Arlm1 is a male-specific modifier of astrocytoma resistance on mouse Chr 12

Jessica C. Amlin-Van Schaick, Sungjin Kim, Christina DiFabio, Min-Hyung Lee, Karl W. Broman, and Karlyne M. Reilly

Mouse Cancer Genetics Program, National Cancer Institute, Frederick, Maryland (J.C.A.-V.S., C.D.F., M.-H.L., K.M.R.); Institute for Biomedical Sciences, George Washington University, Washington, DC (J.C.A.-V.S.); Department of Biostatistics and Medical Informatics, School of Medicine and Public Health, University of Wisconsin, Madison, Wisconsin (S.K., K.W.B.)

While many cancers show a sex bias, the genetic basis and molecular mechanisms underlying sex bias are not always clear. Astrocytoma and glioblastoma show male predominance in humans. We have shown previously that glial tumors forming in the Nf1−/−; Trp53−/−cis (NPcis) mouse model also show a sex bias in some genetic contexts. Using cross-species comparisons we have identified candidate male-specific modifiers of astrocytoma/glioblastoma. Linkage analysis of B6X(B6X129)-NPcis mice identifies a modifier of astrocytoma resistance specific to males, named Arlm1, on distal mouse Chr 12. Arlm1 is syntenic to human Chr 7p15, 7p21, 7q36, and 14q32 regions that are altered in human glioblastoma. A subset of these genes shows male-specific correlations to glioblastoma patient survival time and represents strong candidates for the Arlm1 modifier gene. Identification of male-specific modifier genes will lead to a better understanding of the molecular basis of male predominance in astrocytoma and glioblastoma.

Keywords: astrocytoma, glioblastoma, modifier, sex differences.

Astrocytic gliomas, including astrocytoma and glioblastoma multiforme (GBM), are the most common primary tumors of the CNS in adults and are currently incurable. Both astrocytoma and GBM show male predominance in the population, with a male to female ratio of 1.31:1 for astrocytoma and 1.26:1 for GBM. A better understanding of the susceptibility factors for astrocytoma/GBM and the differences between males and females may lead to better approaches for prevention and treatment. Genome-wide association studies (GWASs) are beginning to enlighten the genetic risk factors for astrocytoma and GBM; however, the heterogeneity of the human population makes it difficult to find risk factors affecting subpopulations.

Mouse models of cancer offer a complementary approach to finding genetic susceptibility factors to human GWASs. Mice can be bred as homogeneous populations under controlled environmental conditions to identify cancer risk modifiers using smaller populations. These identified loci can be tested for their relevance to human disease. We have shown previously that tumorigenesis in the NPcis mouse model is dependent on genetic background and is additionally influenced by sex and parental inheritance of the NPcis mutation.

We are using this model of both astrocytoma/GBM and malignant peripheral nerve sheath tumors to understand how genetic factors, epigenetic factors, and sex interact to determine an individual’s risk for nervous system tumors. We can compare these factors with human data on GBM to identify promising candidates for further study.

The NPcis mouse model carries mutations in the Nf1 gene, encoding the RasGAP protein neurofibromin, and the Trp53 gene, encoding the tumor suppressor p53 (the TP53 gene in humans), together in cis on mouse Chr 11. Both Nf1 and TP53 are implicated in sporadic human GBM and this mouse model resembles human secondary GBM progressing from astrocytoma, as well as the mesenchymal subtype of GBM involving mutations in Nf1. NPcis mice on the 129S4/SvJae (129) strain are resistant to astrocytoma and GBM compared with NPcis mice on the C57BL/6J (B6) strain. Although B6X129 F1 hybrids from reciprocal crosses are genetically identical, they show varying susceptibility depending on how the NPcis mutation and the strain are inherited from the parents. This suggests that susceptibility polymorphisms interact with the inheritance of the
NPcis mutation on Chr 11 from the mother or the father or with additional parent-of-origin effects. These findings point to an interaction of genetic factors (strain background) with epigenetic factors (inheritance of genes from the mother or father) in determining the overall susceptibility to astrocytoma/GBM.

In addition to a genetic-by-epigenetic interaction, we have also found an interaction between the sex of the mouse and whether the mutant Chr 11 is inherited from the mother or the father (sex-by-epigenetic interaction). In inbred B6-NPcis mice where the NPcis mutation is inherited from the father (NPcis\textsuperscript{pat}), males and females have a similar incidence and grade spectrum of astrocytoma, with 52% of females and 49% of males developing astrocytoma/GBM.\textsuperscript{8,9} In B6-NPcis mice where the mutation is inherited from the mother (NPcis\textsuperscript{mat}), there is a gender bias in astrocytoma incidence and grade. Females develop 51% astrocytomas overall (71%) but no GBM.\textsuperscript{9} These data show that the increased incidence of astrocytoma in males is dependent on the inheritance of the parental NPcis allele. We therefore have examined the interaction between sex and genetic factors (sex-by-genetic interactions) in a backcross of B6 and 129 strains where the parental inheritance of the NPcis mutation is held constant.

