CASC15-S Is a Tumor Suppressor IncRNA at the 6p22 Neuroblastoma Susceptibility Locus

Mike R. Russell1, Annalise Penikis1, Derek A. Oldridge1, Juan R. Alvarez-Dominguez2,3, Lee McDaniel1, Maura Diamond1, Olivia Padovan4, Pichai Raman1,5, Yimei Li1, Jun S. Wei6, Shile Zhang6, Janahan Gnanchandran7, Robert Seeger7, Shahab Asgharzadeh7, Javed Khan6, Sharon J. Diskin1,8,9, John M. Maris1,8,9, and Kristina A. Cole1,8,9

Tumor and Stem Cell Biology

Abstract

Chromosome 6p22 was identified recently as a neuroblastoma susceptibility locus, but its mechanistic contributions to tumorigenesis are as yet undefined. Here we report that the most highly significant single-nucleotide polymorphism (SNP) associations reside within CASC15, a long noncoding RNA that we define as a tumor suppressor at 6p22. Low-level expression of a short CASC15 isoform (CASC15-S) associated highly with advanced neuroblastoma and poor patient survival. In human neuroblastoma cells, attenuating CASC15-S increased cellular growth and migratory capacity. Gene expression analysis revealed downregulation of neuroblastoma-specific markers in cells with attenuated CASC15-S, with concomitant increases in cell adhesion and extracellular matrix transcripts. Altogether, our results point to CASC15-S as a mediator of neural growth and differentiation, which impacts neuroblastoma initiation and progression.

Introduction

Neuroblastoma, a cancer of the developing autonomic nervous system, is the most common malignancy diagnosed in the first year of life and accounts for approximately 10% of all pediatric cancer mortality (1–3). While the majority of low-risk neuroblastoma patients are cured with surgery alone, 50% of patients have the high-risk form of the disease, and only about half of these children survive despite highly intensive therapy (1). Neuroblastomas are thought to develop from cells derived from the neural crest committed to the sympathoadrenal lineage (peripheral autonomic nervous system; refs. 1–4). Because malignant transformation can occur at any point during sympathetic development, tumors may arise throughout the developing sympathetic nervous system (most commonly the adrenal gland), contributing to the hallmark heterogeneity observed in this disease (4). To address the etiology of sporadic neuroblastoma, we conducted the first pediatric cancer genome-wide association study (GWAS), leading to the identification of numerous validated susceptibility loci in several populations (5–12). Moreover, we showed that many of these susceptibility alleles are specifically associated with disease phenotype and patient outcomes. The majority of these SNPs act in cis to influence expression of protein coding genes at these loci, and several of these transcripts, such as LMO1, BARD1, and LIN28B, appear to play an oncogenic role in established tumors (5–12).

The first identified neuroblastoma susceptibility locus identified by GWAS, and the one that remains most significant, mapped to chromosome 6p22.3 and robustly replicated in three independent cohorts [rs6939340: \( P = 9.33 \times 10^{-15} \); allelic OR 1.97, 95% confidence interval (CI), 1.58–2.45; ref. 5]. Like other subsequently identified loci, we observed a highly significant association with neuroblastoma susceptibility and clinically aggressive presentation. The minor allele (G) was overrepresented in neuroblastoma cases compared with controls, and presence of the G allele was further enriched in the high-risk subset of neuroblastoma (\( P = 0.007 \)), tumors with MYCN amplification (\( P = 0.002 \)) and stage IV disease (\( P = 0.025 \)), implying the risk alleles were associated with a more malignant neuroblastoma phenotype. On the basis of HapMap data available at the time of this initial discovery, the associated SNPs tagged a 94.2 kb linkage disequilibrium (LD) block; this LD block overlapped two hypothetical genes (FLJ22536 and FLJ44180; ref. 5). However, both FLJ22536 and FLJ44180 lacked protein-coding potential, impeding further characterization of this region in neuroblastoma initiation.

Recent data obtained from whole-genome sequencing studies of neuroblastoma have illustrated far fewer recurrent mutations in protein-coding genes than previously predicted (13–16); however, it is now clear that as much as 70% of the genome is transcribed into products other than traditional protein-coding mRNAs (17, 20).

**Note:** Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

**Corresponding Author:** Kristina A. Cole, Children’s Hospital of Philadelphia, 3501 Civic Center Boulevard, Philadelphia, PA 19104. Phone: 267-426-2285; Fax: 267-426-0685; E-mail: colekJ@email.chop.edu

doi: 10.1158/0008-5472.CAN-14-3613

©2015 American Association for Cancer Research.
Although many of these transcriptionally active loci produce RNA species involved in translation (i.e., ribosomal and transfer RNAs), several other RNA classes have been functionally validated as bona fide regulatory molecules. The recently identified long noncoding RNAs (lncRNA), defined as RNA species >200 nt in length that lack a functional open reading frame, have been increasingly implicated in a wide variety of cellular functions (19). lncRNAs share several transcriptional features in common with mRNAs, they are often spliced, demonstrate RNA polymerase II occupancy, contain a 5' methylguanosine cap, and are commonly (though not always) polyadenylated (20, 21). Although lncRNA function is highly context dependent, they commonly play a prominent role in the spatiotemporal regulation of gene expression during developmental processes (22–24), and therefore exhibit a tendency to be located throughout the genome, in sites proximal to developmentally critical protein-coding genes (25). Indeed, several lncRNAs reside near protein-coding genes known to regulate lineage commitment in neural crest cells (26), serving as an attractive hypothesis to explain the etiology of embryonal cancers such as neuroblastoma.

