ORIGINAL RESEARCH

Variants in *LRRC34* reveal distinct mechanisms for predisposition to papillary thyroid carcinoma

Daniel Forrest Comiskey Jr. (ORCID) ,[1] Huiling He,[1] Sandya Liyanarachchi,[1] Mehek S Sheikh,[1] Luke K Genutis,[1] Isabella V Hendrickson,[1] Lianbo Yu,[2] Pamela L Brock,[3] Albert de la Chapelle[1]

ABSTRACT

Background Papillary thyroid carcinoma (PTC) demonstrates high heritability and a low somatic mutation burden relative to other cancers. Therefore, the genetic risk predisposing to PTC is likely due to a combination of low penetrance variants. A recent genome-wide association study revealed the association of PTC with a missense variant, rs6793295, at 3q26 in a gene called Leucine Repeat Rich Containing 34 (*LRRC34*).

Methods We report the mechanisms of PTC risk at 3q26 using a combination of overexpression, mass spectroscopy, knockdown, transcriptome profiling, migration assays and genetic analysis.

Results We observed differential binding of wild-type and missense LRRC34 to RANBP1. Overexpression of missense LRRC34 reduced RanGTP levels and increased apoptosis. We also identified a second linkage disequilibrium (LD) block upstream of *LRRC34* containing regulatory variants with allele-specific expression. Transcriptome profiling of *LRRC34* knockdown cells showed changes in genes involved with cellular movement. *LRRC34* knockdown reduced the migration of thyroid cancer cell lines. Lastly, we assessed the relative contribution of PTC risk from each locus using haplotype analysis.

Conclusions Our study demonstrates two separate mechanisms, one in G protein signalling and the other in transcriptional control, dictating PTC risk at 3q26 using both biochemical and genetic techniques.

INTRODUCTION

Thyroid cancer is the most common malignancy of the endocrine system. It is the 12th most common cancer in the USA and it is estimated that 52 070 individuals in the USA will be diagnosed with thyroid cancer in 2019.[1] Non-medullary thyroid cancer (NMTC) comprises the vast majority of thyroid cancers and arises from thyroid cells of follicular origin. Papillary thyroid carcinoma (PTC) is the most common histological subtype and represents approximately 80% of all NMTC cases.[2] The overall standard incidence ratio of PTC is increasing worldwide in contrast to most other cancers. This increase is almost entirely attributable to PTC, especially microPTC, and is believed to stem, at least in part, from improved imaging and diagnostic techniques.[3]

It has been demonstrated that the overall heritability of thyroid cancer is high compared with other cancers.[4–7] Classic linkage studies in familial papillary thyroid cancer have revealed a small number of culpable loci; however, heritable germline pathogenic variants have been largely absent from familial and sporadic cases.[8–19] Therefore, we hypothesised that the genetic risk predisposing to thyroid cancer is due to complex genetic factors and a combination of low penetrance alleles. In the past 10 years, we have collaborated with deCODE to perform three genome-wide association studies (GWASs) of differentiated thyroid cancer in individuals of European descent and uncovered 10 loci that show genome-wide significance.[20–22] The most recent meta-analysis revealed the association of NMTC with genetic variants at 1q42, 3q26, 5q22, 10q24 and 15q22. Among these five new loci, four are intronic or intergenic. However, our study uncovered association at 3q26 belonging to a missense variant, rs6793295, in the Leucine Rich Repeat Containing 34 (*LRRC34*) gene.

LRRC34 belongs to a ribonuclease inhibitor subfamily of leucine-rich repeat (LRR) proteins. LRRC34 RNA and protein expression is decreased in thyroid carcinoma[23] and its subcellular localisation is cytoplasmic, occuring in the cytoskeleton, plasma membrane and cytokinetic bridges.[24] LRRC34 is predicted to be involved in protein–protein interactions and nucleotide triphosphate binding.[25] A structurally related protein, RanGAP1, controls CRM1-related nuclear export and recycling of Ran–GTP complexes.[26]

The purpose of our work is to understand how *LRRC34* variants contribute to the genetic risk of thyroid cancer and characterise the network of genes through which it functions. Overall, our findings underscore protein–protein interaction and transcriptional control of LRRC34 in the genetic risk for PTC predisposition.

METHODS

Cell lines

TPC1, BCPAP and HEK 293T cell lines were incubated in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Thermo Fisher), 1× antibiotic-antimycotic (Thermo Fisher) and 1× Plasmocin prophylactic (InvivoGen) at 37°C in humidified air with 5% CO_2. The cell lines were obtained from Rebecca Schweppe (University of Colorado Cancer Center, Denver, CO).

Generation of plasmid constructs

LRRC34 regulatory variants were PCR amplified from genomic DNA and cloned into XhoI and EcoRV sites of the pGL4.10-E4TATA vector. *LRRC34* was cloned from genomic DNA in a multiplex PCR reaction into EcoRI and HindIII sites of the pCMV10 vector. Both missense variants were induced using the GENEART Site-Directed Mutagenesis System (Thermo Fisher). *RANBP1* was cloned from genomic DNA using a PCR reaction into HindIII and XhoI sites of the pCMV6 vector. All cloning was performed using InFusion HD (Clontech) enzyme.

Co-Immunoprecipitation and mass spectroscopy

Cells were transiently transfected with reporter plasmids using Lipofectamine 2000 (Thermo Fisher). For co-immunoprecipitation, 1.5 mg of protein input was precleared with agarose mIgG beads (Sigma) for 1 hour, then incubated with magnetic-FLAG (M2) beads (Sigma) for 4 hours at 4°C. Beads were washed four times in 1× tris-buffered saline (TBS) with Halt protease/phosphatase inhibitor (Thermo Fisher) and twice with phosphate-buffered saline (PBS). Proteins were boiled and eluted in 2× Laemmli buffer (Bio-Rad) and supplemented with 2-mercaptoethanol. Eluates were run on a Western blot as described below. For mass spectroscopy, samples were digested with trypsin and run on an Orbitrap Fusion Tribrid Mass Spectrometer using a 3-hour gradient.

Western blotting

Cells were lysed in NP-40 buffer with 1× Halt protease/phosphatase inhibitor (Thermo Fisher) and equal amounts of protein were loaded on a 4–20% TGX Stain-Free precast SDS-PAGE gel with Unstained or All Blue Protein Standards (Bio-Rad). Gels were activated under UV light using a ChemiDoc (Bio-Rad), then transferred to low-fluorescence polyvinylidene fluoride membranes using RTA transfer kit (Bio-Rad) and Trans-Blot Turbo system (Bio-Rad). Total protein was visualised under UV light using a ChemiDoc (Bio-Rad). Blots were blocked in 5% milk in tris-buffered saline with Tween 20 (TBS-T) at room temperature for 1 hour, probed with primary antibody (Ran Cell Signaling 1:500, RanGTP (AR-12), FLAG (M2) Sigma, MYC-HRP (9E10) Santa Cruz, GAPDH-HRP (0411) Santa Cruz) at 4°C overnight, washed in TBS-T, incubated in secondary antibody at room temperature for 1 hour (anti-mIgG-HRP Sigma or anti-rIgG-HRP Cell Signaling), washed in TBS-T, then subjected to Clarity ECL (Bio-Rad). Chemiluminescence was visualised using a ChemiDoc (Bio-Rad). Blots were reprobed after they were washed in TBS-T, then quenched in 15% H_2O_2 in PBS at room temperature for 15 min.

Microscopy

Cells were seeded on coverslips pretreated with 0.1 mg/mL poly-D-lysine. After 24 hours, cells were fixed in 4% paraformaldehyde and permeabilised in 0.25% Triton X-100. Cells were then blocked in 10% donkey serum and incubated with primary antibodies (LRRC34 Novus, FLAG (M2) Sigma or alpha tubulin (DM1A) Abcam) in 5% donkey serum at 4°C overnight. Cells were incubated in secondary antibody at room temperature for 1 hour in the dark (anti-mIgG Alexa Fluor 647 Thermo Fisher or anti-rIgG Alexa Fluor 488 Thermo Fisher). Coverslips were mounted with 1 drop of Diamond ProLong Antifade with 4',6-diamidino-2-phenylindole (Thermo Fisher) and cured overnight at room temperature. Cells were then imaged on an FV3000 confocal laser scanning microscope (Olympus).

Flow cytometry

Cells were washed twice with PBS and then resuspended in 1× Annexin V binding buffer (BD Biosciences). Cell suspensions were passed through a 100 μm cell strainer into polystyrene tubes and either left cells unstained or stained with FITC–Annexin V, propidium iodide or both. Cells were left incubating on shaker in darkness for 15 min and analysed on LSR II flow cytometer (BD Biosciences).

Luciferase reporter assay

For the luciferase reporter assay, cells were transiently transfected with reporter plasmids using Lipofectamine 2000 (Thermo Fisher). Each well was transfected with a luciferase reporter plasmid and Renilla plasmid pRL-TK (Promega) as an internal control for each well. Cells were lysed 24 hours after transfection with 100 μL passive lysis buffer (Promega). A 20 μL aliquot of cell lysate was assayed for luciferase activity using the GloMax 96 Microplate Luminometer (Promega).

Linkage disequilibrium (LD) and haplotype analyses

LD analysis was performed using the genotype data of 503 samples from 1000 Genomes Project (phase III), European population. Haploview V.4.2 software was applied and haplotype blocks were generated using the CI method. Haplotypes of the five selected variants were generated by using the genotyping and computer imputation data from the European descendants in an Ohio cohort of the recent GWAS.[22] The SHAPEIT V.2 program was used to estimate the haplotype frequencies in 1359 PTC cases and 1605 controls. P values and OR are provided using Fisher's exact test to compare each haplotype with the rest of the haplotypes.

Statistical analysis

Calculation of per cent splicing inclusion was determined using Image Lab V.6.0.1 (Bio-Rad). All graphical data are represented by mean±SEM and all p values reported are calculating from a two-sided Student's t-test. Gene canonical pathway analysis using *LRRC34* knockdown was performed using Ingenuity Pathway Analysis software (Qiagen). For the gene expression array analysis, a filtering method based on percentage of arrays above noise cut-off was applied to filter out low-expression genes. A linear model was employed to detect differentially-expressed genes between conditions. A variance smoothing method with moderated t-statistic was employed. The significance level was adjusted by controlling the mean number of false positives. Statistical software SAS V.9.4 and R were used for analysis.

