	-			GG	90	1.00	-			CC	79	1.00	-	
	AG	22	1.03	0.31,3.44	0.96		GA	23	1.29	0.42,3.97	0.66		CT	33	1.03	0.36,2.97	0.95
	GG	0	-	-			AA	3	c	c	0.99		TT	6	0.93	0.10,8.58	0.95
MAF = 0.09						MAF = 0.13						MAF = 0.19					
Trend test					***0.96***	*Trend test*					***0.90***	*Trend test*					***1.00***

rs2290983 T>C						rs10020464 G>A						rs11204894 T>C					
	TT	39	1.00	-			GG	52	1.00	-			TT	70	1.00	-	
	TC	43	0.54	0.17,1.69	0.29		GA	48	0.57	0.20,1.59	0.28		TC	36	1.79	0.66,4.82	0.25
	CC	17	0.44	0.09,2.32	0.34		AA	10	0.37	0.04,3.23	0.37		CC	4	c	c	0.99
MAF = 0.39						MAF = 0.31						MAF = 0.20					
Trend test					***0.24***	*Trend test*					***0.20***	*Trend test*					***0.66***
rs3797102 T>C						rs11941492 G>A						rs11578418 C>T					
	TT	33	1.00	-			GG	67	1.00	-			CC	94	1.00	-	
	TC	60	1.63	0.47,5.59	0.44		GA	41	1.11	0.41,3.00	0.84		CT	16	2.39	0.73,7.89	0.15
	CC	25	2.29	0.57,9.20	0.24		AA	10	0.51	0.06,4.41	0.45		TT	0	-	-	
MAF = 0.47						MAF = 0.26						MAF = 0.07					
Trend test					***0.24***	*Trend test*					***0.75***	*Trend test*					***0.15***
rs307823 A>G						rs17711073 A>G						rs6693413 T>C					
	AA	81	1.00	-			AA	96	1.00	-			TT	35	1.00	-	
	AG	33	0.56	0.17,1.82	0.34		AG	20	0.45	0.10,2.11	0.31		TC	57	0.96	0.33,2.76	0.93

	GG	4	1.35	0.13,13.90	0.80		GG	1	c	c	1.00		CC	22	0.63	0.15,2.75	0.54
MAF = 0.17						MAF = 0.09						MAF = 0.44					
Trend test					***0.57***	*Trend test*					***0.27***	*Trend test*					***0.82***
rs3797104 A>G						rs2034965 C>T						rs3790515 G>A					
	AA	91	1.00	-			CC	66	1.00	-			GG	94	1.00	-	
	AG	25	0.35	0.08,1.64	0.18		CT	38	2.62	0.89,7.73	0.08		GA	22	0.40	0.09,1.84	0.24
	GG	1	c	c	1.00		TT	9	2.41	0.42,13.94	0.33		AA	2	c	c	0.99
MAF = 0.12						MAF = 0.25						MAF = 0.11					
Trend test					***0.64***	*Trend test*					***0.11***	*Trend test*					***0.18***
rs307826 A>G						rs17085326 G>A						rs1521186 C>T					
	AA	89	1.00	-			GG	96	1.00	-			CC	30	1.00	-	
	AG	28	0.99	0.33,3.00	0.99		GA	13	2.22	0.61,8.11	0.23		CT	50	1.06	0.28,3.97	0.93
	GG	1	c	c	1.00		AA	0					TT	25	1.63	0.39,6.85	0.51
MAF = 0.13						MAF = 0.06						MAF = 0.48					
Trend test					***0.86***	*Trend test*					***0.23***	*Trend test*					***0.51***

rs307827 G>A						rs7654599 A>G						rs939595 G>T					
	GG	88	1.00	-			AA	46	1.00	-			GG	51	1.00	-	
	GA	27	1.02	0.34,3.11	0.97		AG	49	1.22	0.43,3.42	0.71		GT	52	0.59	0.20,1.79	0.35
	AA	2	c	c	0.99		GG	21	0.76	0.19,3.34	0.75		TT	14	0.48	0.06,4.10	0.50
MAF = 0.13						MAF = 0.39						MAF = 0.34					
Trend test					***0.78***	*Trend test*					***0.88***	*Trend test*					***0.29***
rs2242208 C>T						rs13135562 G>C						rs10494269 G>C					
	CC	103	1.00	-			GG	107	1.00	-			GG	51	1.00	-	
	CT	7	1.89	0.34,10.52	0.47		GC	8	0.62	0.07,5.34	0.66		GC	52	1.11	0.20,1.79	0.35
	TT	0	-				CC	0	-	-			CC	14	0.78	0.06,4.10	0.50
MAF = 0.03						MAF = 0.03						MAF = 0.23					
Trend test					***0.47***	*Trend test*					***0.66***	*Trend test*					***0.29***
rs13172346 C>T						rs12505758 A>G						rs1521177 A>C					
	CC	69	1.00	-			AA	88	1.00	-			AA	32	1.00	-	
	CT	43	0.57	0.19,1.72	0.32		AG	19	0.46	0.10,2.16	0.32		AC	55	0.52	0.16,1.65	0.27