As might be expected, lncRNAs have been increasingly implicated in a variety of oncogenic processes through association with epigenetic complexes and modification of chromatin accessibility, ultimately influencing gene expression (27–30). To date, there are few reports concerning the role of lncRNAs in the initiation and progression of solid pediatric neoplasms, despite the fact that many childhood cancers are fundamentally defects of normal human development (31). Here we describe the identification and characterization of a novel lncRNA, CASC15, which contributes to the GWAS association signal on 6p22.3 by functioning as a tumor suppressor in neuroblastoma.

Materials and Methods

GWAS and imputation

In an effort to refine the association signal and search for a causal variant at the 6p22 locus, we performed genotype imputation in a previously described European ancestry cohort of 2,101 neuroblastoma cases and 4,202 controls (10) using the 1000 Genomes Phase I Release 3 as a reference. Genotyping and quality control methods have been previously published (5). GWAS imputation and statistical tests are detailed in Supplementary Materials and Methods.

Data sources

The human February 2009 (GRCh37/hg19) genome assembly was used throughout the study. Transcript scripts and annotations were obtained from GENCODE version 19. Details on the various data sources used are available in Supplementary Materials and Methods.

Neuroblastoma data. The neuroblastoma RNAseq, SNP profiling and HuEx datasets are part of the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative, supported by NCI Grant U10 CA98543. The low-level sequence data have been deposited in the Sequence Read Archive at the National Center for Biotechnology Information, and are further accessible through the database of genotypes and phenotypes (dbGAP, accession number phs000218). The gene expression and copy number data, as well as clinical information on the NBL cases studied, is available via the TARGET Data Matrix (32).

5’/3’ Rapid amplification of cDNA ends

5’ and 3’ Rapid amplification (RACE) was performed via the First Choice RLM-RACE Kit (Ambion) using 10 μg of RNA obtained from fetal brain or NB69 neuroblastoma cells following the manufacturer’s protocol. Specificity for CASC15-S was achieved by nested PCR using the following gene-specific primer (GSP) pairs:

- 5’ RACE: Outer GSP: 5’-TGAGCCCATGCTGCTGCTGTCG-3’
- 5’ RACE: Inner GSP: 5’-TTCACCGTGTCTCCGTCGTCG-3’
- 3’ RACE: Outer GSP: 5’-TGAGAGTCTGAGGAGTGCCCTG-3’
- 3’ RACE: Inner GSP: 5’-CTACCCCTGTCGCAACACCA-3’

Gene products were cloned into a pcR4-TOPO vector for sequencing.

RNA sequencing

Neuroblastoma. PolyA selected RNA libraries obtained from 108 high-risk neuroblastoma patients as part of the NCI TARGET project were prepared using TruSeq v3 (Illumina) for RNA sequencing on Illumina HiSeq2000 sequencers. The 101 bp paired-end reads generated were aligned to the hg19 build of the human reference genome using TopHat v2.2.0, yielding a median of 115 million total aligned reads per patient sample (range, 48–250 million total aligned reads). The HTSeq package (v0.6.1) was used to map aligned reads to the VISTA-annotated enhancer region, hs1335, as well as all transcripts annotated in Ref-Seq (v66) and/or the UCSC Genome Browser. Transcript expression values were normalized and quantified using the metric of reads per kilobase per million reads (RPKM). For isof orm-specific quantitation of CASC15 and CASC15-S, only exons and exon-exon junctions that were unique to each isoform were used in computing RPKM.

Other. Mapped RNA-seq reads from 16 primary tissues (Illumina Human Body Map) and 51 cell lines (Human ENCODE) were used to quantify gene- and transcript-level expression, based on GENCODE v9 annotations, using Cufflinks v.2.2.0 with default parameters and “--min-frags-per-transfrag 0 –compatible-hits-norm –min-isof orm-fraction 0.0.” Expression data heat maps were generated using the heatmap.2 function of the gplots R package.

RNA FISH

We performed RNA FISH and counted RNA in single cells as described in ref. 33. Fluorescently labeled, nonoverlapping oligonucleotide probes (20 mers) were designed to tile RNA transcripts, including CASC14 (LOC729177), hs1335, CASC15, and CASC15-S. Probes were then divided into odd and even pools and hybridized to neuroblastoma cells. Details of image acquisition are available in Supplementary Materials and Methods.