PCR and quantitative real-time PCR assay

PCR of *LRRC34* splice isoforms and *GAPDH* was performed using AmpliTaq Gold (Thermo Fisher) under standard PCR conditions. Quantitative real-time PCR assay was performed in three technical replicates on an ABI Prism 7900 HT Sequence Detection System (Thermo Fisher). For RNA expression analysis, RNA was extracted using RNeasy (Qiagen) then treated with DNase I (Thermo Fisher). One microgram RNA was used for cDNA synthesis using High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher). All primer sets used were detected by Fast SYBR Green Master Mix kit (Thermo Fisher) using standard conditions followed by a dissociation stage. For a list of primer pairs, see online supplementary table S3.

siRNA treatment and microarray hybridisation

A non-targeting siRNA pool (Dharmacon) or siRNA pool against *LRRC34* (Dharmacon) was transfected into the cells using

Lipofectamine RNAiMAX (Thermo Fisher). After 72 hours, cells were harvested and total RNA was extracted using an RNeasy kit (Qiagen), then treated with DNase I (Thermo Fisher). The integrity of the RNA samples was assessed by BioAnalyzer (Agilent Technologies). Clariom D Human arrays (Thermo Fisher) were used to assess gene expression. In total, 100 ng RNA was used to generate the single-stranded cDNA samples for hybridisation. Then, cDNA was enzymatically fragmented and biotinylated using the WT Terminal Labeling kit (Thermo Fisher). The cDNA samples were hybridised to the array at 45°C for 16 hours. The arrays were washed and scanned with the Affymetrix GeneChip Scanner 3000 7G system (Thermo Fisher Scientific) using Affymetrix GeneChip Command Console software. Signal intensities were processed by the robust multiarray average method using Affymetrix Expression Console software.

Migration assay

Cells were transfected with a non-targeting siRNA pool or siRNA against *LRRC34* as described above. Fifty-four hours after transfection, cells were seeded in 8 µm MilliCell chambers (Millipore) for a migration assay in DMEM containing 1% FBS surrounded by DMEM containing 10% FBS. MilliCell chambers were washed in ddH$_2$O, then fixed and stained in Crystal Violet staining solution (0.05% Crystal Violet in 10% formaldehyde). MilliCell chambers were then washed in ddH$_2$O and wiped both from the top and bottom of the membrane with a Q-tip. Crystal Violet was eluted using methanol and measured in a SpectraMax M2 spectrometer (Molecular Devices) at 570 nm.

RESULTS

LRRC34 is differentially alternatively spliced

The most significant GWAS variant at 3q26 from our combined meta-analysis in PTC is rs6793295.[22] This missense variant lies within an alternatively-spliced exon (online supplementary figure S1A). Therefore, we sought to understand what levels of *LRRC34* alternative splicing were present in thyroid tissue and whether rs6793295 displayed any allele-specific alternative splicing pattern. By RT-PCR, there was very little alternative splicing of *LRRC34*, especially in thyroid tissue as compared with other tissues such as kidney and heart (online supplementary figure S1B). We then performed RT-PCR of matched-normal thyroid and tumour tissue from 16 patients with PTC (online supplementary figure S1C). We observed a small but significant increase in exon 7 inclusion in tumour over matched-normal tissue (online supplementary figure S1D). In addition, we observed no allele-specific alternative splicing pattern (online supplementary figure S1E). Overall, we conclude that the incremental changes in alternative splicing of *LRRC34* are not a significant feature in PTC due to the low magnitude of isoform switching.

LD analysis reveals a second missense variant in LD with rs6793295

In order to assess the functional significance of rs6793295 and identify other functional variants in this region, we performed LD analysis. Our analysis revealed an additional missense variant, rs10936600, in exon 8 of *LRRC34* in high LD (D′=1.0; r^2=0.89) with rs6793295 (figure 1A). In our previously-published GWAS, both markers were strongly correlated with PTC and difficult to distinguish statistically; therefore, we decided to focus on both missense variants together for functional validation. In silico predictions indicated that together, these amino acid substitutions (S246G; S286I) were deleterious
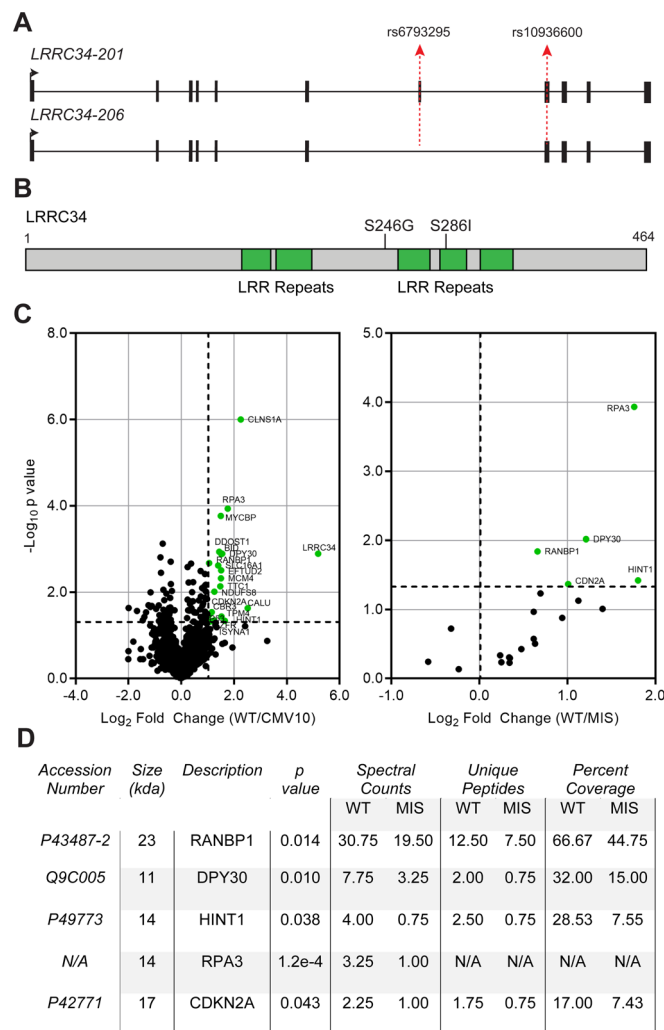

Figure 1 Mass spectrometry identifies differential binding between wild-type and missense LRRC34. (A) Schematic of protein-coding *LRRC34* transcripts. Exons are depicted as black boxes and introns as a solid line. Transcription start site and direction is indicated by a solid black arrow. The positions of the missense variants rs6793295 and rs10936600 are depicted by dotted red arrows. (B) Linear schematic of LRRC34 protein sequence with leucine-rich repeat (LRR) domains (green) and missense mutations (line). (C) Volcano plots from mass spectrometry of LRRC34 co-immunoprecipitation. The x-axes represent fold-change value and y-axes are −log$_{10}$ p values. Absolute twofold change and p value=0.05 is used as the threshold cut-off and are represented by black dotted lines. Significant proteins are labelled as green dots. (D) Table of statistically-significant targets that are bound differentially between wild-type (WT) and missense (MIS) LRRC34.

Accession Number	Size (kda)	Description	p value	Spectral Counts		Unique Peptides		Percent Coverage	
				WT	MIS	WT	MIS	WT	MIS
P43487-2	23	RANBP1	0.014	30.75	19.50	12.50	7.50	66.67	44.75
Q9C005	11	DPY30	0.010	7.75	3.25	2.00	0.75	32.00	15.00
P49773	14	HINT1	0.038	4.00	0.75	2.50	0.75	28.53	7.55
N/A	14	RPA3	1.2e-4	3.25	1.00	N/A	N/A	N/A	N/A
P42771	17	CDKN2A	0.043	2.25	1.00	1.75	0.75	17.00	7.43

to the folding of LRRC34. Specifically S286I, which lies in the middle of several LRR domains (figure 1B), was not well tolerated according to both SIFT (0.044) and PolyPhen (0.999).[27 28] As LRRC34 is predicted to be involved in mediating protein–protein interaction, we hypothesised that missense variants in LRRC34 would change the protein function through cellular localisation or affinity for its binding partners.

LRRC34 is localised to the cytoplasm

In pluripotent stem cells, LRRC34 is localised to the nucleolus.[22] However, in U2-OS, A-431 and HEK 293 cells, it is reported as being localised to cytoplasm.[24] To reconcile this discrepancy,

we performed immunofluorescence staining on a PTC-derived cell line, TPC1, using HEK 293T cells as a positive control. In concordance with data from the Protein Atlas, we observed staining of LRRC34 in the cytoplasm. Further, we observed co-localisation of LRRC34 with α-tubulin, a marker for microtubules (online supplementary figure S2A). To understand if the missense variants in LRRC34 might have an impact on its subcellular localisation, we cloned *LRRC34* into a 3xFLAG vector (wild-type LRRC34) and then induced both missense variants (S249G; L286I) together (missense LRRC34). Ectopically-expressed LRRC34 was localised to the cytoplasm along the cellular membrane in HEK 293T cells (online supplementary figure S2B). However, we did not observe differential localisation between the wild-type and missense LRRC34, indicating that LRRC34 missense mutations do not have an impact on the localisation of this protein.

LRRC34 missense variants impair protein–protein interaction
To test what effects missense variants exert on LRRC34 binding partners, we expressed either wild-type or missense LRRC34 constructs along with an empty vector negative control in HEK 293T cells (online supplementary figure S3A). We first verified equivalent levels of overexpression between both wild-type and missense LRRC34 (online supplementary figure S2B), then subjected the lysates to FLAG co-immunoprecipitation and mass spectroscopy. We discovered 20 proteins in complex with LRRC34 (figure 1C left, online supplementary table S1). Importantly, we found five proteins that were bound differentially between wild-type and missense LRRC34: RANBP1, DPY30, HINT1, RPA3 and CDKN2A (figure 1C right, D). Given the predictions of LRRC34 function from its structure, we focused on RANBP1 because this molecule also binds to RANGAP1 and it was our most abundant target next to LRRC34 itself (figure 1D, online supplementary table S1).