### Materials and Methods

#### Mouse Breeding

The C57BL/6J and 129S4/SvJae strains were used for breeding experiments. We described the B6-NPcis\textsuperscript{pat} and 129-NPcis\textsuperscript{pat} datasets previously.\textsuperscript{6–9} The B6-NPcis\textsuperscript{pat} cohort consists of 45 males and 47 females. The 129-NPcis\textsuperscript{pat} cohort consists of 18 males and 16 females. We generated the F1(B6X129)-NPcis\textsuperscript{pat} cohort by crossing wild-type B6 females to 129-NPcis males and examined 25 males and 32 females. This cohort includes 23 F1 mice reported previously.\textsuperscript{7} We generated the B6X(B6X129)-NPcis\textsuperscript{mat} cohort by crossing wild-type B6 females to the B6X129-NPcis\textsuperscript{pat} F1 males. The B6X(B6X129)-NPcis\textsuperscript{mat} cohort consists of 28 males and 30 females. Mice were maintained at the National Cancer Institute–Frederick Laboratory at Science Applications International Corporation.

#### Genome-wide SNP Genotyping

We prepared tail DNA using the Promega Wizard SV Genomic DNA Purification System. We concentrated the DNA by ethanol precipitation to a range of 5 ng/\(\mu\)L to 155 ng/\(\mu\)L, with 75 ng/\(\mu\)L to 150 ng/\(\mu\)L considered optimal. Thirty-five \(\mu\)L of DNA was plated according to Center for Inherited Disease Research (CIDR) standard procedures and sent to CIDR for genotyping using the Illumina 1440K single nucleotide polymorphism (SNP) panel containing 880 SNPs polymorphic for B6 and 129.

#### Statistical Analysis of Linkage

We performed binary trait interval mapping using the genetic mapping software R/qtl\textsuperscript{14} to identify the locations of quantitative trait loci (QTLs). Statistical significance was determined by permutation testing with 1,000 replicates. A significance threshold of 0.05 was used to declare statistical significance. We analyzed males and females separately.

#### Bioinformatic Analysis of Mouse and Human Data

We identified polymorphic loci between B6 and 129 in the Arlm1 region using the SNP Wizard of the Mouse Phenome Database (http://phenome.jax.org/SNP/), accessing data from mouse genome build 37.1/mm9, dbSNP128, and Ensembl 48. We considered loci that were polymorphic between B6 and either 129S1 or 129X1. We analyzed publicly available gene expression datasets previously.\textsuperscript{6–9} The C57BL/6J and 129S1/SvImJ for each probe corresponding to a polymorphic gene in Arlm1. Probes that were up- or downregulated ≥1.5-fold were considered candidate hits. We identified human homologs of the polymorphic Arlm1 genes and chromosome cytobands using the National Center for Biotechnology Information (NCBI) Homology Maps Viewer (http://www.ncbi.nlm.nih.gov/projects/homology/maps/) and NCBI Unigene (http://www.ncbi.nlm.nih.gov/unigene), accessing mouse genome build 37.1 and human genome build 37.2. We assessed chromosome aberrations in human astrocytomas and GBM using NCBI Cancer Chromosomes Mittleman Data (http://www.ncbi.nlm.nih.gov/sites/entrez?db= cancerchromosomes), excluding cases of pilocytic astrocytoma. The effect of changes in gene expression levels on patient survival time or age of diagnosis was determined using Kaplan–Meier analysis in the Repository of Molecular Brain Neoplasia Data (REMBRANDT; https://caintegrator.nci.nih.gov/rembrandt/; Application Release 1.5.5, 7/27/10) and the Cancer Molecular Analysis (CMA) portal (https://cma.nci.nih.gov/cma-tcgag/; Version 2.0, 2/23/09), subdividing patient
data by sex and by whether the gene of interest was 2-fold
over- or underexpressed. Of the 100 genes polymorphic
between B6 and 129 in the Arlm1 locus, we identified 64
human homologs, of which 58 had probe sets available in
REMBRANDT and 53 had probe sets available in The
Cancer Genome Atlas (TCGA). We examined gene ampli-
ification or deletion using the Cancer Genome Workbench
Heatmap Viewer (https://cgwb.bni.nih.gov/cgi-bin/
heatmap; Update 9/2/10) to analyze the REMBRANDT
AST project (18 female, 28 male, 10 sex unknown
samples), the REMBRANDT GBM project (17 female,
59 male, 53 sex unknown samples), and the TCGA
project (793 combined, unpaired samples), with an
amplification threshold of 1. Copy number variation was also
determined using Cancer Genome Workbench Heatmap
Viewer.

Reverse Transcription qPCR for Crip2 and Cdca7l
RNA was isolated from the brains of age-matched
wild-type B6 and 129 mice or from cultured tumor cell
lines using TRIzol (Invitrogen Life Technologies) and
treated with TURBO DNAfree (Ambion). Mouse
tumor cell lines were described previously,15 and only
NPcispat cell lines were used for analysis. RNA was
reverse-transcribed using SuperScript II (Invitrogen Life
Technologies). Quantitative (q)PCR was carried out on
the Mx3000P qPCR machine (Agilent Technologies),
using Brilliant SYBR Green master mix (Agilent
Technologies). Quantitative (q)PCR was carried out on
the Mx3000P qPCR machine (Agilent Technologies),
using Brilliant SYBR Green master mix (Agilent
Technologies). Reverse Transcription qPCR for Crip2 and Cdca7l