Quantification of CASC15-S in neuroblastoma cells

Quantification of RNA transcripts was performed on a panel of neuroblastoma cancer lines using TaqMan RT-PCR. For quantification of CASC15-S, we developed a custom primer/probe set spanning exons 1–2 and consisting of the following primers:

Forward: GCTGTCGACGAAGAGCTGAT
Reverse: GTCCAAGTCAAAAGTCTCATCCAAGA
Reporter: GCCGGCGCTGCTAGCCC

Primer/probe sets for CASC15 (HS01371949_m1) and CASC14 (HS04275511_s1) are commercially available (Life Technologies).
Quantification was normalized to the geometric mean of housekeeping genes TBP and GUSB.

Cell growth and siRNA assays
ON-TARGETplus SmartPool siRNAs containing four constructs per target (Thermo Scientific) were used for GAPDH and PLK-1 knockdown. ON-TARGETplus Non-targeting Control Pool (containing four constructs) was used as a control. Two constructs were used for each lncRNA, and were purchased as Silencer Select siRNA or custom siRNAs from Life Technologies.

siRNA constructs were transfected into neuroblastoma cells in triplicate, using 50 nmol/L of siRNA in 0.1%–0.2% DharmaFECT 1 (Thermo Scientific). Knockdown was assessed by TaqMan RT-PCR, comparing levels 72 hours after transfection to nontargeting control-transfected cells. For all siRNA experiments, the minimum knockdown achieved was 71.5% (average = 78.6, maximum = 92.0). For growth assays, while both tested, the siRNA with the best knockdown efficiency for each target is shown. Cell growth assays were conducted using the xCELLigence real-time growth (RT-CES, ACEA Biosciences) and/or Cell-TiterGlo (Promega) assays according to manufacturer protocols.

CASC15-S add-back experiments
For CASC15-S add-back experiments, expression plasmids containing the cDNA for either CASC15-S or GFP were transfected in triplicate in a 96-well plate using 250 ng of DNA and 3 μL of Lipofectamine 2000 per well. These experiments were conducted using the growth assays described above, and the expression levels of the CASC15-S transcript were confirmed by TaqMan RT-PCR to be highly expressed in the addback condition. Transfection efficiency was estimated to be approximately 90% based on the GFP-transfected condition.

shRNA expression and lentiviral constructs
The siRNA construct for CASC15-S (targeting exon 1) was used to create a double-stranded shRNA construct and was subsequently cloned into the pLenti-GFP DEST lentiviral vector (Addgene). A total of 5 × 10⁶ 293T cells were transfected with 15 μg of the lentiviral transfer vector, 15 μg of pLP1, 6 μg of pLP2, and 3 μg of pSV-G vectors. Lentiviral particles were collected at 48 and 72 hours after transfection. Lentiviral transduction of neuroblastoma cells was carried out overnight at 37°C using 3 μg/mL of polybrene. For each neuroblastoma line used, serial dilutions of lentiviral particles in basal media were carried out in a 6-well plate, and cells were assessed at 48 hours after transduction for GFP expression. The minimum lentiviral concentration resulting in >95% GFP expression was chosen for expansion and knockdown was validated by TaqMan PCR. For all shRNA experiments, knockdown of at least 70.3% (average = 82.6, maximum = 91.8) was achieved. Each neuroblastoma line with confirmed knockdown was then frozen at low passage (<5) and thawed as needed for subsequent experiments.

Luciferase reporter assay
The respective risk (A_A) and non-risk (T_T) genotypes of 1p and CHP-134 at rs9295534 were verified by PCR and Sanger sequencing. Subsequently, a 1,500 bp fragment was cloned from each cell line (750 bp on either side of the rs9295534) and was purified using the Qiaquick Gel Extraction Kit (Qiagen). The fragment was inserted into the pGJL4.23 vector upstream of a minimal promoter used to drive firefly luciferase expression. Fragment-containing vectors were then transfected into HEK-293 cells, along with a Renilla luciferase vector (pGL4.75) to control for transfection efficiency. Seventy-two hours after transfection, luciferase activity was assayed by Dual-Glo reporter assay (Promega) using the GloMax Multi Detection System (Promega). Firefly luciferase was normalized to Renilla luciferase and a minimum of three replicates was used for each condition.

In vitro transcription/translation assay
CASC15-S was cloned into a T7-promoter containing plasmid (T7-CFE-Chis) and subsequently verified by Sanger sequencing. Coding potential of CASC15-S was assayed using the TnT Quick Coupled Transcription/Translation System according to the manufacturer's protocol. Incorporation of biotinylated lysine residues was visualized via Western blot analysis using a 1:10,000 antibiotin HRP-labeled antibody (Cell Signaling Technology) and chemiluminescent detection (Thermo Scientific).

Wound-healing assays
Scratch assays were carried out on SK-N-BE2 and SK-N-SH cells stably depleted of CASC15-S plated at 85% confluence in 60 mm dishes, and were scratched using a sterile 200 μL pipette tip. Cells were photographed at regular intervals using a previously calibrated 5× light microscope (Nikon). Assessment of cell migration was carried out by measuring scratch closure as a percentage of initial scratch size in ImageJ, and was compared with control cells using a linear regression function in GraphPad Prism 6.