To confirm the findings of our mass spectroscopy screen, we designed an overexpression construct for RANBP1 (figure 2A). We co-transfected either empty vectors (C), wild-type (W) or missense (M) 3xFLAG-LRRC34 with 2xMYC-RANBP1

constructs in HEK 293T cells. We then performed a FLAG co-immunoprecipitation in the same manner as was done for the mass spectroscopy to confirm the interaction between RANBP1 and LRRC34 (figure 2B). Specifically, we observed that missense LRRC34 co-immunoprecipitated less RANBP1 than wild-type LRRC34 (figure 2C). These findings are in concordance with the mass spectroscopy data (figure 1D, online supplementary table S1) and suggest that this differential binding could have an effect on GTP signalling in the cell.

LRRC34 missense variants reduce RanGTP levels and induce apoptosis
As RANBP1 helps catalyse RanGTP, we hypothesised that a loss of binding to RANBP1 by missense LRRC34 might affect overall RanGTP levels. We overexpressed either an empty vector (C), wild-type (W) or missense (M) LRRC34 in HEK 293T cells. We performed a Western blot of cell lysates and immunoblotted using an antibody that specifically recognises the GTP-bound form of Ran (figure 3A) or total Ran (figure 3C). We observed a significant decrease in the amount of RanGTP on missense LRRC34 (LRRC34 MIS) overexpression compared with our negative control (CMV10; figure 3B). Importantly, this change in RanGTP levels was not accompanied by a change in total Ran levels between any of our overexpression conditions (figure 3D).

We then subjected LRRC34-overexpressing cells to flow cytometry and examined the proportion of cells undergoing apoptosis (figure 3E–G). We observed a small but significant increase in the proportion of cells undergoing late apoptosis (PI+/Annexin V+) in cells expressing wild-type LRRC34 (LRRC34 WT) compared to the empty vector (CMV10; figure 3H). This phenotype was exacerbated significantly on expression of missense LRRC34 (LRRC34 MIS; figure 3H).

LRRC34 upstream variants show allele-specific transcriptional activity
The region upstream of *LRRC34* shows several variants with significant allele-specific expression in normal thyroid tissue according to the GTEx Consortium (figure 4A). Likewise, our genetic association analysis in an Ohio PTC cohort uncovered a population-specific PTC locus that is genetically distinct from the missense variant, rs6793295 (figure 4B). To evaluate the allele-specific activity of variants in this LD block, we inserted eight variants from this region into a luciferase vector (figure 4C). We transfected these constructs into TPC1 (figure 4D), as well as HEK 293T and BCPAP cells (online supplementary figure S4A, B). The only variant that showed significant allele-specific expression differences in all three cell types was rs7632291. Specifically, the other allele (OA), or rs7632991 [A], showed approximately a 33% reduction in relative fluorescence units compared with the effect allele (EA) [G] between the three cell lines.

In addition, we performed qRT-PCR for *LRRC34* on unaffected thyroid tissue samples of an Ohio cohort of patients with PTC (n=122), which were also genotyped for missense and regulatory variants: rs6793295, rs10936600, rs7632991, rs1920119 and rs7628572. We report that for both missense (rs6793295 (figure 4E)) and regulatory variants (rs7632991 (figure 4E), rs1920119 (figure 4H) and rs7628572 (figure 4I)), the risk allele was associated with significantly higher expression of *LRRC34*. We did not observe any pattern of allele-specific expression for rs10936602 (figure 4F). These data were consistent with those reported by the GTEx Consortium.
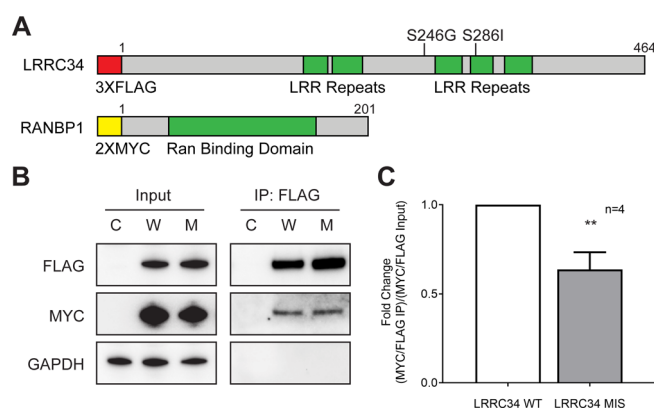

Figure 2 RANBP1 has reduced affinity for missense LRRC34. (A) Linear schematic of LRRC34 and RANBP1 exogenous protein constructs with epitope tags (red or yellow), functional domains (green) and indicated missense mutations. (B) Representative Western blot of input and co-immunoprecipitation of CMV10 and CMV6 empty vectors (C), 3XFLAG-LRRC34 WT and 2XMYC-RANBP1 (W), or 3XFLAG-LRRC34 MIS and 2XMYC-RANBP1 (M). (C) Quantitation of FLAG co-immunoprecipitation as determined by ratio of immunoprecipitated 3XFLAG-LRRC34 divided by co-immunoprecipitated 2XMYC-RANBP1 normalised to the ratio of 3XFLAG-LRRC34 and 2XMYC-RANBP1 expression from input.
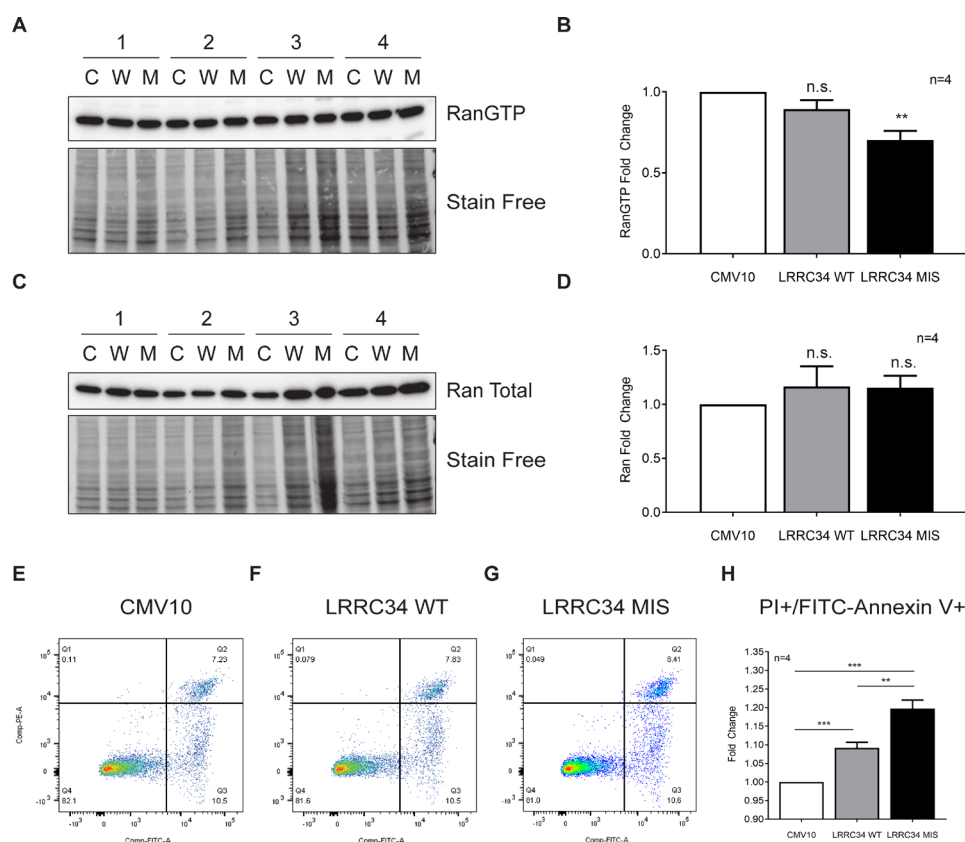
Figure 3 Missense LRRC34 reduces the expression of RanGTP and induces apoptosis. (A) Western blot of RanGTP expression from HEK 293T cells expressing either CMV10 empty vector (C), wild-type LRRC34 (W) or missense LRRC34 (M). Below each immunoblot is a representative section of the stain-free image of the total protein from 25 to 100 kDa. (B) Quantification of relative RanGTP levels were normalised to the total protein present in each lane of the stain-free image. (C) Western blot of total Ran expression from HEK 293T expressing either CMV10 empty vector (C), wild-type LRRC34 (W) or missense LRRC34 (M). Below each immunoblot is a representative section of the stain-free image of the total protein from 25 to 100 kDa. (D) Quantification of relative total Ran levels were normalised to the total protein present in each lane of the stain-free image. All results are shown as mean±SEM, n=4. **p<0.01 (two-tailed t-test). Flow cytometry histograms from HEK 293T cells transfected with (E) CMV10 empty vector (CMV10), (F) 3XFLAG-LRRC34 (LRRC34 WT) or (G) 3XFLAG-LRRC34 S249G; L286I (LRRC34 MIS) for 48 hours. Cells were stained with propidium iodide (PI) and FITC–Annexin V. (H) Quantitation of the percentage of cells in late apoptosis (Q2). Results are shown as mean±SEM, n=4. **p<0.01; ***p<0.001 (two-tailed t-test).

Minor regulatory variants abolish PTC risk from missense variants

In order to determine the relative contribution of each locus at 3q26 to PTC risk, we performed phasing of our Ohio PTC cohort (table 1). For our analysis, we chose variants on the basis of their mechanism and therefore included both missense variants (rs6793295 and rs10936600) as well as regulatory variants that demonstrated allele-specific expression in at least two cell lines (rs7632991, rs7658592 and rs1920119). Of the seven haplotypes, we observed three that were statistically significant (table 1). Haplotypes 1 and 7 are the major haplotypes, demonstrating a risk and protective effect, respectively. Interestingly, haplotype 3 bears EAs for the missense variants, but the OA for the rs7632991 regulatory variant, with an overall OR of 0.72. Thus, our haplotype, expression and luciferase data suggest that the protective effect of the OA at rs7632991 might negate or mitigate the risk of bearing both missense variants by downregulating the expression of *LRRC34*.

siRNA knockdown of *LRRC34* results in suppression of genes involved in cell motility

We determined that lower expression of *LRRC34* is associated with the protective allele [A] at rs7632991 and also the protective missense allele [C] at rs6793295. In order to characterise the

effect of reduced *LRRC34* expression, we treated BCPAP cells with non-targeting siRNA or siRNA against *LRRC34* followed by transcriptome-wide analysis of changes in gene expression. Ingenuity Pathway Analysis of our transcriptome data found that the most significant disease network was cancer (figure 5A), and top biological functions were cellular movement and cell cycle (figure 5B). We found that many of the top 10 differentially-downregulated genes on *LRRC34* knockdown were genes involved in cell movement, including *BMP2*, *CXLC5*, *LCN2* and *PTGS2* (online supplementary figure S5, online supplementary table S2). We confirmed the differential expression of molecules involved in cell motility in BCPAP cells by qPCR (figure 5D) and found concordance between our microarray and qPCR data. In addition, we performed *LRRC34* knockdown in TPC1 cells and observed that most targets had the same significant changes in gene expression (online supplementary figure S6A).

