Published in final edited form as:

Genes Chromosomes Cancer. 2009 August ; 48(8): 661–672. doi:10.1002/gcc.20673.

Genomic and Clinical Analyses of 2p24 and 12q13-q14 Amplification in Alveolar Rhabdomyosarcoma: A Report from the Children's Oncology Group

Frederic G. Barr^{1,*}, Fenghai Duan^{2,†}, Lynette M. Smith², Donna Gustafson¹, Mandy Pitts³, Sue Hammond^{3,4}, and Julie M. Gastier-Foster^{3,4,5}

¹Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA 19104

²Department of Biostatistics, College of Public Health, Nebraska Medical Center, Omaha, NE 68198

³Department of Laboratory Medicine, Nationwide Children's Hospital, Columbus, OH 43205

⁴Department of Pathology, Ohio State University College of Medicine, Columbus, OH 43205.

⁵Department of Pediatrics, Ohio State University College of Medicine, Columbus, OH 43205.

Abstract

Alveolar rhabdomyosarcoma (ARMS) is an aggressive pediatric cancer that is related to the skeletal muscle lineage and characterized by recurrent chromosomal translocations. Within the ARMS category, there is clinical and genetic heterogeneity, consistent with the premise that “primary” genetic events collaborate with “secondary” events to give rise to subsets with varying clinical features. Previous studies demonstrated that genomic amplification occurs frequently in ARMS. In the current study, we used oligonucleotide arrays to localize two common amplicons to the 2p24 and 12q13-q14 chromosomal regions. Based on the copy number array data, we sublocalized the minimum common regions of 2p24 and 12q13-q14 amplification to a 0.83 Mb region containing the *DDX1* and *MYCN* genes, and a 0.55 Mb region containing 27 genes, respectively. Using fluorescent *in situ* hybridization assays to measure copy number of the 2p24 and 12q13-q14 regions in over 100 cases, we detected these amplicons in 13% and 12% of cases, respectively. Comparison with fusion status revealed that 2p24 amplification occurred preferentially in cases positive for *PAX3-FOXO1* or *PAX7-FOXO1* while 12q13-q14 amplification occurred preferentially in *PAX3-FOXO1*-positive cases. Expression studies demonstrated that *MYCN* was usually overexpressed in cases with 2p24 amplification while multiple genes were overexpressed in cases with 12q13-q14 amplification. Finally, although 2p24 amplification did not have a significant association with clinical outcome, 12q13-q14 amplification was associated with significantly worse failure-free and overall survival that was independent of gene fusion status.

Keywords

alveolar rhabdomyosarcoma; amplification; CDK4; MYCN

*Corresponding author: Frederic G. Barr, M.D., Ph.D., Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, 505C Stellar Chance Laboratories, 422 Curie Boulevard, Philadelphia, PA 19104; USA; Phone: 215 898-0884; Fax: 215 898-4227; barrfg@mail.med.upenn.edu.

†Present address: Center for Statistical Sciences, Brown University, Providence, RI 02912

INTRODUCTION

Alveolar rhabdomyosarcoma (ARMS) is a soft tissue sarcoma related to the skeletal muscle lineage (Wexler et al., 2006). This cancer is characterized by its propensity to occur in the extremities in adolescents and young adults. Furthermore, ARMS is known for its aggressive behavior with early dissemination, poor response to treatment, and frequent relapses after treatment. From a genetic perspective, the most notable feature of ARMS is the presence of recurrent chromosomal translocations (Barr, 2001). One translocation – t(2;13)(q35;q14) – occurs in the majority of ARMS cases while a variant translocation – t(1;13)(p36;q14) – occurs in a smaller subset. These translocations break and rejoin cellular genes – *PAX3* and *PAX7*, on chromosome 2 and chromosome 1, respectively, and *FOXO1* (*FKHR*) on chromosome 13 – to generate *PAX3-FOXO1* and *PAX7-FOXO1* fusion genes. The wild-type genes encode transcription factors, and the resulting fusion proteins are chimeric transcription factors with oncogenic activity. Assays of the fusion products revealed that *PAX3-FOXO1*, *PAX7-FOXO1*, and the fusion-negative scenario occur in 55%, 22%, and 23% of ARMS tumors, respectively (Sorensen et al., 2002). Further studies also showed phenotypic differences between the *PAX3-FOXO1* and *PAX7-FOXO1* subtypes of ARMS (Kelly et al., 1997).

Despite the importance of the recurrent 2;13 and 1;13 translocations in ARMS, there is evidence that additional genetic events occur during ARMS pathogenesis and contribute to the biological and clinical behavior of this tumor type (Pandita et al., 1999; Gordon et al., 2000; Bridge et al., 2002). From a biological standpoint, these additional events would potentially function as collaborating factors, and from a clinical standpoint, these additional events would potentially function as modifying factors. Comparative genomic hybridization (CGH) is a technology that can find such additional events by surveying the genome for copy number gains and losses. In studies of ARMS, CGH identified the frequent occurrence of gene amplification events, finding at least one amplicon in 26-73% of ARMS cases. The most commonly amplified chromosomal regions were 2p24 and 12q13-q15, with corresponding frequencies of 24% and 18% of cases, respectively (Gordon et al., 2000). Other amplicons reported in smaller subsets of cases include 1p36, 2q34-qter, 13q14, and 13q31.