Cell culture
Neuroblastoma cell lines were obtained from the neuroblastoma cell line bank maintained at the Children's Hospital of Philadelphia, Philadelphia, PA. Cell line identity is routinely confirmed via AmpFLSTR Identifier (Applied Biosystems), last done in November 2013. Non-neuroblastoma cell lines were purchased from ATCC and all cell lines are routinely tested for mycoplasma. All cell lines are maintained in basal media (either RPMI1640 or DMEM) supplemented with 10% FBS and 1% gentamycin and cultured at 5% CO₂.

Statistical analysis
All group comparisons were conducted using nonparametric testing (Mann–Whitney U) in GraphPad Prism.

For Kaplan–Meier analysis: optimal cutoff was determined by employing a scanning approach to the Kaplan–Meier method by iteratively splitting the ordered genes expression values across samples into two groups and calculating the P value by the Mantel–Haenszel log-rank test. The lowest P value corresponds to the optimal breakpoint. A Benjamini–Hochberg correction was applied to reflect the presence of multiple hypotheses testing.

For multivariate analyses, a Cox Proportional Hazard model was used to evaluate the effect of each gene expression on overall survival, adjusting for clinical factors such as age (as a continuous variable), MYCN amplification, and 1p/11q LOH status. Proportional hazard assumption was evaluated by the log–log plot and there was no evidence of violation of this assumption.

Results
Fine mapping of the 6p22.3 neuroblastoma susceptibility locus
The initial GWAS study illustrating 6p22.3 as a neuroblastoma susceptibility locus identified three common polymorphisms
clustered on chromosome 6p22.3 that were highly associated with aggressive disease and a significantly increased risk of neuroblastoma development (5). These SNPs were shown to tag a 94.2 kb LD block overlapping two hypothetical genes and flanked upstream by SOX4. To map this region with finer detail, and identify putative causal/functional SNPs, we performed genotype imputation in our recently published discovery cohort of 2,101 cases and 4,202 controls (10). We first applied SHAPEIT (34) to infer haplotypes, and then utilized IMPUTE2 (35) with default parameters and Ne = 20,000, along with a multipopulation reference panel from the worldwide 1000 Genomes Project Phase 1 release to impute genotypes across the region. Genotyped and imputed variants were tested for association with neuroblastoma using the frequentist association test under the additive model (9). Imputed variants were tested for association with neuroblastoma risk using the “score” method implemented in SNPTEST (36). Variables with minor allele frequency $<$ 1% and/or IMPUTE2-info quality score $<$ 0.8 were excluded for quality control purposes. We identified 32 polymorphisms with P values less than 1 $\times$ 10$^{-10}$ ($P = 8.26 \times 10^{-10}$–1.88 $\times$ 10$^{-15}$), 12 of which were in high LD ($r^2 > 0.8$) and localized to an intronic region of an annotated gene locus formerly titled FLJ22536/LINC00340 and more recently renamed cancer-associated susceptibility candidate 15 (CASC15, Fig. 1A; Supplementary Table S1). These 32 SNPs identify a 34.9-kb linkage disequilibrium (LD) block based on the 1000 Genomes Northern and Western European (CEU) population (Fig. 1B), significantly refining this relatively small locus.

Identification of CASC15 isoform expression

CASC15 is annotated as an IncRNA, prompting us to examine the difference in IncRNA transcript levels between 220 high-risk and 30 low-risk primary tumors, for which we possessed microarray expression data. This analysis revealed that only 3.5% of annotated IncRNAs (hg19: 8/229) appear significantly differentially regulated between high- and low-risk tumors (Table 1). Furthermore, high-risk neuroblastomas exhibit a 4.4-fold lower level of CASC15 expression as compared with low-risk disease, the most highly significant ($P = 3.60 \times 10^{-17}$) differentially expressed IncRNA we observed. Several computationally predicted IncRNAs map to the 6p22.3 locus, including multiple CASC15 isoforms and an antisense-strand IncRNA, CASC14 (formerly LOC729177) (Supplementary Fig. S1A).

Figure 1.

Fine mapping of 6p22 identifies CASC15-S as a candidate co-acting neuroblastoma susceptibility gene. A, regional association plot of SNPs generated with LocusZoom software using genome-wide imputation data from 2,817 neuroblastoma cases and 7,473 controls. The pair-wise LD ($r^2$) for each SNP is denoted by color, and the log-transformed P values for each SNP are shown on the y-axis. Thirty-two SNPs with highly significant P values ($< 1 \times 10^{-16}$) ranging from $4.67 \times 10^{-10}$ to $4.81 \times 10^{-17}$ were found to cluster as a narrow peak within a genomic region containing the IncRNAs CASC14 and CASC15. The exons and transcribed regions of genes are shown as solid vertical lines. B, the haplotype structure of this region in Northern European (CEU) population demonstrates a signal that overlaps with these SNPs, refining and initial 94.2 kb LD block (boundaries denoted by black vertical dashed lines) down to 34.9 kb, demonstrated by the dashed red vertical line and rightmost black vertical dashed line. SNP annotations: color, representative of LD denoted by color, and the log-transformed coding or 3' UTR region; A, splice; A, nonsynonymous; *, fbcons, conserved motif at transcription factor binding site; mcs44placental, highly conserved region in placental mammals.
We therefore first utilized RNA sequencing (RNASeq) and RNA-paired end tagged (RNA-PET) data available from the ENCODE project to identify the products transcribed from this locus in SK-N-SH and SK-N-BE2 neuroblastoma cell lines. These data demonstrated the existence of two capped and polyadenylated nuclear CASC15 transcripts, a long (hg19, chr6: 21,666,675-22,194,616; Ensembl: CASC15-003) and a short isoform (hg19, chr6: 22,146,883-22,194,616; Ensembl: CASC15-004; Fig. 2A and Supplementary Figs. S1B and S2A). These nuclear noncoding transcripts are highly conserved in vertebrates and readily detected in several brain regions and neuroblastoma cell lines (Supplementary Fig. S2) with putative promoter regions separated by 480 kb indicative of independent transcriptional regulation.