Knockdown of *LRRC34* reduces cell migration

To query whether siRNA-mediated knockdown of *LRRC34* had any effect on cell migration, we performed a transwell migration assay. We transfected BCPAP and TPC1 cells with non-targeting siRNA or siRNA against *LRRC34*, then fixed and stained cells, and quantitated the relative number of migrated versus non-migrated cells (figure 5E, online supplementary figure S6B). In
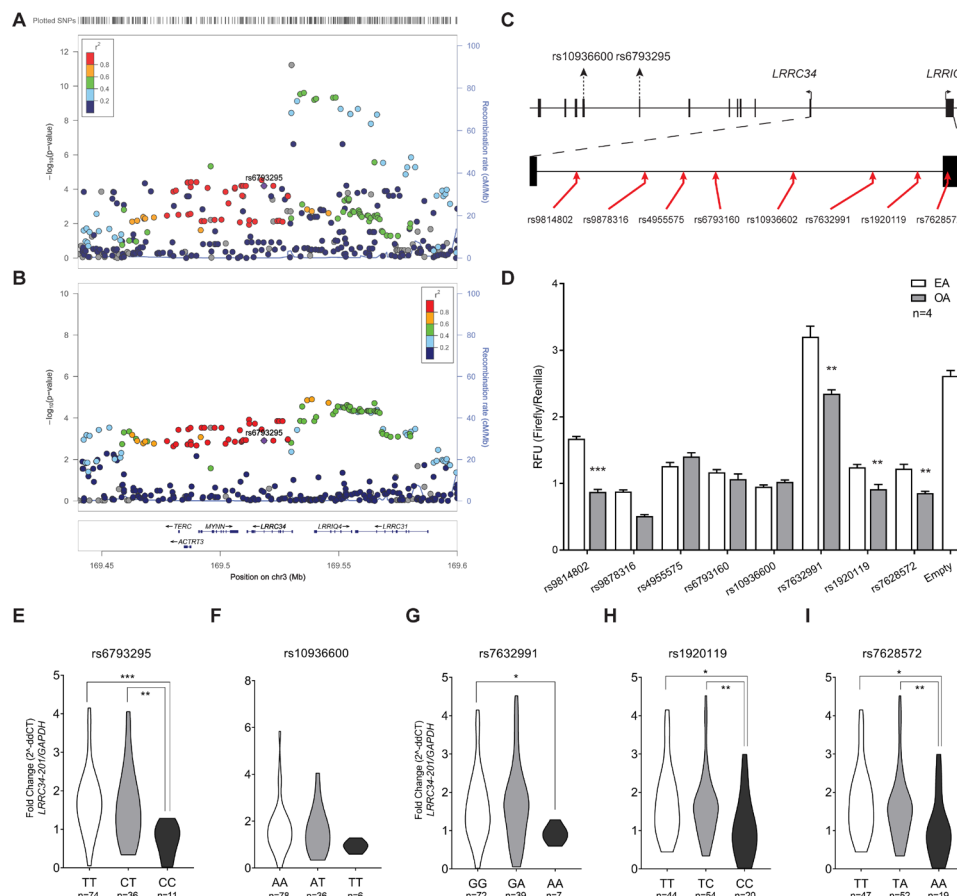
Figure 4 *LRRC34* upstream variants show allele-specific expression. (A) Graph of negative $\log_{10}$-transformed p values for Expression Quantitative Trait Loci (cis-eQTL) in thyroid tissue at the 3q26 locus. (B) Graph of negative $\log_{10}$-transformed GWAS p values from the Ohio PTC cohort at the 3q26 locus. (C) Schematic of the position of variants used in dual luciferase assay (red arrow) relative to the open reading frames for *LRRIQ4* and *LRRC34*. (D) Results of dual luciferase assay in TPC1 cells of upstream *LRRC34* variants. Results are shown as mean±SEM of one experiment with four technical replicates. For each variant, the left allele is the effect allele (EA) and the right is the other allele (OA). Asterisks indicate that experiment was repeated four times with similar results. Allele-specific expression of *LRRC34-201* in unaffected thyroid tissue between genotypes for (E) rs6793295, (F) rs10936600, (G) rs7632991, (H) rs1920119 and (I) rs7628572. **p<0.01; ***p<0.001 (two-tailed t-test).

BCPAP cells, we saw a reduction in the fold change of migrated cells with *LRRC34* siRNA knockdown as compared with the non-targeting control siRNA (figure 5F). In TPC1 cells, we observed a similar reduction in the percentage of migrated cells in *LRRC34* knockdown cells as compared with control siRNA (online supplementary figure S6C).

DISCUSSION
In the present study, we have characterised *LRRC34* variants in the predisposition to PTC. We showed that the first missense variant identified through GWAS analysis, rs6793295, is in LD with a

second missense variant, rs10936600. Furthermore, the wild-type alleles from these loci represent a risk PTC risk haplotype with an OR of 1.23. These findings allowed us to hypothesise that the missense alleles for these variants are protective against PTC, and might stem from the differential activity or binding partners for the two protein isoforms. We performed co-immunoprecipitation followed by mass spectroscopy of wild-type and missense LRRC34 isoforms and uncovered many previously unreported binding partners for LRRC34, but importantly we observed differential binding of RANBP1 to LRRC34 wild-type and missense isoforms using both mass spectroscopy and co-immunoprecipitation.

Table 1 Haplotype analysis of *LRRC34* in PTC cases vs controls

| HapID	Control	Case	OR	P value	rs10936600	rs6793295	rs7632991	rs7658592	rs1920119
Haplotype 1	0.569	0.618	1.23	4.73E−06	A	T	G	T	T
Haplotype 2	0.107	0.107	1.01	0.93	A	T	G	A	C
Haplotype 3	0.037	0.043	0.72	0.02	A	T	A	A	C
Haplotype 4	0.020	0.024	1.22	0.28	A	C	G	A	C
Haplotype 5	0.011	0.011	1.02	1.00	T	C	G	T	T
Haplotype 6	0.013	0.011	0.79	0.41	T	C	G	A	C
Haplotype 7	0.234	0.195	0.79	1.80E−04	T	C	A	A	C

PTC, papillary thyroid carcinoma.
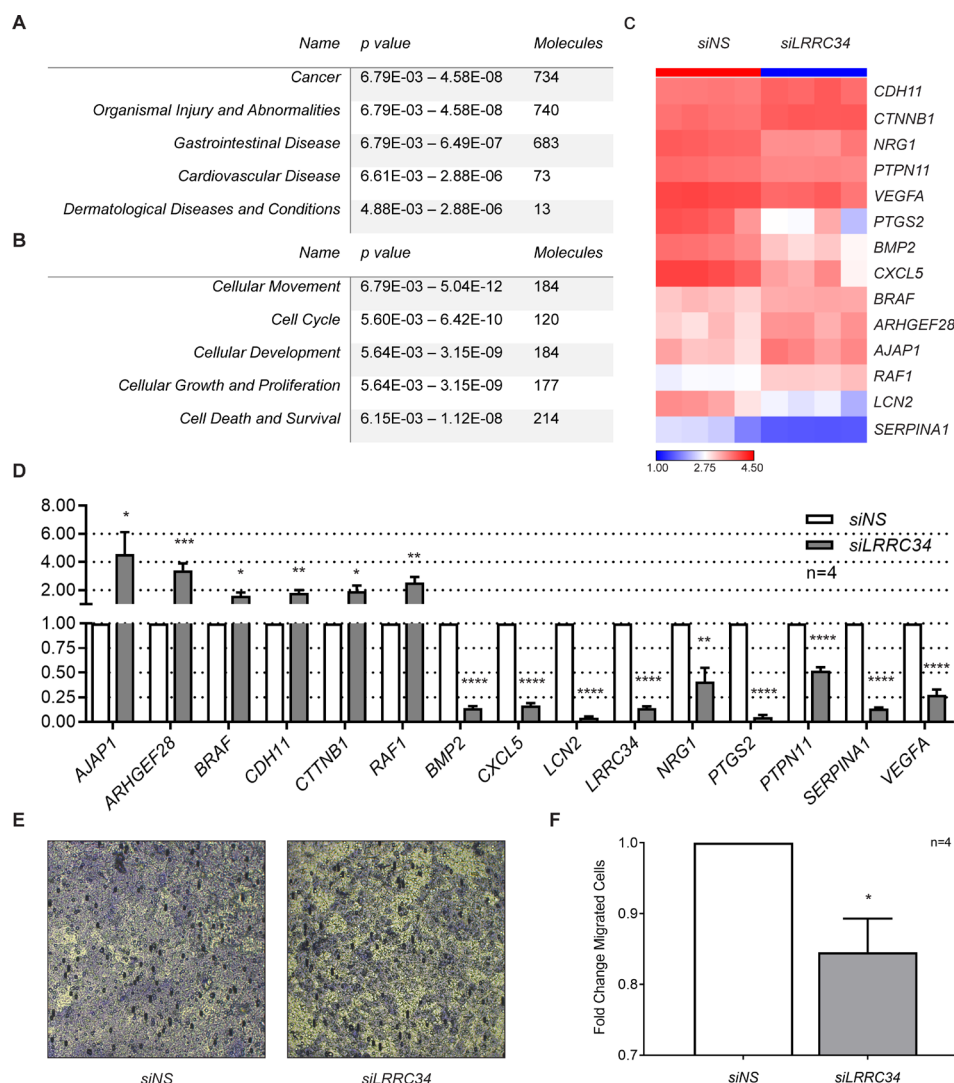
Figure 5 siRNA knockdown of *LRRC34* reduces migration of BCPAP cells. Ingenuity Pathway Analysis depicting the top (A) 'Disease and Disorders' and (B) 'Molecular and Cellular Functions' affected on *LRRC34* knockdown. (C) Expression of the candidate genes involved in cellular movement candidate from siRNA knockdown of *LRRC34* is plotted with a heat-map colour scale using relative expression fold changes (fold change >2.0, p<0.0001, n=4). (D) Quantitative real-time PCR validations for 15 targets in BCPAP cells treated with *siNS* or *siLRRC34*. (E) Representative bright-field micrographs of migrated cells in BCPAP cells. (F) Average fold change in migration *siLRRC34*-treated cells relative to *siNS* in BCPAP cells. Results are shown as mean±SEM, n=4. ****p<0.0001; ***p<0.001; **p<0.01; *p<0.05 (two-tailed t-test).