In this manuscript, we conduct a detailed analysis of the amplification events involving the 2p24 and 12q13-q15 chromosomal regions. Using high-density oligonucleotide copy number arrays to analyze a panel of ARMS cases, we determine the minimum common amplified regions, and then identify the human genes within these regions. This information is then used to design directed FISH and DNA-based PCR assays to screen larger panels of ARMS cases for the amplification status of these regions. RNA-based assays are then performed to analyze the expression status of the genes within these regions. Finally, the amplification and expression data are compared with other pathologic, molecular, and clinical data to investigate the significance of these amplification events.

MATERIALS AND METHODS

Tumor collection

For copy number arrays, 57 ARMS tumors were received from the Children's Oncology Group (COG) Soft Tissue Sarcoma (STS) Tumor Bank. The criteria for selection of these cases was a review diagnosis of ARMS by current diagnostic indicators (including the presence of 50% ARMS histology) (Qualman et al., 1998) and the presence of sufficient frozen tissue for preparation of DNA, RNA, and touch imprint slides for fluorescence *in situ* hybridization (FISH). For subsequent studies, samples from an additional non-overlapping 151 cases from the COG STS Tumor Bank were analyzed; at the time of the current study,

the available material from these latter cases varied and included formalin-fixed paraffin-embedded blocks, frozen blocks, or isolated RNA and DNA. As above, all of these cases also had a review diagnosis of ARMS using current criteria.

Of the 208 samples in this study, 155 samples were derived from patients enrolled on Intergroup Rhabdomyosarcoma Study (IRS) Group or COG STS clinical trials. These cases were distributed among the following trials: IRS IV (n = 54), IRS V pilot (n = 5), D9501 (n = 5), D9502 (n = 7), D9602 (n = 5), D9802 (n = 20), and D9803 (n = 59). These trials have similar chemotherapy control arms with comparable outcomes, and experimental arms did not show statistically different outcomes when compared with the control arms. Therefore there is justification for combining the study populations for outcome analysis in the present study.

Fusion status was determined in most of the cases using either real-time RT-PCR or standard gel-based RT-PCR assays (Barr et al., 1995; Barr et al., 2006). One of the original 57 cases was deemed to be indeterminate for fusion status due to the finding of a low but significant *PAX3-FOXO1* expression in one set of readings and no measurable expression in a second set of readings.

DNA preparation and microarray analysis

DNA from the initial 57 tumors and selected additional cases was extracted using the QIAamp DNA Mini Kit (Qiagen, Valencia, CA), and analyzed on Affymetrix GeneChip Human Mapping 50K *Xba*I arrays of the 100K Mapping Panel (Affymetrix, Santa Clara, CA). The isolated DNA was processed and labeled using the protocol recommended by the manufacturer. The labeled samples were submitted to the Penn Microarray Core Facility for hybridization and analysis using protocols and reagents recommended by the manufacturer. Selected samples were reanalyzed on Affymetrix GeneChip Human 250K *Sly*I arrays of the 500K mapping panel.

The 50K and 250K array data were analyzed using Affymetrix GeneChip DNA Analysis Software version 3.0, Affymetrix Copy Number Analysis Tool (CNAT) versions 2.2 and 3.0, Affymetrix GeneChip Operating Software version 1.4, Affymetrix GeneChip Genotyping Analysis Software version 4.1, and Partek Genomics Suite version 6.3 (Partek, St. Louis, MO). Original 50K copy number data were calculated in CNAT using an algorithm developed to identify genome-wide gains and losses on high density oligonucleotide arrays based on SNP genotyping methods and whole genome sampling analysis. The raw data were analyzed with CNAT to calculate a copy number at each locus, which was then normalized relative to a reference set of normal individuals. The copy number estimates were displayed with a genome-smoothed analysis, which incorporates Gaussian kernel smoothing over the default genomic filter length of 0.5 Mb. Data from the 250K array were analyzed with the Hidden Markov Model default settings in the Copy Number Workflow in Partek Genomics Suite. The working definition of amplification is a segment of multiple contiguous loci in which the relative copy number starts from a background level of 2, rises to greater than 4, and then returns to the background level.

Fluorescent *in situ* hybridization (FISH)

Interphase FISH was performed on touch imprints made from snap frozen or OCT-embedded tumor. A minimum of 100 interphase cells was counted for each probe with a DAPI II counterstain (Abbott Molecular; Abbott Park, IL). To test for amplification of the 2p24 chromosomal region containing the *MYCN* gene, the Vysis LSI N-MYC Spectrum Green/CEP2 SpectrumOrange probe (Abbott Molecular; Abbott Park, IL) was used. Nuclei were analyzed for the number of green test signals (corresponding to a 200 kb 2p24

chromosomal region surrounding *MYCN* gene) and red control signals (chromosome 2 centromere). A BAC probe (RP11-571M6) containing 15 genes from the 12q13.3-q14.1 region, including *CDK4*, was direct-labeled with SpectrumOrange-dUTP or SpectrumGreen-dUTP using the Nick Translation Kit (Abbott Molecular; Abbott Park, IL). This probe was hybridized in combination with a probe from the 12q15 chromosomal region (RP11-611O2 and RP11-450G15) or a Vysis CEP12 (Abbott Molecular; Abbott Park, IL) control probe from the chromosome 12 centromere.