We next examined RNA expression data from a panel of 16 primary human tissues as part of the Illumina Human Body Map project, where we found that the short CASC15 isoform was expressed in abundance in the brain, but at modest to low levels in other tissues (Fig. 2B). We subsequently generated RNA sequencing data from 108 primary neuroblastoma tumors using expressed in abundance in the brain, but at modest to low levels of both the 12-exon 1.9 kb CASC15 transcript (NR_015410.1, Ensembl: CASC15-003) and the 4-exon short (1.2 kb) variant (Ensembl: CASC15-004), hereafter referred to as CASC15-S. CASC15-S resembles a known cDNA clone (GenBank: AK094718), containing a unique first exon, yet sharing its remaining sequence with the last three exons of CASC15 (Fig. 2A and Supplementary Fig. S3A). Despite several attempts, we were unable to isolate a 5′-capped product for the intron-less transcript (Ensembl: CASC15-006) predicted to overlap exon 1 of CASC15-S and extend to the noncoding enhancer element (hs1335). Indeed, although the presence of this isoform is indicated on the SK-N-SH RNASeq track (Fig. 2A), it is absent in RNA-PET data (Supplementary Fig. S1B), supporting this result. Finally, RNA-FISH experimentally confirmed the presence, relative abundance, and cellular localization of these transcripts in both NB-69 and NGP neuroblastoma cells; using nonoverlapping, strand-specific probes to label CASC15-S, CASC15-C, and hs1335.

Table 1. Differential expression analysis of lncRNAs between high- and low-risk neuroblastomas

<table>
<thead>
<tr>
<th>Gene</th>
<th>Ratio</th>
<th>Disease state</th>
<th>P</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINC00340</td>
<td>4.4-fold lower in</td>
<td>High risk</td>
<td>3.60E–17</td>
<td>chr6:21666675-2294616</td>
</tr>
<tr>
<td>LINC00174</td>
<td>2.4-fold lower in</td>
<td>High risk</td>
<td>2.28E–15</td>
<td>chr7:6854703-6856539</td>
</tr>
<tr>
<td>LINC01296</td>
<td>8.9-fold higher in</td>
<td>High risk</td>
<td>5.80E–15</td>
<td>chr149980209-1995329</td>
</tr>
<tr>
<td>LINC00260</td>
<td>2.8-fold lower in</td>
<td>High risk</td>
<td>1.75E–13</td>
<td>chr120369970-20700979</td>
</tr>
<tr>
<td>LINC00221</td>
<td>2.4-fold higher in</td>
<td>High risk</td>
<td>7.04E–12</td>
<td>chr140693445-10691529</td>
</tr>
<tr>
<td>LINC00478</td>
<td>2.5-fold higher in</td>
<td>High risk</td>
<td>1.74E–04</td>
<td>chr217442842-17920094</td>
</tr>
<tr>
<td>LINC00514</td>
<td>2.4-fold lower in</td>
<td>High risk</td>
<td>3.70E–04</td>
<td>chr16039055-3044510</td>
</tr>
<tr>
<td>LINC00355</td>
<td>2.9-fold higher in</td>
<td>High risk</td>
<td>5.01E–04</td>
<td>chr156450650-6465044</td>
</tr>
</tbody>
</table>

Note: Differential gene expression analysis (|log fold|; P < 0.05; FDR < 0.05) was conducted using 220 high- and 30 low-risk primary neuroblastoma tumors, focusing on differences in lncRNA expression. The top differentially regulated lncRNA was CASC15 (annotated here as LINC00340), which was significantly lower in high-risk disease (4.4-fold decrease; ANOVA, P = 3.60 x 10^–17).
Finally, we looked for SNPs with evolutionary conservation, resulting in a single candidate polymorphism, rs9295534 (P = 3.51 × 10^{-12}; OR, 1.63; 95% CI, 1.4–1.89) upstream of CASC15-S, and this variant localizes to an expanse of regulatory chromatin and dense transcription factor binding sites in several cell lines (Fig. 3B). Moreover, this region exhibits enhancer activity, evidenced by H3K27Ac CHIP-Seq data in several fetal tissues available from the NIH Roadmap Epigenomics Mapping Consortium (Fig. 3C).