Characterization of B6X(B6X129)-NPcis Progeny
Compared with B6-NPcis, 129-NPcis, and
B6X129-NPcis Progeny
We compared the astrocytoma/GBM-free survival of
the NPcispat F1(B6X129) hybrid (F1) and the
B6X(B6X129) (BC) backgrounds with mice on the par-
ental B6 and 129 strains to better understand the genet-
ics of susceptibility. By comparing F1 animals with the
parental lines, we can determine whether the B6 or
129 strain is dominant over the other strain.
129-NPcispat mice show a significant delay in astrocy-
toma latency compared with B6-NPcispat mice in
both males (P = .0047, log-rank test) and females
(P = .0008, log-rank test). When the B6 and 129
strains are intercrossed, both the F1 and BC

Mapping of Susceptibility to Astrocytoma
in B6X(B6X129)-NPcispat Mice
We mapped the presence or absence of astrocytoma
(WHO grades II–IV, including GBM) as a binary trait
in B6X(B6X129)-NPcispat mice. Males and females
were considered separately and combined, but no in-
crease in linkage was observed in the combined sexes
data not shown), suggesting that genetic susceptibility
to astrocytoma is linked to different loci in males
(Fig. 2A) and females (Fig. 2B). A significant linkage
peak (LOD 3.68; P = .024 genome-wide) was found on
distal Chr 12 in males (Fig. 2A and C). This locus
has no effect on female susceptibility to astrocytoma
(Fig. 2C and D), whereas males carrying the B6/B6
allele were resistant to astrocytoma (Fig. 2D) compared
with males carrying the B6/129 allele. The peak is
located at 3.63 cM on the far distal end of Chr 12.
The 1.5-LOD support interval, which defines the
region of interest, covers 53–63 cM on Chr 12. We
have named this the astrocytoma resistance locus in
males 1 (Arlm1).

Arlm1 Affects Astrocytoma Incidence, but not Latency,
in Males
We analyzed the B6X(B6X129)-NPcispat progeny based
on their genotype at Arlm1. Arlm1B6/B6 NPcispat males
developed significantly fewer astrocytomas (8%; n = 13;
P = .0002, Fisher’s 2-tailed test) than Arlm1B6/129
NPcispat males (30%; n = 15) (Fig. 2D). Arlm1B6/B6
NPcispat females developed similar numbers of astrocy-
tomas (65%; n = 17) as Arlm1B6/129 NPcispat females
(54%; n = 13) (Fig. 2D). In contrast, the astrocytomas
forming in Arlm1B6/129 males arise at the same age as
those in Arlm1B6/129 males, with no difference in
astrocytoma-free survival in either males (Supplementary Fig. S1A) or females (Supplementary Fig. S1B). These data are consistent with Arlm1 being a recessive resistance locus in B6 acting on astrocytoma incidence in males.

Bioinformatic Analysis of Arlm1 in Mouse and Human to Prioritize Candidates

We are interested in comparing the Arlm1 locus identified in mouse with human risk factors for astrocytoma and GBM. The region syntenic to Arlm1 has not yet been reported as a significant risk locus in GWASs of glioma (www.genome.gov/GWASStudies)\(^\text{16}\); however, this may be due to pooling of males and females in the analysis. We therefore took a broader approach to identify candidate Arlm1 genes likely to have a role in glioma based on a variety of data (Fig. 3). We used bioinformatic approaches combining mouse haplotype analysis, mouse strain-specific brain expression, human chromosomal synteny, and GBM patient data (summarized in Table 1 and described below) to prioritize Arlm1 candidate genes. For GBM patient data, we focused on genes within tumors that may alter the survival time of GBM patients as a surrogate for genes that affect GBM prior to tumor formation. These candidates can be tested directly in association studies for their role in glioma susceptibility in future studies.

Arlm1 Genes Are Differentially Expressed in B6 and 129 Male Mouse Brains

The Arlm1 locus covers the region of Chr 12 from 105 to 121 Mb and contains 100 genes that are polymorphic between B6 and 129S1/SvImJ and/or 129X1/SvJ (Mouse Phenome Database, http://phenome.jax.org/SNP/), including the Igh cluster. Of the polymorphic genes, 28 carry at least one nonsynonymous coding SNP predicted to change the sequence of the protein (Cn SNPs) (Table 2), and an additional 43 carry a SNP in the
untranslated region or first intron sequence, which are most likely to play a role in transcription regulation or transcript stability. For 73 of the 100 polymorphic genes in the Arlm1 locus, expression data were available on 10- to 12-week-old male brains from the B6 and 129S1/SvImJ strains (http://phenogen.ucdenver.edu/PhenoGen/web/datasets/geneData.jsp). Forty-one genes showed a ≥1.5-fold difference in expression levels between B6 and 129S1/SvImJ (Fig. 4). In combining coding differences and expression level differences, at least 55 genes are likely to vary between the B6 and 129 strains in the Arlm1 locus.

Arlm1 Is Syntenic with Human Chr 14q32, Chr 7q36, and Chr 7p15-21

We compared the Arlm1 locus on mouse Chr 12 with the human genome using the NCBI Homology Maps Viewer (http://www.ncbi.nlm.nih.gov/projects/homology/maps/). Arlm1 is syntenic with human Chr 14 98.8–106 Mb, corresponding to cytobands 14q32.1–0.33; Chr 7 20–22.4 Mb, corresponding to cytobands 7p21.2–7p15.3; and Chr 7 157.3–158.9 Mb, corresponding to cytoband 7q36.3. Importantly, it excludes the EGFR locus frequently amplified in human GBM.