Quantitative PCR (qPCR) assay

DNA copy number of the *CDK4* gene was quantified along with the control gene *MGP* (located at chromosomal region 12p13) by multiplex qPCR. The primer-probe sets were designed with Primer Express software (Applied Biosystems) and the PCR product spanned exon/intron boundaries to ensure specificity for genomic DNA. These qPCR assays were performed on an ABI Prism 7700 Sequence Detection System (Foster City, CA).

Quantitative RT-PCR (qRT-PCR) assays

RNA was extracted from the frozen tumors using RNA STAT-60 (Tel-Test, Friendswood, TX). RNA was screened for *CDK4* expression by qRT-PCR using a primer-probe set designed with Primer Express software (Applied Biosystems). A previously published primer-probe set was used to assay for *MYCN* expression (Williamson et al., 2005). RNA was screened for expression of other selected genes (*CENTG1*, *DDIT3*, *DDX1*, *DTX3*, *MBD6* and *PIP5K2C*) using premade qRT-PCR assays from Applied Biosystems. All qRT-PCR assays were run on 384-well plates in the ABI Prism 7900 Sequence Detection System. The control gene 18S RNA was assayed in separate wells of the same run using a premade qRT-PCR assay (Applied Biosystems). The association of 18S RNA expression with ARMS clinical parameters, such as stage, clinical group, tumor size, or nodal status, was studied and no statistically significant associations were found.

Statistical analysis

The Fisher's exact test was used to test for an association between two discrete variables in a contingency table. For a contingency table other than 2×2, the Freeman-Halton test (an extension of Fisher's exact test) was applied instead (Freeman and Halton, 1951).

Kaplan-Meier's method was used to generate plots to evaluate survival distributions and the non-parametric log rank test was applied to compare survival curves in several groups. Cox proportional hazards model was further fitted to determine predictive values of interesting variables in univariate and multivariate survival analyses.

In this study, all gene expression levels were logarithmic transformed (specifically, log₂ transformed) prior to analyses. The main reason was that the transformed values were found to be more normally distributed than the original ones (data not shown). To study the expression of *MYCN* and *CDK4* varying across different categories (Supplementary Table 2 and Supplementary Table 4), the repeated measurements ANOVA was used to account for correlation between sample duplicates.

RESULTS

Mapping amplicons using oligonucleotide arrays

We assembled a panel of 57 ARMS cases from the COG STS Tumor Bank consisting of 31 *PAX3-FOXO1*-positive (55%), 9 *PAX7-FOXO1*-positive (16%), 16 fusion-negative (29%), and one indeterminate fusion status. DNA was isolated from all cases and analyzed for genome-wide copy number on the 50K *XbaI* array of the Affymetrix GeneChip Human

Mapping 100K Set. In this study, we focused on two chromosomal loci that are frequently amplified in ARMS, 2p24 and 12q13-q15. The initial goals of this study were to evaluate the incidence and composition of these amplification events.

For the 2p24 chromosomal region, an amplification event was identified in five of the 57 cases (3 *PAX3-FOXO1*, 1 *PAX7-FOXO1*, 1 indeterminate). Though the frequency of 2p24 amplification in this initial series is low, our subsequent analysis of a larger series of cases will demonstrate a higher frequency of 2p24 amplification. In the copy number array analysis, the endpoints of each amplicon were operationally defined as the markers where the copy number returns to the background level. When the endpoints of the five 2p24 amplicons were compared, we identified a minimum common region of amplification measuring 1.79 Mb (Fig. 1A). To further localize this minimum common region, three of the five amplified ARMS cases were further analyzed for copy number on one of the 250K Affymetrix GeneChip arrays from the 500K mapping set. These arrays have 5-fold higher density of markers across the genome and permit a higher resolution analysis of the precise boundaries of the amplification units. By moving both sides of the minimal common region inward, the size of this minimal common region of amplification for the 2p24 amplicon was reduced to 0.83 Mb. To identify the human genes located within this region, we accessed the corresponding human genetic map (NCBI Build 36.2) and superimposed it to the map of the amplification units. In the 2p24 region, though 18 genes are distributed among the various amplified regions, only *DDX1* and *MYCN* are localized within the minimum common region.

For the 12q13-q15 chromosomal region, we identified two distinct amplification events in nine of the 57 cases. Eight of the amplified cases (7 *PAX3-FOXO1*, 1 *PAX7-FOXO1*) had an amplification event involving the 12q13.3-q14.1 region (henceforth referred to as 12q13-q14), and three of the amplified cases (2 *PAX3-FOXO1*, 1 fusion-negative) had an amplification event involving the 12q15 region. The 12q13-q14 and 12q15 amplification events occurred simultaneously, with an intervening non-amplified region, in two of the nine amplified cases (2 *PAX3-FOXO1*). Due to the presence of a low number of cases with the 12q15 amplicon, the minimal common amplified could only be localized to a 2.2 Mb region that contains 27 genes, including the well-recognized proto-oncogene *MDM2* (data not shown). In contrast, a series of eight cases with the 12q13-q14 amplicon permitted localization of the minimal common amplified region to a small segment of 0.86 Mb, based on 50K Affymetrix GeneChip mapping data (Fig. 1B). The 12q13-q14 amplicons from six of the cases were further examined on 250K GeneChip arrays and the high density mapping data permitted the minimal common amplified region to be further refined to a 0.55 Mb region. However, when this region was superimposed on the human genome map, we were surprised to find that this relatively small region contains 27 human genes, including *CDK4*.