We next verified rs9295534 genotypes in Chp134 (homozygous non-risk) and Lan5 (homozygous risk) neuroblastoma cells by Sanger sequencing of a 1.5 kb region encompassing this SNP, and subsequently cloned risk and non-risk fragments from these lines. To assess the impact of rs9295534 genotype on transcriptional activity, we inserted risk (A/A) and non-risk (T/T) fragments into luciferase reporter constructs upstream of a minimal promoter. Results from these experiments demonstrated significantly decreased transcriptional reporter activity following insertion of the risk genotype fragment (Fig. 3D), suggesting this region possesses an enhancer-like function that is disrupted following the inclusion of a neuroblastoma risk allele.

CASC15-S is differentially expressed in neuroblastoma and highly associated with disease outcome

Because the rs9295534 homozygous risk genotype, by virtue of its linkage with rs6939340, is associated with aggressive neuroblastoma, poor survival (5), and demonstrates decreased...
transcriptional activity in a minimal promoter assay, we postulated that patients with high-risk disease would have reduced CASC15-S expression. Indeed, we observed significantly lower CASC15-S expression in high-risk patient tumors \( (n = 230) \) compared with low-risk patient tumors \( (n = 30; \text{Fig. 4A}) \). This finding appeared independent of MYCN amplification, a known oncogenic driver in neuroblastoma \( (P = 0.29, \text{Supplementary Fig. S5A}) \). In addition, patients with tumors enriched in CASC15-S expression exhibited superior survival when compared with patients harboring neuroblastoma with low CASC15-S expression, both when low- and high-risk patients were included in the analysis \( (adj. \text{ } P = 3.2 \times 10^{-10}, \text{Fig. 4B}) \) but also when comparing expression from only high-risk patients \( (adj. \text{ } P = 0.002, \text{Supplementary Fig. S5B}) \). Furthermore, this finding remained significant \( (P = 0.0084) \) after multivariate analysis adjusting for clinical factors such as age, MYCN amplification, and 1p/11q deletion status, all known prognostic variables in this disease \( (\text{Supplementary Table S2}) \). Although expressed at much lower levels than CASC15-S, the long isoform of CASC15 demonstrated a similar pattern in patient tumors \( (\text{Supplementary Fig. SSC–SSE}) \). Taken together, these data indicate that low CASC15-S expression correlates with a more aggressive phenotype in neuroblastoma and decreased overall survival probability.

To understand the contribution of CASC15-S in the initiation and progression of neuroblastoma, we performed gene set enrichment analysis \( (\text{GSEA}) \) restricted to high-risk neuroblastoma samples \( (n = 220) \), where we utilized the same 1.9-fold difference in median CASC15-S expression as the Kaplan–Meier analysis to define 146 low- and 74 high-expressing CASC15-S samples \( (\text{Fig. 4C}) \). The top differential gene expression profile \( (\text{normalized enrichment score} = -2.69, \text{nominal} \text{ } P \text{ } value < 0.0001, \text{FDR} \text{ } Q \text{ } value < 0.0001) \) that emerged from these analyses consists of a 55 gene signature previously shown by Asgharzadeh and colleagues to be downregulated in neuroblastoma patients with poor outcomes \( (39) \). This result indicates that tumors with high CASC15-S expression are enriched in the expression of these genes, and again
Cancer Res; 75(15) August 1, 2015

Cancer Research

3162

expression analyses between patients with high group, and these patients exhibited a significant increase in overall survival. D, the most highly significantly regulated pathway using GSEA was identified to be a set of genes downregulated in poor outcome neuroblastoma (Asgharzadeh neuroblastoma poor survival down). Patients with high CASC15-S expression demonstrated enrichment of these genes, suggesting that CASC15-S acts in a protective manner (***, P < 0.0001).

Figure 4.
Low CASC15-S expression correlates with poor clinical prognosis. A, microarray expression data from clinically annotated primary neuroblastoma tumors (n = 250) obtained at diagnosis reveals that high-risk (stage IV) neuroblastomas (n = 220) demonstrate significantly lower expression of CASC15-S than low-risk (stage I) tumors (n = 30). B, Kaplan–Meier analysis demonstrates significantly poorer overall survival for children with tumors expressing low levels of CASC15-S (n = 163) versus high (n = 87; P = 3.2 × 10⁻⁶). C, relevant metrics for selection of high-risk patient tumors used for differential gene expression analyses between patients with high (n = 163) and low (n = 87) CASC15-S levels. CASC15-S expression was increased 1.9-fold in the “high” group, and these patients exhibited a significant increase in overall survival. D, the most highly significantly regulated pathway using GSEA was identified to be a set of genes downregulated in poor outcome neuroblastoma (Asgharzadeh neuroblastoma poor survival down). Patients with high CASC15-S expression demonstrated enrichment of these genes, suggesting that CASC15-S acts in a protective manner (***, P < 0.0001).