Fig. 2. Linkage analysis for astrocytoma resistance in B6X(B6X129)-NPCis males and females. A graph of the genome-wide linkage analysis LOD scores for males is shown in (A) and for females in (B). The dotted line indicates the 0.05 alpha thresholds for statistical significance using permutation testing. (C) Linkage analysis of Chr 12 for both males (blue) and females (red). (D) The effect size of Arlm1^{B6/129} (SB) and Arlm1^{B6/B6} (BB) at 63.5 cM for males and females.
at cytoband 7p12 (located on mouse Chr 11), suggesting that an alternate gene from Chr 7p is involved. It includes the human IGH cluster on Chr 14q32. We investigated gains and losses of these regions in human GBM using NCBI Cancer Chromosomes Mittelman Data (http://www.ncbi.nlm.nih.gov/sites/entrez?db=cancerchromosomes), and we used the Cancer Genome Workbench Heatmap Viewer (https://cgwb.nci.nih.gov/cgi-bin/heatmap) to examine the REMBRANDT glioma and TCGA GBM datasets. Details of the numbers of samples analyzed are given in Table 1. Data from NCBI Cancer Chromosomes and the Cancer Genome Workbench Heatmap show that, overall, human Chr 7 is commonly amplified in GBM, while Chr 14 is commonly deleted. In addition, at the cytoband level there was evidence for specific alterations in the NCBI Cancer Chromosomes data, with both amplification and deletion of 7p15, deletion of 7p21, and amplification of 14q32. The TCGA GBM dataset represents a homogeneous group of GBMs, in which both Chr 7p15-21 and 7q36 regions were amplified in 74% of samples and deleted in 3%. Chr 14q32 was deleted in 27% of samples and amplified in 6%. The REMBRANDT glioma dataset represents a more heterogeneous population but showed similar results. Chr 7p15-21 was amplified in 21% of astrocytoma and 49% of GBM. Chr 7q36 was amplified in 17% of astrocytoma and 34% of GBM, and deleted in 5% of astrocytoma. Chr 14q32 was deleted in 17% of astrocytoma and 22% of GBM. In summary, the Arlm1 locus overlaps regions of the human genome that are both amplified (Chr 7) and deleted (Chr 14) in GBM, and we hypothesize that modifier candidates syntenic with Chr 7 are most likely to act in an oncogenic manner, while candidates syntenic with Chr 14 are most likely to act in a tumor-suppressive manner. Because the 129 allele
increases susceptibility to astrocytoma, we hypothesize that Arlm1 genes syntenic with human Chr 7 will have higher expression in 129 male brains and lower expression in B6 male brains, whereas genes syntenic with Chr 14 will have lower expression in 129 brains and higher expression in B6 brains. Figure 4 shows the ratio of 129 to B6 expression levels compared with whether genes are syntenic with Chr 7 or 14.

### Arlm1 Genes Show Sex-Specific Effects on GBM Survival Time

Amplification or deletion of genomic regions in cancer can point to regions where there is selective pressure to maintain or lose genes that affect cell growth. However, this selective pressure can apply to broad regions of the genome such that neighboring genes can influence amplifications and deletions of any given gene in cancer, giving rise to a “bystander effect.” As an example, the high rate of amplification of human Chr 7 in GBM is likely driven by strong selective pressure to amplify the EGFR locus, although the chromosome carries many genes that may not all contribute to cancer. Gene expression analysis and comparisons of mouse and human syntenic regions where neighboring relationships are unlinked can help to dissect the varying roles of different loci. We took advantage of GBM gene expression data for males and females from REMBRANDT and TCGA to examine the effects of low and high expression of Arlm1 candidates on patient survival time. The REMBRANDT website (https://caintegrator.nci.nih.gov/rembrandt/menu.do) was used to analyze glioma survival data by gene expression, and the CMA portal (https://cma.nci.nih.gov/cma-tcga/) was used to analyze TCGA datasets. Details of the number of samples analyzed are given in Table 1. Thirty genes show significant differences in survival time based on different gene expression levels. Of these, 16 genes show different effects in males and females. While increases and decreases in expression level may be secondary to copy number changes, we found that not all linked genes have the same correlation to survival time. For example, NCAPG2 and VIPR2 are linked on 7q36.3, but overexpression of NCAPG2 correlates with reduced survival time in males, whereas VIPR2 overexpression correlates with increased survival time (Supplementary Fig. S2). This suggests that specific correlations between gene expression levels and patient survival time are more likely to reflect a causal relationship between a modifier candidate gene and tumorigenesis, compared with gene amplifications and deletions.