RANBP1 interacts with Ran through its RAN binding domain. It forms a complex with CRM1, RanGTP and RanGAP to allow GTP hydrolysis and nuclear export of CRM1-bound cargo. In addition to nuclear transport, RanGTP signalling plays an important role in mitotic spindle assembly.[29 30] During apoptosis, changes in RanGTP levels impact microtubule dynamics and spindle assembly through TPX2.[31] It was previously observed that treatment of cells with the topoisomerase II inhibitor, etoposide, led to the redistribution of Ran and a reduction in GTP levels.[32] Altogether, these data suggest that an important step in apoptotic signalling is breakdown in nuclear trafficking and concomitant reduction in RanGTP. Consistent with these findings, we observed that missense LRRC34 bound less to RANBP1 and its overexpression significantly reduced RanGTP levels. One hypothesis for reduced RanGTP levels associated with the missense variants in LRRC34 is that LRRC34 normally competes with RANGAP1 for the binding of RANBP1. Since missense LRRC34 has reduced affinity for RANBP1, RANBP1 and RANGAP1 together drive greater hydrolysis of RanGTP to

RanGDP. This may cause a disruption in nuclear trafficking and allow apoptosis to proceed.

Our transcriptome-wide analysis demonstrated the processes that were affected by *LRRC34* knockdown were cell motility and cell cycle; however, some molecules showed states of activation or repression that were inconsistent with the state of downstream molecules. For instance, many of the molecules that were among the most significantly downregulated on *LRRC34* knockdown, including *LCN2* and *CXCL5*, are upregulated in PTC tumours.[23] However, our genetic association data show that the other alleles in the regulatory region upstream of *LRRC34* make up a protective haplotype and our biochemical data demonstrate that these other alleles result in lower transcriptional output, reducing the migration of thyroid cancer cells. Interestingly, gnomAD reports that there is a truncating mutation in exon 9 of *LRRC34* (p.Ser300Ter) with a heterozygous carrier frequency of 7.2% in Africans, which is nearly absent in Caucasians (0.0001%). The incidence of PTC among individuals of African descent is low compared with Caucasians; therefore, it is interesting to

speculate that a common loss-of-function mutation in *LRRC34* may be responsible for mitigating the risk of PTC in Africans.[33 34]

Coding variants represent the strongest association between genotype and phenotype. However, coding variants, rare or common, can show variable penetrance.[35] One hypothesis for the wide range of phenotypic expression from a single genotype is that non-coding variants modify the levels of gene expression and therefore the penetrance of the coding variant. Recently, the Lappalainen group reported that *cis*-regulatory variants have epistatic effects on the penetrance of pathogenic coding variants on a genome-wide level.[36] In addition to the two missense variants, rs6793295 and rs10936600, we discovered a third upstream regulatory variant, rs7632991. While the former variants are in perfect LD, this regulatory variant lies within a separate LD block and therefore represents a separate association with PTC risk. Our data show that lowered expression of *LRRC34* from the minor allele at rs7632991 mitigates the increased genetic risk of PTC from the missense alleles. Therefore, we hypothesise that the upstream regulatory variants contribute to the variable penetrance of missense variants in LRRC34.

Overall, our study provides functional validation of the recently reported GWAS locus at 3q26 in the predisposition to PTC. Our genetic and biochemical data support two distinct mechanisms for the role of *LRRC34* in thyroid carcinoma, one in G protein signalling and the other in transcriptional control. As one of several low penetrance genes implicated in thyroid cancer, further studies to refine and characterise additional risk loci will enable a comprehensive understanding of the germline genetics of thyroid cancer.

ORCID iD
Daniel Forrest Comiskey Jr. http://orcid.org/0000-0002-7269-4238

REFERENCES

1 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019;69:7–34.
2 Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973–2002. *JAMA* 2006;295:2164–7.
3 Chen AY, Jemal A, Ward EM. Increasing incidence of differentiated thyroid cancer in the United States, 1988–2005. *Cancer* 2009;115:3801–7.
4 Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst* 1994;86:1600–8.
5 Dong C, Hemminki K. Modification of cancer risks in offspring by sibling and parental cancers from 2,112,616 nuclear families. *Int J Cancer* 2001;92:144–50.
6 Risch N. The genetic epidemiology of cancer: interpreting family and twin studies and their implications for molecular genetic approaches. *Cancer Epidemiol Biomarkers Prev* 2001;10:733–41.
7 Czene K, Lichtenstein P, Hemminki K. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer database. *Int J Cancer* 2002;99:260–6.
8 Peiling Yang S, Ngeow J. Familial non-medullary thyroid cancer: unraveling the genetic maze. *Endocr Relat Cancer* 2016;23:R577–95.
9 McKay JD, Lesueur F, Jonard L, Pastore A, Williamson J, Hoffman L, Burgess J, Duffield A, Papotti M, Stark M, Sobol H, Maes B, Murat A, Kääriäinen H, Bertholon-Grégoire M, Zini M, Rossing MA, Toubert ME, Bonichon F, Cavarec M, Bernard AM, Boneu A, Leprat F, Haas O, Lasset C, Schlumberger M, Canzian F, Goldgar DE, Romeo G. Localization of a susceptibility gene for familial nonmedullary thyroid carcinoma to chromosome 2q21. *Am J Hum Genet* 2001;69:440–6.
10 Canzian F, Amati P, Harach HR, Kraimps JL, Lesueur F, Barbier J, Levillain P, Romeo G, Bonneau D. A gene predisposing to familial thyroid tumors with cell oxyphilia maps to chromosome 19p13.2. *Am J Hum Genet* 1998;63:1743–8.
11 Bignell GR, Canzian F, Shayeghi M, Stark M, Shugart YY, Biggs P, Mangion J, Hamoudi R, Rosenblatt J, Buu P, Sun S, Stoffer SS, Goldgar DE, Romeo G, Houlston RS, Narod SA, Stratton MR, Foulkes WD. Familial nontoxic multinodular thyroid goiter locus maps to chromosome 14q but does not account for familial nonmedullary thyroid cancer. *Am J Hum Genet* 1997;61:1123–30.
12 Diquigiovanni C, Bergamini C, Evangelisti C, Isidori F, Vettori A, Tiso N, Argenton F, Costanzini A, Iommarini L, Anbunathan H, Pagotto U, Repaci A, Babbi G, Casadio R, Lenaz G, Rhoden KJ, Porcelli AM, Fato R, Bowcock A, Seri M, Romeo G, Bonora E. Mutant MYO1F alters the mitochondrial network and induces tumor proliferation in thyroid cancer. *Int J Cancer* 2018;143:1706–19.
13 Lesueur F, Stark M, Tocco T, Ayadi H, Delisle MJ, Goldgar DE, Schlumberger M, Romeo G, Canzian F. Genetic heterogeneity in familial nonmedullary thyroid carcinoma: exclusion of linkage to RET, MNG1, and tCo in 56 families. *J Clin Endocrinol Metab* 1999;84:2157–62.
14 Tsilchorozidou T, Vafiadou E, Yovos JG, Romeo G, McKay J, Lesueur F, Bonora E. A Greek family with a follicular variant of familial papillary thyroid carcinoma: tCo, MNG1, fPTC/PRN, and NMTC1 excluded as susceptibility loci. *Thyroid* 2005;15:1349–54.
15 Tomsic J, He H, Akagi K, Liyanarachchi S, Pan Q, Bertani B, Nagy R, Symer DE, Blencowe BJ, de la Chapelle A. A germline mutation in SRRM2, a splicing factor gene, is implicated in papillary thyroid carcinoma predisposition. *Sci Rep* 2015;5:10566.
16 He H, Bronisz A, Liyanarachchi S, Nagy R, Li W, Huang Y, Akagi K, Saji M, Kula D, Wojcicka A, Sebastian N, Wen B, Puch Z, Kalemba M, Stachlewska E, Czetwertynska M, Dlugosinska J, Dymecka K, Ploski R, Krawczyk M, Morrison PJ, Ringel MD, Kloos RT, Jazdzewski K, Symer DE, Vieland VJ, Ostrowski M, Jarząb B, de la Chapelle A. Srgap1 is a candidate gene for papillary thyroid carcinoma susceptibility. *J Clin Endocrinol Metab* 2013;98:E973–80.
17 He H, Li W, Wu D, Nagy R, Liyanarachchi S, Akagi K, Jendrzejewski J, Jiao H, Hoag K, Wen B, Srinivas M, Waidyaratne G, Wang R, Wojcicka A, Lattimer IR, Stachlewska E, Czetwertynska M, Dlugosinska J, Gierlikowski W, Ploski R, Krawczyk M, Jazdzewski K, Kere J, Symer DE, Jin V, Wang Q, de la Chapelle A. Ultra-rare mutation in long-range enhancer predisposes to thyroid carcinoma with high penetrance. *PLoS One* 2013;8:e61920.
18 He H, Nagy R, Liyanarachchi S, Jiao H, Li W, Suster S, Kere J, de la Chapelle A. A susceptibility locus for papillary thyroid carcinoma on chromosome 8q24. *Cancer Res* 2009;69:625–31.
19 Köhler A, Chen B, Gemignani F, Elisei R, Romei C, Figlioli G, Cipollini M, Cristaudo A, Bambi F, Hoffmann P, Herms S, Kalemba M, Kula D, Harris S, Broderick P, Houlston R, Pastor S, Marcos R, Velázquez A, Jarzab B, Hemminki K, Landi S, Försti A. Genome-wide association study on differentiated thyroid cancer. *J Clin Endocrinol Metab* 2013;98:E1674–81.
20 Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Masson G, He H, Jonasdottir A, Sigurdsson A, Stacey SN, Johannsdottir H, Helgadottir HT, Li W, Nagy R, Ringel MD, Kloos RT, de Visser MCH, Plantinga TS, den Heijer M, Aguillo E, Panadero A, Prats E, Garcia-Castaño A, De Juan A, Rivera F, Walters GB, Bjarnason H, Tryggvadottir L, Eyjolfsson GI, Bjornsdottir US, Holm H, Olafsson I, Kristjansson K, Kristvinsson H, Magnusson OT, Thorleifsson G, Gulcher JR, Kong A, Kiemeney LALM, Jonsson T, Hjartarson H, Mayordomo JI, Netea-Maier RT, de la Chapelle A, Hrafnkelsson J, Thorsteinsdottir U, Rafnar T, Stefansson K. Discovery of common variants associated with low TSH levels and thyroid cancer risk. *Nat Genet* 2012;44:319–22.
21 Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Sigurdsson A, Bergthorsson JT, He H, Blondal T, Geller F, Jakobsdottir M, Magnusdottir DN, Matthiasdottir S, Stacey SN, Skarphedinsson OB, Helgadottir H, Li W, Nagy R, Aguillo E, Faure E, Prats E, Saez B, Martinez M, Eyjolfsson GI, Bjornsdottir US, Holm H, Kristjansson K, Frigge ML, Kristvinsson H, Gulcher JR, Jonsson T, Rafnar T, Hjartarson H, Mayordomo JI, de la Chapelle A, Hrafnkelsson J, Thorsteinsdottir U, Kong A, Stefansson K. Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nat Genet* 2009;41:460–4.
22 Gudmundsson J, Thorleifsson G, Sigurdsson JK, Stefansdottir L, Jonasson JG, Gudjonsson SA, Gudbjartsson DF, Masson G, Johannsdottir H, Halldorsson GH, Stacey SN, Helgason H, Sulem P, Senter L, He H, Liyanarachchi S, Ringel MD, Aguillo E, Panadero A, Prats E, Garcia-Castaño A, De Juan A, Rivera F, Xu L, Kiemeney LA, Eyjolfsson GI, Sigurdardottir O, Olafsson I, Kristvinsson H, Netea-Maier RT, Jonsson T, Mayordomo JI, Plantinga TS, Hjartarson H, Hrafnkelsson J, Sturgis EM, Thorsteinsdottir U, Rafnar T, de la Chapelle A, Stefansson K. A genome-wide association study yields five novel thyroid cancer risk loci. *Nat Commun* 2017;8:14517.