Directed copy number studies

Based on the gene composition of the amplicons, FISH assays were set up to detect the corresponding 2p24 and 12q13-q14 amplicons. We analyzed 53-54 of the 57 ARMS cases with each probe to determine the sensitivity and specificity of the copy number arrays for detecting these amplification events compared to FISH. For both the 2p24 and 12q13-q14 amplicons, there was 100% concordance between the FISH results and copy number array results. In particular, the FISH assays detected all five cases with 2p24 amplicons and all eight cases with 12q13-q14 amplicons that were detected by the copy number arrays. These FISH assays further revealed that the actual cells with the amplification events contained numerous copies of the amplified locus (≥ 20 copies for 2p24 and >10 copies for 12q13-q14). Furthermore, these FISH assays did not detect the 2p24 or the 12q13-q14 amplicons in any additional cases. Therefore, these arrays demonstrated high sensitivity and specificity for detecting these amplification events.

We then proceeded to assay copy number in an additional series of cases. Utilizing the FISH assay described above to assay copy number of the 2p24 region, amplification was detected in a total of 17 of 126 ARMS cases (13%) (Table 1A). A comparison of 2p24 copy number with fusion subtype indicates a significantly higher frequency of amplification in fusion-positive tumors. The frequencies of 2p24 amplification in *PAX3-FOXO1* and *PAX7-FOXO1*-positive tumors were 18% and 22%, respectively, while the frequency in fusion-negative tumors was ~8-fold lower (2.4%, $p = 0.018$).

In a similar analysis of 12q13-q14 copy number in a large series of cases, we utilized the FISH assay described above for most cases and a qPCR assay to measure *CDK4* copy number in a small number of cases for which only isolated DNA was available. Using these methodologies, we detected 12q13-q14 amplification in 13 of 109 ARMS cases (12%) (Table 1B). Among the three fusion subtypes, 12q13-q14 amplification was significantly higher in *PAX3-FOXO1*-positive cases (12/39, 24%) than in *PAX7-FOXO1*-positive cases (1/27, 4%, $p = 0.028$) or in fusion-negative cases (0/31, 0%, $p = 0.0027$).

Expression of genes in minimal common amplified region

To further study downstream consequences of the amplification event, we measured mRNA expression of genes within the minimal common amplified region. For the 2p24 amplicon, we used qRT-PCR assays to measure *MYCN* and *DDX1* expression in cases for which we had available RNA and had previously measured amplification status. As a group, cases with 2p24 amplification have significantly higher mean *MYCN* and *DDX1* expression levels than non-amplified cases. In a closer analysis of 106 cases with *MYCN* and *DDX1* expression data, 10 of 15 amplified cases were in the first (highest) quintile for *MYCN* expression while only seven of 15 amplified cases were in the first quintile for *DDX1* expression (Fig. 2). The other amplified cases without high expression had *MYCN* expression levels scattered among the second, third and fourth expression quintiles and *DDX1* expression levels scattered among the third, fourth, and fifth expression quintiles. A notable divergence between *MYCN* and *DDX1* expression is found in three amplified cases in which *MYCN* levels were in the first quintile while the *DDX1* levels were in the third or fourth quintiles. This finding indicates that *MYCN* can be overexpressed without *DDX1* overexpression, and thus provides support that *MYCN* is more likely the target gene of the 2p24 amplicon in ARMS. However, these findings also indicate that 2p24 amplification is not always associated with *MYCN* overexpression.

To analyze expression of genes in the 12q13-q14 amplicon, we selected six of the 27 genes (*CDK4*, *CENTG1*, *DDIT3*, *DTX3*, *MBD6*, *PIP5K2C*) and developed qRT-PCR assays. These six genes were selected based on the criterion of a biological function potentially related to oncogenesis. For five of these genes (*CDK4*, *DDIT3*, *DTX3*, *MBD6*, *PIP5K2C*), statistical analysis of the *PAX3-FOXO1*-positive cases indicated that the amplified cases expressed significantly higher levels of each gene relative to the cases without amplification ($p < 0.0001$) (Fig. 3). We noted that for three of these five genes (*DTX3*, *PIP5K2C*, *MBD6*), one of the eight amplified *PAX3-FOXO1* cases did not express high levels, though this low expressing case was not the same for all three genes. Furthermore, consideration of the *PAX7-FOXO1*-positive and fusion negative cases, which were not amplified (except for one *PAX7-FOXO1*-positive case), also revealed significantly lower expression of these five genes compared to the *PAX3-FOXO1* amplified cases. In contrast to the first five genes, qRT-PCR assay of *CENTG1* expression did not demonstrate evidence of overexpression in the amplified cases. Therefore, from this initial sampling of genes in the minimal common region of 12q13-q14 amplification, we conclude that many but not all of the genes are overexpressed as a result of amplification.