TCGA samples showed far less variation in expression levels of the different genes compared with REMBRANDT samples, likely due to the more homogeneous population being assessed. Because of this, there were very few samples in the overexpressed and underexpressed category for each gene, and the statistics were not as robust. Nonetheless, 11 genes showed a significant effect on survival times in males vs females when stratified by expression level. In comparing the 2 independent datasets, we find 22 genes in which male-female differences are detected in at least 1 human dataset, and 7 genes where male-female differences are significant in both datasets. Of these 7 genes, only 2—CRIP2 on Chr 14q32.3 (Fig. 5) and CDCA7L on Chr 7p15 (Fig. 6)—show similar effects in both datasets, correlating with shortened survival time when overexpressed, and show a significant effect in males, but not females. Because of the small variation in the TCGA dataset, we do not want to exclude promising candidates found in only the REMBRANDT data. These include NCAPG2 and WDR60 on Chr 7q36; and Dicer1, WDR25, WDR20, Xrcc3, Tdrd9, and CDCA4 on Chr 14q32. Advanced GBMs, such as those in the TCGA dataset, may evolve to overcome the resistance

### Table 1. Databases and samples numbers used to identify candidate Arlm1 genes

<table>
<thead>
<tr>
<th>Description</th>
<th>Database</th>
<th># Samples</th>
<th># Genes Examined</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genes in Arlm1 locus</td>
<td>NCBI Map Viewer</td>
<td>NA</td>
<td>NA</td>
<td>503 genes</td>
</tr>
<tr>
<td>Mouse genes polymorphic B6 vs 129</td>
<td>Mouse Phenome Database</td>
<td>NA</td>
<td>267 genes</td>
<td>100 genes</td>
</tr>
<tr>
<td>Differential exp in male B6 vs 129 brains</td>
<td>PhenoGen</td>
<td>4–6 per strain</td>
<td>73 probe sets</td>
<td>41 genes</td>
</tr>
<tr>
<td>Human amp/del of Chr 7 or 14 in astrocytoma</td>
<td>NCBI Cancer Chromosomes</td>
<td>642 grade II-IV, NOS</td>
<td>NA</td>
<td>39% Chr 7</td>
</tr>
<tr>
<td>Human amp/del of Chr 7 or 14 in GBM</td>
<td>Cancer Genome Workbench (REMBRANDT)</td>
<td>56 astrocytoma (28 males, 18 females, 10 NOS)</td>
<td>NA</td>
<td>21% Chr 7</td>
</tr>
<tr>
<td>Human amp/del of Chr 7 or 14 in GBM</td>
<td>Cancer Genome Workbench (REMBRANDT)</td>
<td>130 GBM (59 male, 17 female, 53 NOS)</td>
<td>NA</td>
<td>49% Chr 7</td>
</tr>
<tr>
<td>Human amp/del of Chr 7 or 14 in GBM</td>
<td>Cancer Genome Workbench (TCGA)</td>
<td>793 GBM</td>
<td>NA</td>
<td>77% Chr 7</td>
</tr>
<tr>
<td>Correlation of expression with survival time</td>
<td>REMBRANDT</td>
<td>248 glioma (153 males, 95 females)</td>
<td>58 homologs</td>
<td>30 genes</td>
</tr>
<tr>
<td>Correlation of expression with survival time</td>
<td>Cancer Molecular Analysis Portal</td>
<td>210 GBM (131 males, 79 females)</td>
<td>53 homologs</td>
<td>11 genes</td>
</tr>
</tbody>
</table>

Abbreviations: NOS, not otherwise specified.
mechanisms of modifier genes, thus making it more difficult to identify subtle effects. Finally, because of the complexity of the \(Igh/IGH\) cluster, we did not subject it to the same level of analysis; however, the association of the immune system with glioma risk is clear from previous studies\(^1\)–\(^4\), and \(IGH\) is also a strong candidate in the \(Arlm1\) locus. Table 3 summarizes the evidence data from mouse and human for the highest priority candidates for a male-specific modifier at \(Arlm1\). Individual Kaplan–Meier curves of top candidates are shown in Supplementary Figure S2.

The genotype at the \(Arlm1\) locus does not affect astrocytoma-free survival in the \(NPcis\) mice (Supplementary Fig. S1A and B); however, the top candidates we have chosen from the human datasets correlate with changes in the length of survival after diagnosis in patients. To more directly compare \(Arlm1\) with the human gene candidates, we examined whether patients with different levels of \(CDCA7L\) or \(CRIP2\) in their tumors were diagnosed at different average ages (more analogous to the identification of mouse tumors that are not treated). Using Kaplan–Meier time-to-event analysis, we found no difference in the age of diagnosis in either male or female patients with different levels of

![Graph of the expression level of \(Arlm1\) candidate genes in B6 and 129 male mouse brains, taken from publicly available microarray data. Genes are organized along the Y-axis by their synteny with human Chr 7p (cyan), 7q (magenta), or 14q (green), corresponding to the cartoon in Fig. 3. Genes with no clear human homolog are shown in gray. The X-axis shows the ratio of 129/B6 on the log2 scale. Red dotted lines indicate the 1.5-fold up or down threshold used to prioritize candidates. The \(Igh\) locus showed a lot of variability between different probes in the dataset and was not included in the graph. Abbreviation: N.H., no homolog.](https://academic.oup.com/neuro-oncology/article-abstract/14/2/160/1039228)
CDCA7L or CRIP2, consistent with the findings in mice (Supplementary Fig. S1C–F).