23 Agrawal N, Akbani R, Aksoy BA, Ally A, Arachchi H, Asa SL, Auman JT, Balasundaram M, Balu S, Baylin SB, Behera M, Bernard B, Beroukhim R, Bishop JA, Black AD, Bodenheimer T, Boice L, Bootwalla MS, Bowen J, Bowlby R, Bristow CA, Brookens R, Brooks D, Bryant R, Buda E, Butterfield YSN, Carling T, Carlsen R, Carter SL, Carty SE, Chan TA, Chen AY, Cherniack AD, Cheung D, Chin L, Cho J, Chu A, Chuah E, Cibulskis K, Ciriello G, Clarke A, Clayman GL, Cope L, Copland D, Covington K, Danilova L, Davidsen T, Demchok JA, DiCara D, Dhalla N, Dhir R, Dookran SS, Dresdner G, Eldridge J, Eley G, El-Naggar AK, Eng S, Fagin JA, Fennell T, Ferris RL, Fisher S, Frazer S, Frick J, Gabriel SB, Ganly I, Gao J, Garraway LA, Gastier-Foster JM, Getz G, Gehlenborg N, Ghossein R, Gibbs RA, Giordano TJ, Gomez-Hernandez K, Grimsby J, Gross B, Guin R, Hadjipanayis A, Harper HA, Hayes DN, Heiman DI, Herman JG, Hoadley KA, Hofree M, Holt RA, Hoyle AP, Huang FW, Huang M, Hutter CM, Ideker T, Iype L, Jacobsen A, Jefferys SR, Jones CD, Jones SJM, Kasaian K, Kebebew E, Khuri FR, Kim J, Kramer R, Kreisberg R, Kucherlapati R, Kwiatkowski DJ, Ladanyi M, Lai PH, Laird PW, Lander E, Lawrence MS, Lee D, Lee E, Lee S, Lee W, Leraas KM, Lichtenberg TM, Lichtenstein L, Lin P, Ling S, Liu J, Liu W, Liu Y, LiVolsi VA, Lu Y, Ma Y, Mahadeshwar HS, Marra MA, Mayo M, McFadden DG, Meng S, Meyerson M, Mieczkowski PA, Miller M, Mills G, Moore RA, Mose LE, Mungall AJ, Murray BA, Nikiforov YE, Noble MS, Ojesina AI, Owonikoko TK, Ozenberger BA, Pantazi A, Parfenov M, Park PJ, Parker JS, Paull EO, Pedamallu CS, Perou CM, Prins JF, Protopopov A, Ramalingam SS, Ramirez NC, Ramirez R, Raphael BJ, Rathmell WK, Ren X, Reynolds SM, Rheinbay E, Ringel MD, Rivera M, Roach J, Robertson AG, Rosenberg MW, Rosenthal M, Sadeghi S, Saksena G, Sander C, Santoso N, Schein JE, Schultz N, Schumacher SE, Seethala RR, Seidman J, Senbabaoglu Y, Seth S, Sharpe S, Shaw KRM, Shen JP, Shen R, Sherman S, Sheth M, Shi Y, Shmulevich I, Sica GL, Simons JV, Sinha R, Sipahimalani P, Smallridge RC, Sofia HJ, Soloway MG, Song X, Sougnez C, Stewart C, Stojanov P, Stuart JM, Sumer SO, Sun Y, Tabak B, Tam A, Tan D, Tang J, Tarnuzzer R, Taylor BS, Thiessen N, Thorne L, Thorsson V, Tuttle RM, Umbricht CB, Van Den Berg DJ, Vandin F, Veluvolu U, Verhaak RGW, Vinco M, Voet D, Walter V, Wang Z, Waring S, Weinberger PM, Weinhold N, Weinstein JN, Weisenberger DJ, Wheeler D, Wilkerson MD, Wilson J, Williams M, Winer DA, Wise L, Wu J, Xi L, Xu AW, Yang L, Yang L, Zack TI, Zeiger MA, Zeng D, Zenklusen JC, Zhao N, Zhang H, Zhang J, Zhang Jiashan (Julia), Zhang W, Zmuda E, Zou L. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* 2014;159:676–90.

24 Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Åsa, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA-K, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist P-H, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Pontén F, Proteomics PF. Proteomics. tissue-based map of the human proteome. *Science* 2015;347:1260419.

25 Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Hönigschmid P, Schafferhans A, Roos M, Bernhofer M, Richter L, Ashkenazy H, Punta M, Schlessinger A, Bromberg Y, Schneider R, Vriend G, Sander C, Ben-Tal N, Rost B. PredictProtein—an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res* 2014;42:W337–43.

26 Ritterhoff T, Das H, Hofhaus G, Schröder RR, Flotho A, Melchior F. The RanBP2/RanGAP1*SUMO1/Ubc9 SUMO E3 ligase is a disassembly machine for CRM1-dependent nuclear export complexes. *Nat Commun* 2016;7:11482.

27 Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 2012;40:W452–7.

28 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;7:248–9.

29 Gruss OJ, Vernos I. The mechanism of spindle assembly: functions of Ran and its target TPX2. *J Cell Biol* 2004;166:949–55.

30 Quimby BB, Dasso M. The small GTPase Ran: interpreting the signs. *Curr Opin Cell Biol* 2003;15:338–44.

31 Moss DK, Betin VM, Malesinski SD, Lane JD. A novel role for microtubules in apoptotic chromatin dynamics and cellular fragmentation. *J Cell Sci* 2006;119:2362–74.

32 Wong C-H, Chan H, Ho C-Y, Lai S-K, Chan K-S, Koh C-G, Li H-Y. Apoptotic histone modification inhibits nuclear transport by regulating RCC1. *Nat Cell Biol* 2009;11:36–45.

33 Miller BA, Kolonel LN. *National Cancer Institute (U.S.). Cancer Control Research Program. Racial/ethnic patterns of cancer in the United States, 1988–1992*. Washington, D.C.: U.S. Department of Health and Human Services, National Institute of Health, 1996.

34 Weeks KS, Kahl AR, Lynch CF, Charlton ME. Racial/ethnic differences in thyroid cancer incidence in the United States, 2007–2014. *Cancer* 2018;124:1483–91.

35 Chen R, Shi L, Hakenberg J, Naughton B, Sklar P, Zhang J, Zhou H, Tian L, Prakash O, Lemire M, Sleiman P, Cheng W-Y, Chen W, Shah H, Shen Y, Fromer M, Omberg L, Deardorff MA, Zackai E, Bobe JR, Levin E, Hudson TJ, Groop L, Wang J, Hakonarson H, Wojcicki A, Diaz GA, Edelmann L, Schadt EE, Friend SH. Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. *Nat Biotechnol* 2016;34:531–8.