Association of amplification and gene expression with clinical parameters

To determine the association of amplification and/or expression of these genes with differential aspects of ARMS tumorigenesis or tumor progression, we investigated the correlation of 2p24 and 12q13-q14 amplification with specific clinical characteristics (Supplementary Tables 1-4). In addition, we also correlated expression of one gene from each amplicon whose overexpression may serve as a surrogate marker for amplification, and thus selected *MYCN* and *CDK4*. In particular, we examined age, gender, race, stage, group, tumor size, nodal status, tumor invasiveness, and tumor site. In a comparison with tumors lacking 2p24 amplification, 2p24-amplified tumors showed a trend towards occurrence in non-white individuals (60% vs. 40%, $p = 0.084$), less Stage 3 (20% vs. 47%, $p = 0.063$), and smaller tumors (40% vs. 68%, $p = 0.09$), but these findings were not statistically significant (Supplementary Table 1). When *MYCN* expression was compared to these parameters, there was a trend for higher *MYCN* expression in association with metastatic sites and extremity lesions, again without reaching statistical significance (Supplementary Table 2). In survival analysis, 2p24 amplification was not associated with a significant difference in failure-free ($p = 0.57$) or overall survival ($p = 0.87$) (Table 2 and Fig. 4A). However, *MYCN* expression demonstrated an interesting relationship with outcome such that increases in expression are associated with improved outcome. Though this relationship is not statistically significant for failure free survival, ($p = 0.29$), a 2-fold increase in *MYCN* expression was associated with a 13% decrease in the hazard ratio ($p = 0.024$) for overall survival.

When 12q13-q14 amplified cases were compared to cases without this amplification, several interesting associations were found (Supplementary Table 3). In particular, amplification was found more often in males (91% vs. 59%, $p = 0.049$) and was associated with nodal metastasis (73% vs. 41%, $p = 0.056$) and increased tumor invasiveness (91% vs. 59%, $p = 0.048$). When *CDK4* expression was examined, there was a trend towards higher expression in males, and higher expression could still be seen in the nodal metastasis and higher invasiveness groups, but with much higher p values (Supplementary Table 4). In survival studies, there was a statistically significant association of 12q13-q14 amplification in both failure-free and overall survival (Table 2 and Fig 4B). In particular, amplification of the 12q13-q14 region was associated with a worse outcome; the hazard ratio in failure-free and overall survival was increased to 2.24 ($p = 0.023$) and 2.26 ($p = 0.040$), respectively. As a representative gene whose expression is increased by this amplicon, analysis of *CDK4* expression level also showed a significant association of increased expression with poor outcome. In particular, a 2-fold increase in *CDK4* expression was associated with a 40% increase in the hazard ratio ($p = 0.038$) for failure-free survival.

Findings in previous studies suggested an association of the *PAX3-FOXO1* fusion subtype with a poor outcome (Sorensen et al., 2002). Since the far majority of 12q13-q14-amplified cases occur in *PAX3-FOXO1*-positive cases, we investigated whether the predictive effect of the 12q13-q14 amplicon is independent of fusion status. Towards this end, we performed a subsequent analysis of the clinical features of 45 *PAX3-FOXO1*-positive ARMS cases with and without 12q13-q14 amplification (Supplementary Table 5). Comparison of these two groups revealed that the 12q13-q14 amplified subset still has an increased frequency of male patients and a trend towards increased nodal status and tumor invasiveness. Most importantly, outcome analysis demonstrated that, among *PAX3-FOXO1*-positive cases, 12q13-q14 amplification is still associated with a significant decrease in failure-free survival (10% vs. 37% 5 year survival, $p = 0.042$) (Fig. 4C). In the analysis of overall survival, there is still a marked trend toward a worse outcome associated with amplification ($p = 0.056$).

Finally, a multivariate analysis was performed to determine the predictive value of 12q13-q14 amplification in the presence of the known prognostic variables for ARMS. The variables of clinical group, tumor size, lymph node status, and primary site were included in

this analysis. After adjusting for the effects of these four clinical variables, the effect of 12q13-14 amplification was not statistically significant for failure-free survival ($p = 0.19$) (Table 3) or overall survival ($p = 0.30$). Therefore, with the current number of test cases, 12q13-q14 amplification was not yet a significant independent predictor of outcome.

DISCUSSION

In this paper, we used genomic tools to define the common 2p24 and 12q13-q14 amplification events in ARMS, elucidate the expression pattern of the genes within these amplicons, and assess the frequency of these amplicons in a well-annotated panel of ARMS tumors. The results of these experiments emphasize the very different properties of these two amplification units in ARMS. Though the minimal common regions of amplification of the 2p24 and 12q13-q14 amplicons were both localized well under 1 Mb, there was a 10-fold difference in the number of genes within the minimal 2p24 and 12q13-q14 amplicons: two versus twenty-seven. Expression analyses indicate that multiple 12q13-q14 genes appear to be overexpressed in 12q13-q14-amplified cases. In the case of the 2p24 amplicon, *MYCN* is overexpressed in most 2p24-amplified cases, though there is also a small subset of amplified cases without high *MYCN* expression. In correlations with fusion status, both amplicons are more frequent in fusion-positive cases, but the 2p24 amplicon is found at comparable levels in both *PAX3-FOXO1* and *PAX7-FOXO1*-positive cases while the 12q13-q14 amplicon is more specific to the *PAX3-FOXO1* category. Finally, in the assessment of ARMS outcome, the 12q13-q14 amplicon is associated with a poor prognosis (though not independently) whereas the 2p24 amplicon does not have predictive value.