Cdca7l Is Differentially Expressed between B6 and 129 Brains in both Males and Females

To verify the strain differences in expression of Cdca7l and Crip2, we used reverse transcription qPCR to measure the levels of the 2 genes in male and female B6 and 129 brains. There were no probes for Cdca7l in the publicly available microarray dataset for the B6 and 129S1/SvImJ (http://phenogen.ucdenver.edu/PhenoGen/web/datasets/geneData.jsp) (Table 3). We found that the 129S4/SvJae (129) strain that we used for our mapping studies expressed more Cdca7l than the B6 strain (1.6-fold difference for males, \( t \)-test \( P = .0030 \); and 1.5-fold difference for females, \( t \)-test \( P = .056 \)) (Fig. 7A). Crip2 expression in the publicly available microarray dataset was 1.8-fold higher in B6 than in 129S1/SvImJ males; however, we found no difference between B6 and 129 male brains. B6 females brains expressed 1.4-fold higher Crip2 than 129 female brains (\( t \)-test \( P = .049 \)) (Fig. 7B). Taken together these data suggest that Cdca7l is a stronger candidate for the Arlm1 locus and that strain differences in Cdca7l between B6 and 129 may have different effects depending on male or female physiology.

We extended the Cdca7l and Crip2 expression studies to mouse and human tumor cell lines. Mouse astrocytoma cell lines have been isolated from male and female B6-NPcis mice and from female 129-NPcis mice.\(^\text{15}\) We have not yet been able to derive astrocytoma cell lines from male 129-NPcis mice for reasons that are unclear, so we were not able to directly compare male tumors on the B6 and 129 strain background. Expression of Cdca7l in astrocytoma cells from B6-NPcis females and 129-NPcis females supports our findings in whole brain, in that 129 tumor lines express higher levels of Cdca7l compared to B6 tumor line (Fig. 7C), although we noted wide variability between lines from the same strain (Supplementary Fig. S3A). Crip2 expression was even more variable between lines (Supplementary Fig. S3B) and did not show the same trend we found in normal brains (Fig. 3D), again supporting Cdca7l as a stronger candidate for the Arlm1 locus than Crip2.
To examine whether there were male and female differences in \textit{Cdca7l} in tumors, we compared B6-NPcis male and female tumor lines and human GBM male and female tumor lines. We found variable levels of \textit{Cdca7l} expression between different tumor lines (Supplementary Fig. S3A and C), but for both mouse and human the highest expressing lines were male and the lowest expressing lines were female. Additional samples are needed to determine whether male astrocytomas/GBMs have a greater propensity for upregulating \textit{Cdca7l} than female tumors, but this data presents the possibility that the strain-specific increase in \textit{Cdca7l} in 129 mice interacts with male-specific increases in \textit{Cdca7l} in tumors to affect tumor susceptibility.

**Discussion**

We show using an unbiased mouse screen that a polymorphic modifier locus on mouse Chr 12 affects resistance to astrocytoma in males. Although this modifier region includes 100 polymorphic candidate genes, it is difficult to assess all of these individually, and cross-species comparison was used to prioritize the candidates for further study. Two genes, \textit{Cdca7l} and \textit{Crip2}, are of particular interest due to their correlation to GBM patient survival time in 2 separate human datasets; however, an additional 14 genes (\textit{Itgb8}, \textit{Dnab11}, \textit{Ncapg2}, \textit{Serpina}, \textit{Dicer1}, \textit{Bcl11b}, \textit{Evl}, \textit{Ppp2r5c}, \textit{Xrcc3}, \textit{Tdrd9}, \textit{Ccda4}, \textit{Jag2}, \textit{Crip1} and the \textit{Igh} locus) are also of great interest due to a variety of bioinformatic data as well as what is already known about these genes in cancer and the nervous system. Furthermore, it is important to note that in addition to these genes, other polymorphisms in the DNA in the region of \textit{Arlm1}, such as those affecting miRs or chromatin structure, could also affect astrocytoma resistance and were not specifically examined in this study. Nevertheless, prioritization of genes within the region allows a direct comparison with human GWASs and a more rapid test for the role of these genes in human astrocytoma/GBM.

\textit{CDCA7L} encodes the transcription factor R1 (\textit{RAM2/JPO2}) with links to both nervous system function and cancer. R1 was first identified as interacting...
Table 3. Top candidates for the *Arlm1* gene based on B6 vs 129 polymorphisms, differential expression in B6 and 129 male brains, hazard ratios for correlation to changes in GBM survival time in REMBRANDT (REMB) and TCGA datasets for males and females, and percentage of samples with copy number changes in human astrocytoma and GBM. Genes are listed in order of location in human genome.