36 Castel SE, Cervera A, Mohammadi P, Aguet F, Reverter F, Wolman A, Guigo R, Iossifov I, Vasileva A, Lappalainen T. Modified penetrance of coding variants by cis-regulatory variation contributes to disease risk. *Nat Genet* 2018;50:1327–34.
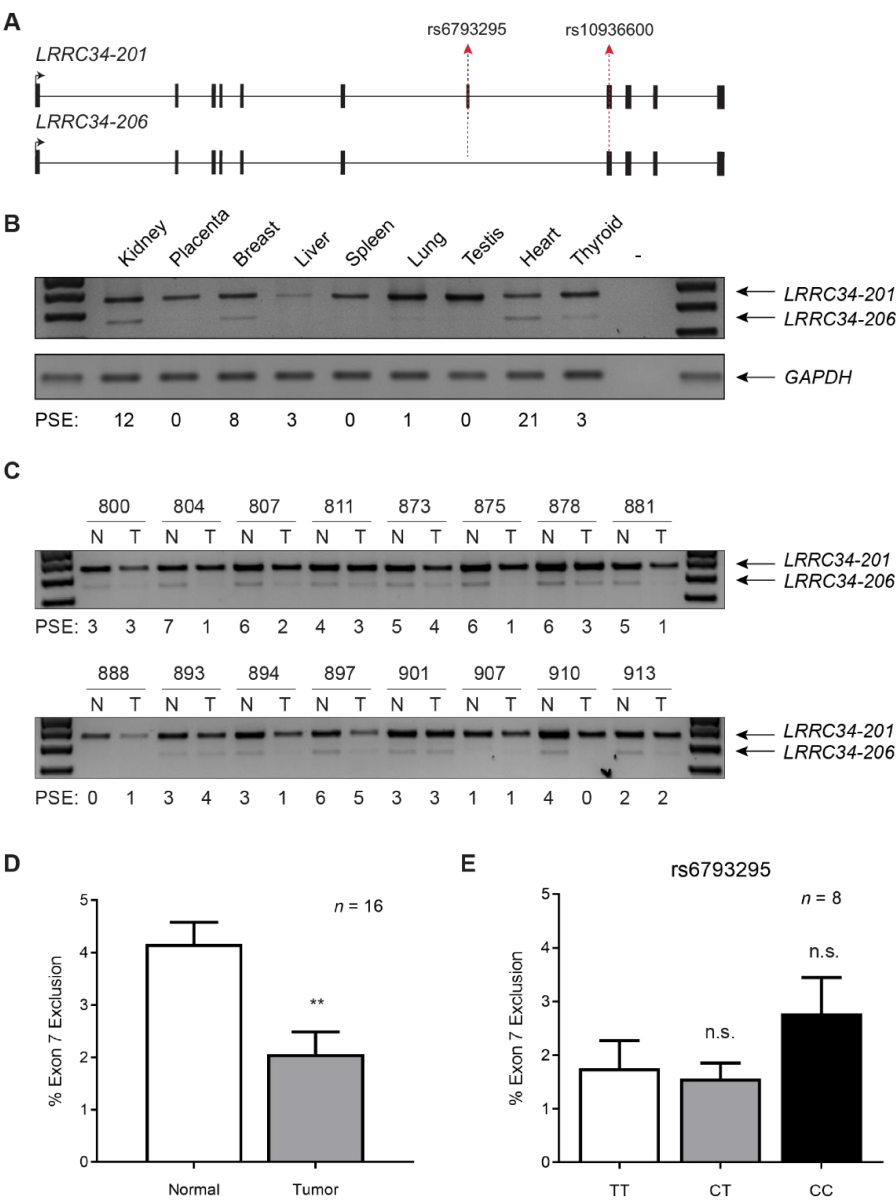
Supplementary Figure S1: *LRRC34* **shows tumor-specific alternative splicing. A.** Schematic of protein-coding LRRC34 transcripts. Exons are depicted as black boxes and introns as a solid line. Transcription start site and direction is indicated by a solid black arrow. The positions of the missense variants rs6793295 and rs10936600 are depicted by dotted red arrows. **B.** RT-PCR of *LRRC34* depicting alternative splicing of *LRRC34* exon 7 in a panel of tissues. **C.** RT-PCR of *LRRC34* depicting alternative splicing of *LRRC34* exon 7 between matched unaffected and tumor tissue from the thyroids of PTC patients. Percent splicing exclusion (PSE), or relative abundance of *LRRC34-206*, was calculated for each sample. **D.** Quantitation of the PSE between matched unaffected and tumor tissue from the thyroids of PTC patients. Results are shown as mean ± SEM, *n* = 16. **E.** Alternative splicing of *LRRC34* from eight normal thyroid samples of each rs6793295 genotype was determined by RT-PCR. Results are shown as mean ± SEM, *n* = 8. **$P < 0.01$ (Two tailed t-test).

1

A

| DAPI | LRRC34 | α-Tubulin | Merge |

B

| | DAPI | FLAG | Merge |

LRRC34 WT

LRRC34 MIS
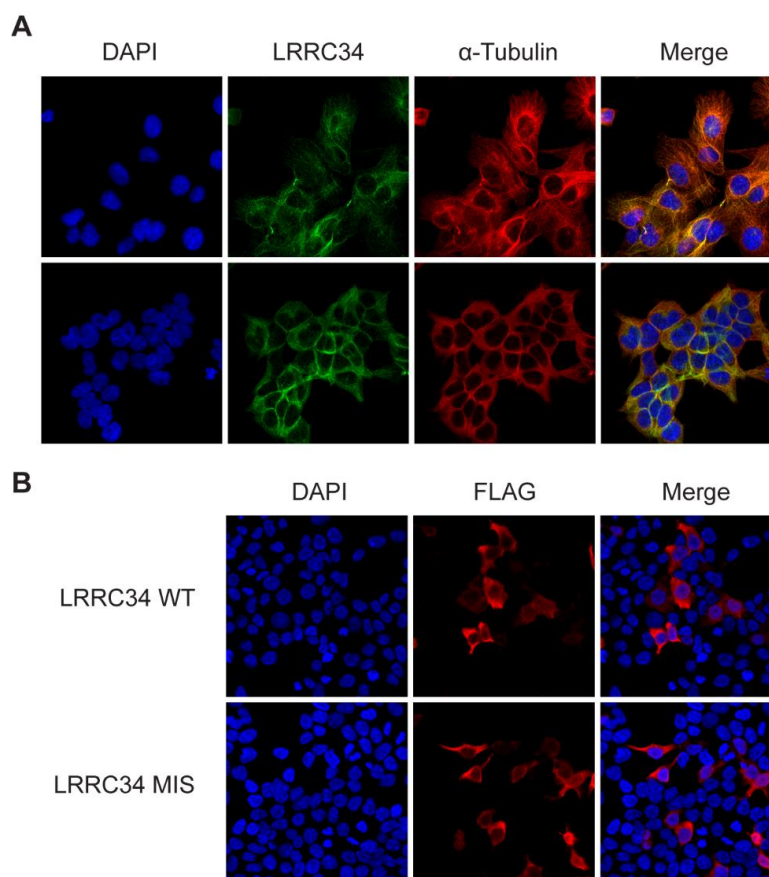
Supplementary Figure S2: LRRC34 is localized to the cytoplasm along microtubules and cell membranes. A. Immunofluorescence of endogenous LRRC34 (green) and α-tubulin (red) in TPC1 (top) and HEK 293T (bottom) cells (60X). Blue color represents staining of cell nuclei with DAPI. **B.** Immunofluorescence of HEK 293T ectopically expressing FLAG-tagged LRRC34 wild-type (top) and missense (bottom) constructs (red, 40X). Blue color represents staining of cell nuclei with DAPI.

2
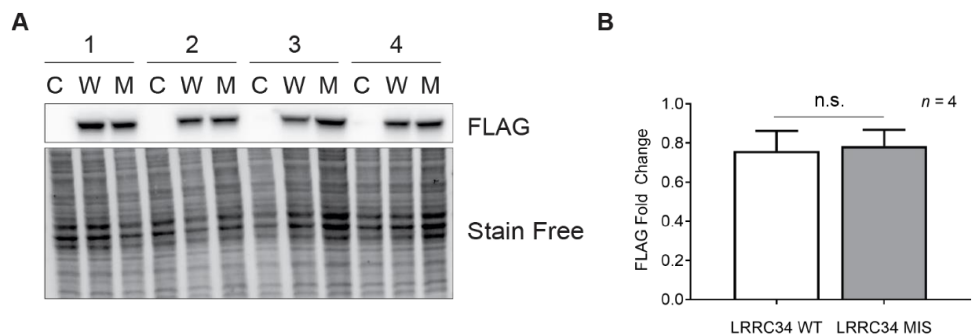
Supplementary Figure S3: Wild-type and missense LRRC34 show equivalent levels of expression.
A. Western blot of FLAG expression from HEK 293T expressing either CMV10 empty vector (C), wild-type LRRC34 (W), or missense LRRC34 (M). Below each immunoblot is a representative section of the stain-free image of the total protein from 25 to 100 kilodaltons. **B.** Quantification of relative FLAG levels were normalized to the total protein present in each lane of the stain-free image. Results are shown as mean ± SEM, *n* = 4.

3
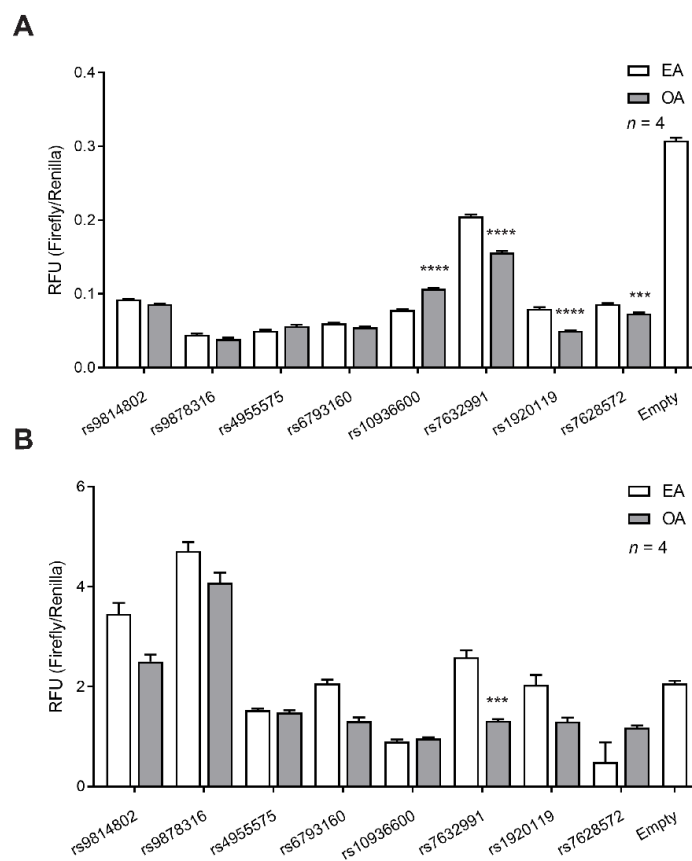
Supplementary Figure S4: *LRRC34* **upstream variants show allele-specific expression.** Results of dual luciferase assay of *LRRC34* upstream variants in **A.** BCPAP and **B.** HEK 293T cells. Results are shown as mean ± SEM of one experiment with four technical replicates. For each variant, the left allele is the effect allele (EA) and the right is the other allele (OA). Asterisks indicate that experiment was repeated four times with similar results. ***$P < 0.001$; ****$P < 0.0001$ (Two tailed t-test).