Amplification of the 12q13-q14 genomic region, including *CDK4*, has been found in a diverse set of tumors (Knuutila et al., 1998; Myllykangas et al., 2007), though these issues have not been well studied in ARMS before the current study. The highest frequency of 12q13-q14 amplification has been found in malignant adipose tumors in which 75-100% of well differentiated and dedifferentiated liposarcomas contain this amplicon while benign adipose tumors are rarely amplified (Shimada et al., 2006). In addition, amplification of this 12q13-q14 region occurs at lower frequencies in other bone and soft tissue tumors, brain tumors (most notably glioblastoma multiforme), hematopoietic tumors (notably B cell lymphomas), and epithelial tumors (such as breast and lung carcinoma) (Knuutila et al., 1998; Myllykangas et al., 2007). In general, the presence of this amplification is mutually exclusive with alterations of the *CDKN2A* locus or other lesions involving genes of the RB1 pathway, thereby indicating that *CDK4* is functional in tumors with this amplicon (Ichimura et al., 1996). The association of this amplification with outcome has also been studied in several tumors, with decreased survival associated with amplification in anaplastic astrocytoma (Backlund et al., 2005) and mantle cell lymphoma (Bea et al., 1999), while no correlation with outcome was found in glioblastoma (Houillier et al., 2006).

The structure of the 12q13-q14 amplicon has been studied in detail in non-RMS sarcomas and gliomas, and our current findings show similarities and differences with these studies. In particular, the previous studies initially used Southern blot technology and found two distinct amplicons in the 12q13-q15 region, a proximal amplicon containing *CDK4* and a more distal amplicon containing *MDM2* (Berner et al., 1996; Reifenberger et al., 1996). The distinct nature of these amplicons was evidenced by cases in which only one region is amplified and other cases in which both regions are amplified with an intervening non-amplified region. In the early studies, the minimal 12q13-q14 amplicon was initially localized in gliomas and non-RMS sarcomas by identifying cases with *CDK4* and *TSPAN31* (*SAS*) amplification and without *OS9*, *GLI* or *DDIT3* (*CHOP*) amplification. More recent studies with a BAC array confirmed these findings and showed that the amplicons in bone and other non-RMS soft tissue tumors span a 0.42 or 0.91 Mb region including the *PIP4K2C*

through *LOC338805* loci (Heidenblad et al., 2006). Similarly CGH analysis of the 12q13-q14 amplicon in glioblastoma on a cDNA array found evidence of amplification of a 0.18 Mb region including the *B4GALNT1* through *AVIL* loci (Ruano et al., 2006). These studies thus indicate that the 12q13-q14 amplicon in these tumor types overlaps but is different than the 12q13-q14 amplicon found in ARMS. In particular, the ARMS minimal region includes a 0.3 Mb segment located closer to the centromere that includes loci spanning from *KIF5A* to *R3HDM2*. Therefore, the presence of overlapping and distinct parts of the ARMS minimal common 12q13-q14 amplicon suggests that there are some critical target genes distinct to ARMS as well as an overlapping set of critical target genes common to ARMS and these other tumor types. Though the focus of studies on this amplicon has previously been on CDK4, this structural data further add to our overall expression findings to implicate additional genes in the oncogenic output of this amplicon in ARMS.

Amplification of the 2p24 chromosomal region has been detected in ~25% of neuroblastomas and a small subset of retinoblastomas, medulloblastomas, and small cell lung cancers (Schwab, 2004). This amplicon has been extensively studied in neuroblastoma, in which the 2p24 amplicon demonstrates some similarities and other striking differences from the corresponding amplicon in ARMS. In neuroblastoma, the copy number is usually in the range of 50- to 100-fold. Structural analyses indicated that the amplified region in neuroblastoma spans from 350 kb to over 1 Mb, but all contain a core 130 kb domain in which only *MYCN* is situated. In addition, ~50% of the amplicons also contain *DDX1* in addition to *MYCN*. Expression studies in neuroblastoma have shown that amplification is usually associated with increased mRNA and protein expression, but a few amplified cases have been reported with low *MYCN* expression (Matsunaga et al., 2000), similar to the finding in the current ARMS study. From a clinical perspective, the occurrence of *MYCN* amplification in neuroblastoma is directly correlated with disease stage, ranging from 5% incidence at low stages to 30% at advanced stages (Brodeur et al., 1997). Furthermore, amplification of this region is associated with poor outcome independent of stage, and is now an important factor used to assign neuroblastoma patients to treatment categories.

Two previous studies have investigated the clinical significance of *MYCN* amplification in ARMS. In one study using FISH on paraffin-embedded formalin-fixed sections, *MYCN* amplification was detected in 9 of 15 (60%) ARMS cases, and the cases with *MYCN* amplification showed a decreased survival rate compared to the non-amplified cases (11% vs. 66%, $p < 0.01$) (Hachitanda et al., 1998). In a second study, quantification of *MYCN* copy number by qPCR detected *MYCN* amplification (> 4-fold increase in copy number) in 12 of 48 (25%) ARMS cases (Williamson et al., 2005). In addition, low-level *MYCN* gains (1.5-4-fold increase in copy number) were found in 26 of 48 (54%) ARMS cases. In clinical correlative studies involving 34-38 of these ARMS cases, *MYCN* copy number gains (amplification and low level gains) and increased *MYCN* expression (greater than the median) were associated with a significantly worse outcome as measured by failure-free or overall survival. These previous two studies therefore differ from the current study finding that 2p24 and *MYCN* amplification has no predictive value. To put this difference in context, it should be noted that our current study of *MYCN* amplification status in 126 total cases, including 93 cases on IRS Group or COG protocols, constitutes the largest such study to date, more than six times as large as the first study and more than twice as large as the latter study. Since *MYCN* amplification occurs in fusion-positive cases, most of which are *PAX3-FOXO1*-positive, the clinical data in the previous studies may in part reflect an association with gene fusion status in a small ARMS case set. In addition, our study used FISH on touch preparations of frozen tissue, an optimal technique for assessing amplification, and confined our attention to the cases with true amplification. It should also be noted that the finding in the second study associating high *MYCN* expression with poor prognosis further contrasts with the finding in the current study that increased *MYCN* expression instead is associated

with improved overall survival. Of note, this same association of increased expression with improved outcome was found in a recent expression microarray study of ARMS cases (Davicioni et al., 2006).