| Genes on Mouse Chr 12 | Mouse B6 vs 129 Cn polymorphic between B6 and 129 | Mouse B6 vs 129 UTR/ Intron 1 SNPs | Mouse Brain Exp Fold Ratio 129/51/B6 | Human Homolog | Human Location | Hazard Ratio Males REMB \( n = 153 \) | Hazard Ratio Females REMB \( n = 95 \) | Hazard Ratio Males TCGA \( n = 131 \) | Hazard Ratio Females TCGA \( n = 79 \) | Percent Amp in REMB Astro | Percent Del in REMB Astro | Percent Amp in TCGA GBM | Percent Del in TCGA GBM | Percent Amp in GBM | Percent Del in GBM |
|-----------------------|---------------------------------|---------------------------------|---------------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Itgb8                 | 1 7 NC                           | NP                              | ITGB8                           | 7p21.1          | NS             | 1.95<sup>c</sup> | 2.02<sup>a</sup> | NS              | NS              | 3               | 0               | 45              | 0               | 70              | 3               |
| Dnac11                | 4 10 1.5                           | NP                              | DNAH11                           | 7p21            | 1.65<sup>b</sup> | 1.44<sup>d</sup> | NS              | NS              | 21              | 0               | 49              | 0               | 72              | 2               |
| Cdca7l                | 1 48 NC                           | NP                              | CDCA7L                           | 7p15            | 1.76<sup>b</sup> | NS              | 1.67<sup>f</sup> | NS              | 19              | 0               | 15              | 0               | 72              | 2               |
| Ncapg2                | 3 21 NC                           | NC                              | NCPG2                            | 7q36.3          | 1.48<sup>b</sup> | NS              | NS              | 3.74<sup>a,g</sup> | 17              | 0               | 13              | 0               | 74              | 2               |
| Wdr60                 | 1 12 NC                           | NC                              | WDR60                           | 7q36.3          | 2.31<sup>d</sup> | NS              | NS              | NS              | 17              | 0               | 34              | 0               | 74              | 2               |
| Vipr2                 | 1 2 NC                           | NC                              | VIPR2                            | 7q36.3          | 0.66<sup>a</sup> | 0.60<sup>b</sup> | 0.70<sup>e</sup> | NS              | 16              | 5               | 21              | 0               | 74              | 2               |
| Serpina1a             | 0 9 4.4                           | NP                              | SERPINA1                         | 14q32.13        | 1.57<sup>c</sup> | 1.96<sup>b</sup> | NS              | 2.19<sup>e</sup> | 0               | 5               | 0               | 4               | 27              |
| Serpina3n             | 0 9 4.1                           | NP                              | SERPINA3                         | 14q32.13        | 1.82<sup>c</sup> | 1.26<sup>e</sup> | NS              | NS              | 0               | 5               | 0               | 4               | 27              |
| Dicer1                | 1 2 0.7                           | NP                              | DICER1                          | 14q32.13        | 0.47<sup>b</sup> | NS              | 0.46<sup>f</sup> | NS              | 0               | 7               | 9               | 3               | 27              |
| Tcl1                  | 0 1 1.5                           | NC                              | TCL1A                            | 14q32.1         | 0.65<sup>a</sup> | 3.91<sup>b,g</sup> | NS              | NS              | 0               | 0               | 0               | 4               | 27              |
| Bcl11b                | 0 1 NC                           | NC                              | BCL11b                           | 14q32.2         | 0.73<sup>a</sup> | 0.47<sup>f</sup> | 0.098<sup>d,g</sup> | NS              | NS              | 0               | 0               | 3               | 4               | 27              |
| Evl                   | 0 2 NC                           | NC                              | EVL                             | 14q32           | 0.60<sup>b</sup> | 0.12<sup>d,g</sup> | NS              | NS              | 0               | 5               | 3               | 4               | 26              |
| Wdr25                 | 0 0 NC                           | NC                              | WDR25                           | 14q32.2         | 0.39<sup>a</sup> | NS              | 10.50<sup>a,g</sup> | NS              | 7               | 5               | 8               | 4               | 26              |
| Ppp2r5c               | 0 0 0.6                           | NP                              | PPP2R5c                         | 14q32.13        | 0.50<sup>a</sup> | 0.59<sup>b</sup> | NS              | NS              | 0               | 8               | 0               | 5               | 26              |
| Wdr20a                | 0 3 NC                           | NC                              | WDR20                          | 14q32.13        | 0.43<sup>d,g</sup> | NS              | NS              | NS              | 0               | 8               | 0               | 5               | 26              |
| Cdc42bp               | 0 1 0.5                           | NC                              | CDC42BP                         | 14q32.3         | 0.61<sup>a</sup> | 0.65<sup>b</sup> | NS              | 0.43<sup>d,g</sup> | 0               | 14              | 21              | 6               | 26              |
| Xrcc3                 | 0 3 NC                           | NC                              | XRCC3                           | 14q32.3         | 1.66<sup>a</sup> | NS              | NS              | 17              | 0               | 22              | 6               | 26              |
| Tdrd9                 | 0 1 1.5                           | NC                              | TDRD9                           | 14q32.33        | 0.76<sup>b</sup> | NS              | NS              | 0               | 8               | 0               | 5               | 5               | 26              |
| Cdc44                 | 0 0 2.1                           | NP                              | CDC44                           | 14q32.33        | 1.70<sup>a</sup> | NS              | NS              | 7               | 0               | 3               | 5               | 25              |
| Jag2                  | 0 0 0.4                           | NP                              | JAG2                            | 14q32           | 4.97<sup>a</sup> | NS              | 0.027<sup>d,g</sup> | NS              | 0               | 7               | 3               | 5               | 25              |
| Crip2                 | 1 0 0.6                           | NP                              | CRIP2                           | 14q32.3         | 2.41<sup>a</sup> | NS              | 4.01<sup>b,g</sup> | NS              | 0               | 7               | 3               | 5               | 25              |
| Crip1                 | 0 8 0.6                           | NC                              | CRIP1                           | 14q32.33        | 1.63<sup>b</sup> | NS              | NS              | NS              | 0               | 7               | 3               | 5               | 25              |

Abbreviations: NP, no probe; NC, no change; NS, not significant.