CASC15 depletion in neuroblastoma cell lines enhances proliferation and invasive capabilities
Having demonstrated a clinically relevant association with patient outcome for CASC15-S in neuroblastoma patients, we next sought functional validation for a role in tumorigenesis. We first assessed CASC15-S levels by qRT-PCR across a well-characterized panel of neuroblastoma cell lines (n = 21) where we observed differential expression similar to our primary tumor panel (Fig. 5A). To investigate the functional role of CASC15-S in neuroblastoma, we utilized an siRNA construct targeting exon 4 near the 3′ end of the gene. Depletion of CASC15-S resulted in a highly reproducible increase in neuroblastoma proliferation as evidenced by real-time cell growth and viability assays (Fig. 5B). We were able to recapitulate these results by subsequently targeting CASC15-S within its unique first exon in several neuroblastoma lines (Fig. 5C and Supplementary Fig. S6A and S6B). Depletion of CASC15 or CASC14 had no observable impact on cell growth or viability, supporting our initial findings of CASC15-S as the functional isoform in neuroblastoma (Supplementary Figs. S6C and S6D and S7A–S7D). We next derived neuroblastoma cell lines stably depleted of CASC15-S (Supplementary Fig. S7E) and found that these cells also exhibited a substantial increase in cellular proliferation identical to what we observed in our transient siRNA-based CASC15-S depletion experiments (Fig. 4D and Supplementary Figs. S6E and S6F and S8). Furthermore, rescue experiments, conducted by ectopically expressing CASC15-S in these cells, were able to revert the accelerated growth (Fig. 5E). Microscopic examination revealed overt morphologic changes in shCASC15-S SK-N-BE2 cells, including striking changes in cell shape and size, with a resultant 3-fold increase in cell area (659.5 ± 50.2 μm² in control vs. 2024 ± 211.1 μm² in shCASC15 cells, P < 0.0001; Fig. 5F and G). Furthermore, both SK-N-BE2 (Fig. 6A and B) and SK-N-SH neuroblastoma cells (Fig. 6C) stably depleted of CASC15-S exhibited an increased migratory capacity and invasiveness as evidenced by wound-healing assays (SK-N-SH-
Taken together, these data suggest that CASC15-S loss promotes increased cellular growth and a more migratory phenotype in neuroblastoma. CASC15-S regulates a subset of genes involved in neural crest development.

To better characterize the phenotypic changes we observed following CASC15-S depletion, we surveyed gene expression signatures of neuroblastoma cells depleted of CASC15-S. For initial studies, SK-N-SH neuroblastoma cells were transfected in triplicate with either a nontargeting construct or siRNA specific for CASC15-S, and microarray gene expression signatures were examined at 48 hours following transfection. We observed substantial upregulation of several known cell adhesion genes, most notably entactin (NID1, \( p = 3.7 \times 10^{-4} \)) and activated leukocyte cell adhesion molecule (ALCAM, \( p = 1.05 \times 10^{-7} \)), after CASC15-S depletion. These gene expression data were used for Gene Ontology (GO) analysis to examine enrichment top level biologic processes. In support of the increased migratory phenotype we observed in wound-healing assays, both locomotion and cellular adhesion pathways were found among the top differentially regulated processes as a result of CASC15-S depletion (Fig. 6D).

We next assayed the result of persistent CASC15-S depletion on gene expression changes by comparing SK-N-BE2 neuroblastoma cells stably silenced for CASC15-S compared with control vector-transfected cells. We again examined gene expression signatures via microarray analysis, where we observed downregulation of
CASC15-S suggests that loss of gene signatures (Fig. 6F, bottom). Taken together, these data indicated increased migratory capacity in wound-healing assays (t = 24 hours). B and C, linear regression comparison of wound closure at regular intervals demonstrates a clear enhancement of migration in silenced SK-N-BE2 (B) or SK-N-SH (C) cells. E, GO analysis of SK-N-SH cells following CASC15-S depletion via siRNA at 48 hours. The top-level biologic processes are shown by percentage of enrichment signal, with both locomotion and cellular adhesion gene sets exhibiting over representation. F, a similar analysis was carried out for SK-N-BE2 cells stably expressing an shRNA construct targeting CASC15-S. The top enrichment observed in these cells, “cellular process,” was primarily the result of the cell differentiation gene subset.

Discussion

High-risk neuroblastoma remains a major challenge due to a relative paucity of somatic mutations hindering the development of targeted therapies (1). The identification of mutations in ALK and PHOX2B has helped explain the origin of familial neuroblastoma; however, an understanding of the basis of sporadic disease has only recently begun to come into focus (40, 41). Here we identify and demonstrate CASC15-S as a neuroblastoma suppressor gene via a post-GWAS mechanistic evaluation of a complex region of the human genome. Despite this robust association, however, eQTL analyses correlating risk genotypes with transcript expression failed to reach statistical significance, a likely...
CASC15 Is a Neuroblastoma Suppressor Gene

A consequence of an underpowered patient data set and/or additional mechanisms capable of impacting expression (such as additional SNPs affecting expression, posttranslational modification/degradation, etc.). To address this shortcoming, we fine mapped this locus, using orthogonal methodologies to identify the likely disease causal SNP, rs9295534, which localizes to the closest upstream enhancer of CASC15-S. Moreover, expression of the rs9295534 risk allele disrupts the enhancer function of this region, proving that genotype can indeed impact transcriptional ability at this locus. More importantly, we provide evidence for a potent effect of this lncRNA on neuroblastoma differentiation and migratory capacities, yielding mechanistic insights into why the GWAS signal at this locus is associated with metastatic disease and poor survival probability.