In summary, this study has extensively characterized two frequent amplicons in ARMS and demonstrated the clinical significance of the 12q13-q14 amplicon. Based on the available data, this amplicon appears to identify an aggressive subset of *PAX3-FOXO1*-positive tumors. The clinical data indicate that the aggressive nature of these tumors is not due to a propensity for increased distal metastasis. Instead the available data suggests that tumors with 12q13-q14 amplification may be more locally invasive and may have an increased frequency of nodal metastasis, though neither of these comparisons reaches statistical significance when confined to the *PAX3-FOXO1*-positive tumors. Studies of larger numbers of ARMS cases are needed to further investigate these issues and to better define the predictive value of this molecular marker. Finally, as described above, there are multiple genes within this amplicon that may be responsible for this aggressive behavior. Multiple molecular and cell biology strategies must be used to carefully assess these genes and define the critical targets, with the potential goal of developing strategies to reverse these phenotypic effects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge the contributions of Julie Moore, Fred Ullrich, Christine Pressey, and Dimple Khurana to this project. We gratefully acknowledge the leadership, organization, advice, and encouragement provided by our friend and colleague, the late Stephen J. Qualman.

Supported by: NIH grants CA104896, CA81659, CA24507 and the NIH Children's Oncology Group Chair's Grant U10 CA98543; Alveolar Rhabdomyosarcoma Research Fund; and the Joanna McAfee Childhood Cancer Foundation. Tissue samples were provided by the Pediatric Division of the Cooperative Human Tissue Network, which is funded by NIH grant CA054021.

REFERENCES

- Backlund LM, Nilsson BR, Liu L, Ichimura K, Collins VP. Mutations in Rb1 pathway-related genes are associated with poor prognosis in anaplastic astrocytomas. *Br J Cancer*. 2005; 93:124–130. [PubMed: 15970925]
- Barr FG. Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma. *Oncogene*. 2001; 20:5736–5746. [PubMed: 11607823]
- Barr FG, Smith LM, Lynch JC, Strzelecki D, Parham DM, Qualman SJ, Breitfeld PP. Examination of gene fusion status in archival samples of alveolar rhabdomyosarcoma entered on the Intergroup Rhabdomyosarcoma Study-III trial: a report from the Children's Oncology Group. *J Mol Diagn*. 2006; 8:202–208. [PubMed: 16645206]
- Barr FG, Xiong QB, Kelly K. A consensus polymerase chain reaction-oligonucleotide hybridization approach for the detection of chromosomal translocations in pediatric bone and soft tissue sarcomas. *Am J Clin Pathol*. 1995; 104:627–633. [PubMed: 8526204]
- Bea S, Ribas M, Hernandez JM, Bosch F, Pinyol M, Hernandez L, Garcia JL, Flores T, Gonzalez M, Lopez-Guillermo A. Increased number of chromosomal imbalances and high-level DNA amplifications in mantle cell lymphoma are associated with blastoid variants. *Blood*. 1999; 93:4365–4374. others. [PubMed: 10361135]
- Berner JM, Forus A, Elkhouloun A, Meltzer PS, Fodstad O, Myklebost O. Separate amplified regions encompassing CDK4 and MDM2 in human sarcomas. *Genes Chromosomes Cancer*. 1996; 17:254–259. [PubMed: 8946207]