<sup>a</sup>\( P < 0.05. \)<sup>b</sup>\( P < 0.01. \)<sup>c</sup>\( P < 0.001. \)<sup>d</sup>\( P < 0.0001. \)<sup>e</sup>\( P < 0.1. \)<sup>f</sup>Underexp vs Intermed (2-fold change) is not significant (\( P = .18 \)); Overexp vs Underexp (4-fold change) is significant (HR = 2.66, \( P = .020 \)).
with Myc and being involved in medulloblastoma malignant transformation. It was further shown to inhibit apoptosis downstream of Bcl2 and to enhance proliferation downstream of Myc and upstream of cyclinD1/E2F1. This oncogenic role for R1 is consistent with its being within the 7p15 amplicon and being correlated with poorer prognosis in males when overexpressed in GBM. We show that the B6 brains carrying the resistance allele of Arlm1 express lower levels of Cdca7l than wt 129 female brains (n = 5) and wt 129 male brains (n = 8) (A). The difference between B6 and 129 females is suggestive (t-test P = .056) and the difference between B6 and 129 males is significant (t-test P = .0030). Wt B6 female brains express higher levels of Crip2 than wt 129 female brains (t-test P = .049), but males show no difference in Crip2 expression between B6 and 129 (B). Brackets indicate which data comparison is statistically significant (*P < .05; **P < .005). Female B6-NPcis astrocytoma lines (n = 2) express lower levels of Cdca7l than female 129-NPcis astrocytoma lines (n = 4) (C), consistent with results in normal brains (A); however, the number of available tumor lines is too low to draw statistically significant conclusions. The level of Crip2 in female B6-NPcis astrocytoma lines (n = 2) and 129-NPcis astrocytoma lines (n = 4) (D) is variable and does not reflect the expression differences seen in normal brain (B). Red bars indicate B6 samples and blue bars indicate 129 samples. Solid bars represent male samples and checked bars represent female samples.

CRIP2 encodes a molecular adaptor protein in the LIM domain family of proteins that has recently been identified as a tumor suppressor in nasopharyngeal carcinoma. CRIP2 was originally identified as a binding partner of the protein tyrosine phosphatase PTP-BL, localized to the cell cortex with a potential role in regulating the actin cytoskeleton. It is highly expressed in brain and shows specific expression patterns in subsets of neurons and during neural crest
glioma risk. Progression of gliomas requires an angiomaliares more strongly than females in one study of glioma grade and prognosis. Four of the candidate ITGB8 characterized by microvascular proliferation. The genic switch and high-grade glioblastomas are characterized by increased microvascular proliferation.1 The IGH candidates are linked to immune system function: the in different hallmarks of cancer, genes involved in DNA damage repair in glioma prevention, 19 and several Arlm1 candidates have been implicated in chromatin dynamics and DNA repair. NCA82 and DNAHC11 encode structural proteins important for chromosome condensation and segregation, and XRCC3 encode proteins involved in homologous recombination and genome stability. XRCC3 has been studied in human association studies for its potential role in susceptibility to glioma. Intriguingly, the rs3212092 SNP in XRCC3 affects males more strongly than females in one study of glioma risk. Progression of gliomas requires an angiogenic switch and high-grade glioblastomas are characterized by microvascular proliferation.1 The Arlm1 candidate ITGB8 regulates tumor angiogenesis in brain cancer 48 and is upregulated in GBM. The association of immune function and glioma risk is now well established,19 with an inverse association between glioma and allergies or asthma, 31–33 associations of cytokine signaling and invasion pathways with glioma risk, and conflicting roles for infiltrating T-cells in glioma grade and prognosis.55–57 Four of the Arlm1 candidates are linked to immune system function: the IGH locus, SERPINA1, SERPINA3, and CRIP1. In addition to the genes directly involved in different hallmarks of cancer, genes involved in regulation of non–protein coding RNAs can have indirect effects on cancer. Non–protein coding RNAs are increasingly recognized as major tumor suppressors and oncogenes in cancer, including GBM. DICER1 and TDRD9 regulate microRNAs and Piwi-interacting RNAs, respectively, and could have wide-ranging, complex effects on cancer.

Through cross-species comparison, we identify Cdc7l as a high priority candidate as well as 15 additional strong candidates for a male-specific modifier of brain cancer. These candidates can be tested directly in human GWAS for glioblastoma risk as they become available. Future studies of the mechanism of action of Arlm1 will give new insights into the male predominance of both astrocytoma and GBM, and may suggest new therapeutic or preventative approaches.

Supplementary Material

Supplementary material is available online at Neuro-Oncology (http://neuro-oncology.oxfordjournals.org/).

Acknowledgments

We thank M. Anvers, R. Tuskan, K. Smith, K. Fox, and E. Truffer for technical assistance and B. Mock, J. Simmons, N. Colburn, and K. Hunter for helpful discussions and comments on the manuscript. J.C.A.-V.S. is a predoctoral student in the Graduate Partnership Program of the N.I.H. and the Institute for Biomedical Sciences at George Washington University. This work is from a dissertation to be prepared in partial fulfillment of the requirements for the PhD degree. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsements by the US Government.

Conflict of interest statement. None declared.

Funding

Intramural Research Program of the National Institutes of Health, National Cancer Institute [ZIA BC 010539 to K.M.R.]; federal funds from the National Cancer Institute to SAIC Frederick [contract N01-CO-12400]; federal contract from the National Institutes of Health to The Johns Hopkins University [contract HHSN268200782096C]; extramural funding from the National Institutes of Health [R01 GM074244 to K.W.B.].
References