4

Supplementary Figure S5: Ingenuity Pathway Analysis (IPA) of *LRRC34* knockdown indicates changes in cellular motility. The IPA top network changed upon knockdown of *LRRC34* is "cell movement of tumor cell lines" (*P* = 5.04 x 10^{-12}). Although the pathway shows mild activation (Z-score = 1.335), three molecules located in the extracellular space are represented among the top downregulated genes (see Table S1).

Supplementary Figure S6: siRNA knockdown of *LRRC34* reduces migration of TPC1 cells. A. Quantitative real-time polymerase chain reaction (qPCR) validations for 15 targets in TPC1 cells treated with siNS or siLRRC34. **B.** Representative bright field micrographs of migrated cells in TPC1 cells. **C.** Average fold change in migration siLRRC34-treated cells relative to siNS in TPC1 cells. Results are shown as mean ± SEM, *n* = 4. ****$P < 0.0001$; ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$ (Two tailed t-test).

6

Supplementary Table S1. Mass Spectroscopy of LRRC34 FLAG immunoprecipitation

Accession Number	Size (kda)	Description	P value	Spectral Counts		Unique Peptides		Percent Coverage	
				WT	MIS	WT	MIS	WT	MIS
Q8IZ02-2	51	LRRC34	0.570	121.50	182.00	30.00	19.50	61.25	46.00
P43487	23	RANBP1	0.014	30.75	19.50	12.50	7.50	66.67	44.75
Q15029	116	EFTUD2	0.510	9.00	3.50	5.00	5.00	N/A	N/A
P53985	54	SLC16A1	0.190	7.25	9.00	4.25	4.50	9.25	9.25
Q9C005	11	DPY30	0.010	7.75	3.50	2.00	0.75	32.00	15.00
Q996114	34	TTC1	0.460	5.00	4.50	3.75	1.75	20.50	14.30
P49773	14	HINT1	0.038	4.00	1.00	2.50	0.75	28.53	7.55
O75828	31	CBR3	0.270	5.50	3.75	1.00	0.50	16.50	9.00
P54105	26	CLNS1A	0.110	5.25	3.50	3.75	2.50	27.00	18.53
N/A	48	DDOST	0.490	3.25	2.50	2.50	2.00	N/A	N/A
O43852	37	CALU	0.098	8.50	3.00	6.00	2.50	24.25	11.03
Q9NPH2-3	55	ISYNA1	0.130	4.00	1.50	2.00	0.50	6.73	2.03
P55957-2	27	BID	0.580	3.75	3.25	2.75	1.75	10.55	8.70
N/A	14	RPA3	1.2E-4	3.25	1.00	2.00	0.25	N/A	N/A
P67936	29	TPM4	0.075	6.50	3.00	1.50	0.75	17.50	5.50
P33991	97	MCM4	0.380	3.00	2.25	2.25	1.50	3.50	2.13
N/A	103	MYCBP	0.058	3.25	1.75	2.00	1.00	N/A	N/A
N/A	210	NDUFS8	0.310	2.75	1.75	2.50	1.25	N/A	N/A
P04637	44	P53	0.590	2.25	1.75	1.75	1.25	9.15	5.20
Q96KR1	117	ZFR	0.730	2.50	2.75	1.75	2.00	2.75	3.03
P42771	17	CDKN2A	0.043	2.25	1.00	2.25	0.75	17.00	7.43

Supplementary Table S2: Top ten differentially-expressed genes upon *LRRC34* knockdown.

Gene	Fold Change	P value
TMEM97	25.91	7.01E-08
TMEM45A	19.05	2.37E-09
TMCO1	11.34	2.58E-10
CRHBP	11.28	2.92E-11
PLA2G16	10.63	6.74E-10
GEMIN5	8.91	2.96E-07
SH3D19	8.75	5.22E-09
BMP4	8.72	1.08E-05
VANGL1	7.94	2.30E-09
FOXM1	7.59	6.17E-06
ITGB8	-18.86	1.26E-07
LCN2	-22.12	9.50E-06
CYTIP	-23.33	2.58E-08
APOD	-23.49	1.71E-07
BMP2	-24.55	2.91E-11
ADGRF1	-25.25	1.83E-10
SAA1	-26.91	2.72E-06
SAA2	-38.14	2.41E-05
CXCL5	-62.71	4.61E-06
PTGS2	-117.46	5.82E-05

Supplementary Table S3: Primer pairs.

Name	Sequence
LRRC34-201 208 F1	gtacagtgaacaggaagagtc
LRRC34-201 208 R1	tacatcagccaaatacacc
LRRC34-206 F1	actgtgatctggaagagtctaca
LRRC34-206 R1	ttgcgcctgcattttctatt
LRRC34 X6 F1	atcggactctgaaatacctaagaa
LRRC34 X8 R1	agcgcaggctactgttcag
GAPDH F1	ggcatggactgtggtcatgag
GAPDH R1	tcttctgggtggcagtgatg
LRRC34 X1 F1	cagccgccgcggccagtgggtgagagg
LRRC34 X9 R1	cttccagggtagtgttgcttttc
LRRC34 X8 F1	tgtaggccgcatgttgaaag
p3xFLAG-CMV10-LRRC34 EcoRI F	ttgcggccgcgaattctatggcagcgcagccgccg
p3xFLAG-CMV10-LRRC34 EcoRI R	atctatcgatgaattctctcatggctgttgacctac
LRRC34_861_FW	cgacctatactgtacggtgaacaggaagagt
LRRC34_861_RV	actcttcctgttcaccgtacagtataggtcg
LRRC34_972_FW	agtggtatacaacagatatgtgatgcactgt
LRRC34_972_RV	acagtgcatcacatatctgttgtataccact
2XMYC-RANBP1 Hind III F	gccagatctcaagctatggaacaaaaactcatctcagaagaggatctggaacaaaaactcatctcagaagaggatctggcggccgccaaggac
2XMYC-RANBP1 Smal R	acagggatgccacccttgcttctcctcagc
rs4955575 F2 EcoRV	aaaaaagatcttgattttgtgatttcttttcttctactt
rs4955575 R2 XhoI	gctcgctagcctcgataatcccagctactcagg
rs4955575 A<C F	ccaggaggcagaggtggcagtgagctgagat
rs4955575 A<C R	atctcagctcactgccacctctgcctcctgg
rs7632991 F2 EcoRV	aaaaaagatcttgatctgggttgggctttagttcattca
rs7632991 R2 XhoI	gctcgctagcctcgacattagtctttggctggtta
rs7632991 G<A F	agaacagagacaggttctgaagtttcatggt
rs7632991 G<A R	accatgaaacttcagaacctgtctctgttct
rs10936602 F1 EcoRI	aaaaaagatcttgattcaccagtttttagaaggaga
rs10936602 R1 XhoI	gctcgctagcctcgaatgcacaaactagataggtaagac
rs10936602 T<C F	tctctctgttagcatgaaaacccatttctag
rs10936602 T<C R	ctagaaatgggttttcatgctaacagagaga
rs9878316 F1 EcoRV	aaaaaagatcttgatccagtttattgagggcttttta
rs9878316 R1 XhoI	agtggcattcatcttaggaagtggcattcatcttagga
rs9878316 T<A F	aatcaatgtgatacatcctatcaacagaatg
rs9878316 T<A R	cattctgttgataggatgtatcacattgatt
rs1920119 F2 EcoRV	aaaaaagatcttgataaataggtgccatcccagaagaga
rs1920119 R2 XhoI	gctcgctagcctcgacgggatggagtggaggaggttgtg
rs7628572 F2 EcoRV	aaaaaagatcttgatatagcaagtctggtggcaggtaaa
rs7628572 R2 XhoI	gctcgctagcctcgattctatcattgacgtgctgtggat
rs6793160 F1 EcoRV	aaaaaagatcttgattattctgcagctgttggatga
rs6793160 R1 XhoI	gctcgctagcctcgaacaataatggctgttgatttcag

9

rs9814802 F1 EcoRV	aaaaaagatcttgatctactagggtcttggggtgtttg
rs9814802 R1 XhoI	gctcgctagcctcgacccggctattctggaggctgag
GAPDH F3	tgcaccaccaactgcttagc
GAPDH R3	ggcatggactgtggtcatgag
RAF1 F4	ctcggattgggtcaggctcttttg
RAF1 R2	tttgcgcagaacagccacctcatt
ARHGEF28 F1	aagcgtgtcaaaagcctggtggtt
ARHGEF28 R1	atgtcgaggggcggcttggatttttt
BMP2 F4	gacactgagacgctgttcc
BMP2 R4	ccatggtcgacctttagg
BRAF F1	aattttatggtgggacgaggat
BRAF R1	ccgattcaaggagggttctg
CDH11 F4	ctttgcagcagaaatccaca
CDH11 R4	tgggagcattatcgttgaca
CTNNB1 F2	gcgtggacaatggctactca
CTNNB1 R2	aaaatccctgttcccactcatac
CXCL5 F1	gagagctgcgttgcgtttgtttac
CXCL5 R1	ccgttcttcagggaggctaccact
LCN2 F2	gaagacaaagacccgcaaaagatg
LCN2 R2	cgggctggcaacctggaacaaaag
NRG1 F4	atctgcatcgccctccttgtggtc
NRG1 R4	gggggtgggttaggatggtgagg
PTGS2 F3	catccccttctgcctgaca
PTGS2 R3	catccccttctgcctgaca
PTPN11 F4	cggagggcgggaggaac
PTPN11 R4	tctcttttaattgcccgtgatgtt
RAF1 F4	ctcggattgggtcaggctcttttg
RAF1 R2	tttgcgcagaacagccacctcatt
SERPINA1 F2	ccgtgaaggtgcctatgatgaagc
SERPINA1 R2	gtgatgcccagttgacccaggac
VEGFA F1	atccgcagacgtgtaaatgttcct
VEGFA R1	gtggtggcggcagcgtggtt

10