	TT	5	6.46	0.98,42,70	0.05		GG	3	c	c	0.99		CC	19	0.67	0.15,2.97	0.60
MAF = 0.23						MAF = 0.11						MAF = 0.44					
Trend test					***0.54***	*Trend test*					***0.21***	*Trend test*					***0.46***
rs6879285 T>C^						rs6838752 A>G						rs12045886 A>G					
	TT	44	1.00	-			AA	60	1.00	-			AA	61	1.00	-	
	TC	57	1.19	0.39,3.63	0.76		AG	50	1.42	0.53,3.82	0.49		AG	49	0.47	0.17,1.33	0.15
	CC	6	3.17	0.47,21.24	0.24		GG	8	1.89	0.33,10.87	0.48		GG	4	c	c	0.99
MAF = 0.32						MAF = 0.28						MAF = 0.25					
Trend test					***0.36***	*Trend test*					***0.38***	*Trend test*					***0.09***
rs17080412 G>C						rs1458831 A>G						rs3828057 G>A					
	GG	104	1.00	-			AA	89	1.00	-			GG	35	1.00	-	
	GC	6	c	c	0.99		AG	25	0.32	0.07,1.48	0.14		GA	56	3.56	0.94,13.43	0.06
	CC	0	-	-			GG	3	c	c	0.99		AA	25	2.03	0.41,10.01	0.38
MAF = 0.03						MAF = 0.13						MAF=0.46					
Trend test					***0.99***	*Trend test*					***0.10***	*Trend test*					***0.34***

rs1565818 C>G					rs7671745 C>T					rs9826 A>G				
	CC	89	1.00	-		CC	44	1.00	-		AA	55	1.00	-	
	CG	14	0.38	0.05,3.13	0.37	CT	50	1.67	0.59,4.71	0.33	AG	48	0.46	0.16,1.33	0.15
	GG	1	c	c	1.00	TT	22	0.53	0.10,2.79	0.45	GG	16	0.46	0.09,2.30	0.35
MAF = 0.08						MAF = 0.41					MAF =0.34				
Trend test					***0.32***	*Trend test*				***0.74***	*Trend test*				***0.16***
rs11747066 C>T^						rs1531289 G>A									
	CC	56	1.00	-		GG	63	1.00	-						
	CT	56	1.30	0.47,3.60	0.61	GA	37	1.70	0.62,4.68	0.30					
	TT	6	6.00	1.03,35.11	0.05	AA	8	0.76	0.08,6.84	0.80					
MAF = 0.29						MAF = 0.25									
Trend test					***0.12***	*Trend test*				***0.64***					
rs6877011 G>C						rs10008360 C>T									
	GG	97	1.00	-		CC	115	1.00	-						
	GC	12	1.69	0.41,6.93	0.47	CT	3	2.38	0.21,27.48	0.49					

	CC	1	c	c	1.00	TT	0	-	-	
MAF = 0.06						MAF = 0.01				
Trend test					***0.11***	*Trend test*				***0.49***
rs2279622 G>A						rs12642307 A>G				
	GG	99	1.00	-		AA	48	1.00	-	
	GA	18	0.97	0.25,3.70	0.96	AG	53	1.02	0.36,2.91	0.97
	AA	1	c	c	1.00	GG	16	1.67	0.43,6.51	0.46
MAF = 0.08						MAF = 0.36				
Trend test					***0.39***	*Trend test*				***0.54***
rs307822 G>A						rs10006115 C>A				
	GG	81	1.00	-		CC	111	1.00	-	
	GA	20	0.92	0.24,3.61	0.91	CA	5	1.14	0.12,10.73	0.91
	AA	4	1.74	0.17,18.09	0.64	AA	0	-	-	
MAF = 0.13						MAF = 0.02				
Trend test					***0.81***	*Trend test*				***0.91***

rs10055319 G>C							
	GG	80	1.00	-			
	GC	35	1.29	0.47,3.57	0.63		
	CC	3	2.58	0.22,30.55	0.45		
MAF = 0.17							
Trend test					***0.44***		
rs11739214 G>C							
	GG	57	1.00	-			
	GC	51	2.34	0.80,6.87	0.12		
	CC	10	5.67	1.24,25.96	0.03		
MAF = 0.30							
Trend test					***0.02***		

[a] Results for tagSNPs with no variation in our sample are not shown, including rs432475 for VEGFR3, and for RORC: rs17582155, rs7546811; [b] Sample sizes vary slightly across tagSNPs due to missing values from random technical errors; [c] Insufficient data for meaningful analysis; [d] MAF refers to minor allele frequency

*according to HapMap, rs307814 and rs307811 are in high LD but below the r^2=0.80 threshold (r^2=0.79). Our analyses reveal the two SNPs to be in high LD (r^2=0.90)

^according to HapMap, rs6879285 and rs11747066 are in LD (r^2=0.83)

Shaded boxes highlight clusters of results (in terms of physical location within the gene) that meet *a priori* criteria of interest, namely p_{trend}<0.05 or extreme odds ratio (<0.5 or >2.0); see results for further description