A growing body of work supports a defined role of lncRNAs as spatiotemporal-specific regulators of gene expression critical for ensuring proper differentiation during development. Thus, predominant expression of CASC15-S in brain (but not other tissues), the derivation of several cDNA clones from brain regions, and its proximity to a validated enhancer element (hst1335) strongly suggest that this lncRNA is uniquely involved in neural tube development. The role of lncRNA-mediated tumorigenesis in embryonal cancers provides a logical hypothesis to contribute to the understanding of the etiology of neuroblastoma tumors, which are typically devoid of activating somatic mutations (1).

A preliminary working model of how this transcript functions in neuroblastoma biology can be proposed from the functional and expression data we have demonstrated. CASC15-S expression strongly correlates with disease stage and overall survival, and patient tumors with high CASC15-S levels are enriched in genes typically lost in poor outcome neuroblastoma, demonstrating a protective role for CASC15-S. Furthermore, ablation of CASC15-S in neuroblastoma cell lines results in increased proliferative and migratory capacities, upregulation of adhesion and migration gene pathways, and a concomitant decrease in neural-specific transcripts. Therefore, reduced CASC15-S expression as the result of an inherited polymorphism would impact neural crest cellular lineage commitment and predispose these cells to undergo malignant transformation.

The phenotypic and gene signatures changes we observed signify that CASC15-S is responsible for maintaining a more differentiated and benign cell state, with CASC15-S loss leading to a poorly differentiated phenotype and expression of genes associated with transformed cells. It has been recently shown that CASC14 (renamed NBAT-1), although very lowly expressed, exerts a similar phenotype (42), suggesting that many lncRNAs in this region may cooperate. The precise mechanism by which CASC15-S (and potentially other transcripts near this locus) exerts its effect is currently under investigation, although it is likely that the observed transcriptional changes may be the result of modulation of cis elements. One possibility, given its proximity to a validated enhancer (hst1335) with neural tube expression and its position downstream of the developmental regulatory gene SOX4, is that CASC15-S functions as an enhancer RNA (43). Future studies to identify direct interaction partners of CASC15-S will undoubtedly strengthen our understanding of lncRNA function and yield key insights into neuroblastoma tumorigenesis.

In summary, our findings support a recent and growing body of evidence that convincingly demonstrates involvement of the noncoding genome in the tumorigenesis of pediatric cancers in general, and neuroblastoma in particular (42, 44–47). While we certify CASC15-S as neuroblastoma suppressor gene, we still have not elucidated all of the stochastic and/or epigenetic events that select for CASC15-S repression. Future studies focused on engineering CASC15-S depletion in vivo will further explore mechanisms for tumor initiation and progression as well as screen for synthetic lethal interactions to be leveraged therapeutically. Finally, this work provides the first highly significant GWAS-supported identification of lncRNAs in neuroblastoma, thus further stimulating exploration of this class of regulatory RNAs in human cancer etiology and clonal evolution.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: M. Russell, J.R. Alvarez-Dominguez, J.M. Maris, K.A. Cole


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M. Russell, A. Penikis, O. Padowan, J.S. Wei, R.C. Seeger, S. Asgharzadeh, I. Khan, J.M. Maris


Writing, review, and/or revision of the manuscript: M. Russell, D. Oldridge, J.R. Alvarez-Dominguez, Y. Li, S. Asgharzadeh, S.J. Diskin, J.M. Maris, K.A. Cole

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M. Russell, A. Penikis, M. Diamond, J.S. Wei, J. Gnanachandran, J.M. Maris

Study supervision: J.M. Maris, K.A. Cole

Acknowledgments

The authors acknowledge the Children’s Oncology Group (U10-CA98543) for providing blood and tumor specimens from neuroblastoma patients.

Grant Support

This work was supported by grant K08CA136979 (K.A. Cole) and Alex’s Lemonade Stand Foundation (K.A. Cole and M.R. Russell).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 USC Section 1734 solely to indicate this fact.

Received December 10, 2014; revised May 4, 2015; accepted May 5, 2015; published OnlineFirst June 22, 2015.

References


www.aacrjournals.org

Cancer Res; 75(15) August 1, 2015

3165

Published OnlineFirst June 22, 2015; DOI: 10.1158/0008-5472.CAN-14-3613

Downloaded from cancerres.aacrjournals.org on January 28, 2018. © 2015 American Association for Cancer Research.
Russell et al.


CASC15-S Is a Tumor Suppressor lncRNA at the 6p22 Neuroblastoma Susceptibility Locus

Mike R. Russell, Annalise Penikis, Derek A. Oldridge, et al.