- Bridge JA, Liu J, Qualman SJ, Suijkerbuijk R, Wenger G, Zhang J, Wan X, Baker KS, Sorensen P, Barr FG. Genomic gains and losses are similar in genetic and histologic subsets of rhabdomyosarcoma, whereas amplification predominates in embryonal with anaplasia and alveolar subtypes. *Genes Chromosomes Cancer*. 2002; 33:310–321. [PubMed: 11807989]
- Brodeur GM, Maris JM, Yamashiro DJ, Hogarty MD, White PS. Biology and genetics of human neuroblastomas. *J Pediatr Hematol Oncol*. 1997; 19:93–101. [PubMed: 9149737]
- Davicioni E, Finckenstein FG, Shahbazian V, Buckley JD, Triche TJ, Anderson MJ. Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas. *Cancer Res*. 2006; 66:6936–6946. [PubMed: 16849537]
- Freeman GH, Halton JH. Note on an exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika*. 1951; 38:141–149. [PubMed: 14848119]
- Gordon AT, Brinkschmidt C, Anderson J, Coleman N, Dockhorn-Dworniczak B, Pritchard-Jones K, Shipley J. A novel and consistent amplicon at 13q31 associated with alveolar rhabdomyosarcoma. *Genes Chromosomes Cancer*. 2000; 28:220–226. [PubMed: 10825007]
- Hachitanda Y, Toyoshima S, Akazawa K, Tsuneyoshi M. N-myc gene amplification in rhabdomyosarcoma detected by fluorescence in situ hybridization: its correlation with histologic features. *Mod Pathol*. 1998; 11:1222–1227. [PubMed: 9872655]
- Heidenblad M, Hallor KH, Staaf J, Jonsson G, Borg A, Hoglund M, Mertens F, Mandahl N. Genomic profiling of bone and soft tissue tumors with supernumerary ring chromosomes using tiling resolution bacterial artificial chromosome microarrays. *Oncogene*. 2006; 25:7106–7116. [PubMed: 16732325]
- Houillier C, Lejeune J, Benouaich-Amiel A, Laigle-Donadey F, Criniere E, Mokhtari K, Thillet J, Delattre JY, Hoang-Xuan K, Sanson M. Prognostic impact of molecular markers in a series of 220 primary glioblastomas. *Cancer*. 2006; 106:2218–2223. [PubMed: 16568472]
- Ichimura K, Schmidt EE, Goike HM, Collins VP. Human glioblastomas with no alterations of the CDKN2A (p16INK4A, MTS1) and CDK4 genes have frequent mutations of the retinoblastoma gene. *Oncogene*. 1996; 13:1065–1072. [PubMed: 8806696]
- Kelly KM, Womer RB, Sorensen PH, Xiong QB, Barr FG. Common and variant gene fusions predict distinct clinical phenotypes in rhabdomyosarcoma. *J Clin Oncol*. 1997; 15:1831–1836. [PubMed: 9164192]
- Knuutila S, Bjorkqvist AM, Autio K, Tarkkanen M, Wolf M, Monni O, Szymanska J, Larramendy ML, Tapper J, Pere H, El-Rifai W, Hemmer S, Wasenius VM, Vidgren V, Zhu Y. DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. *Am J Pathol*. 1998; 152:1107–1123. [PubMed: 9588877]
- Matsunaga T, Shirasawa H, Hishiki T, Yoshida H, Kouchi K, Ohtsuka Y, Kawamura K, Etoh T, Ohnuma N. Enhanced expression of N-myc messenger RNA in neuroblastomas found by mass screening. *Clin Cancer Res*. 2000; 6:3199–3204. [PubMed: 10955804]
- Myllykangas S, Bohling T, Knuutila S. Specificity, selection and significance of gene amplifications in cancer. *Semin Cancer Biol*. 2007; 17:42–55. [PubMed: 17161620]
- Pandita A, Zielenska M, Thorner P, Bayani J, Godbout R, Greenberg M, Squire JA. Application of comparative genomic hybridization, spectral karyotyping, and microarray analysis in the identification of subtype-specific patterns of genomic changes in rhabdomyosarcoma. *Neoplasia*. 1999; 1:262–275. [PubMed: 10935481]
- Qualman SJ, Coffin CM, Newton WA, Hojo H, Triche TJ, Parham DM, Crist WM. Intergroup Rhabdomyosarcoma Study: update for pathologists. *Pediatr Dev Pathol*. 1998; 1:550–561. [PubMed: 9724344]
- Reifenberger G, Ichimura K, Reifenberger J, Elkahoul AG, Meltzer PS, Collins VP. Refined mapping of 12q13-q15 amplicons in human malignant gliomas suggests CDK4/SAS and MDM2 as independent amplification targets. *Cancer Res*. 1996; 56:5141–5145. [PubMed: 8912848]
- Ruano Y, Mollejo M, Ribalta T, Fiano C, Camacho FI, Gomez E, de Lope AR, Hernandez-Moneo JL, Martinez P, Melendez B. Identification of novel candidate target genes in amplicons of Glioblastoma multiforme tumors detected by expression and CGH microarray profiling. *Mol Cancer*. 2006; 5:39. [PubMed: 17002787]
- Schwab M. MYCN in neuronal tumours. *Cancer Lett*. 2004; 204:179–187. [PubMed: 15013217]

- Shimada S, Ishizawa T, Ishizawa K, Matsumura T, Hasegawa T, Hirose T. The value of MDM2 and CDK4 amplification levels using real-time polymerase chain reaction for the differential diagnosis of liposarcomas and their histologic mimickers. *Hum Pathol.* 2006; 37:1123–1129. [PubMed: 16938516]
- Sorensen PH, Lynch JC, Qualman SJ, Tirabosco R, Lim JF, Maurer HM, Bridge JA, Crist WM, Triche TJ, Barr FG. PAX3-FKHR and PAX7-FKHR gene fusions are prognostic indicators in alveolar rhabdomyosarcoma: a report from the children's oncology group. *J Clin Oncol.* 2002; 20:2672–2679. [PubMed: 12039929]
- Wexler, LH.; Meyer, WH.; Helman, LJ. Rhabdomyosarcoma and the undifferentiated sarcomas.. In: Pizzo, PA.; Poplack, DG., editors. *Principles and Practice of Pediatric Oncology*. Fifth ed.. Lippincott Williams & Wilkins; Philadelphia: 2006. p. 971-1001.
- Williamson D, Lu YJ, Gordon T, Sciort R, Kelsey A, Fisher C, Poremba C, Anderson J, Pritchard-Jones K, Shipley J. Relationship between MYCN copy number and expression in rhabdomyosarcomas and correlation with adverse prognosis in the alveolar subtype. *J Clin Oncol.* 2005; 23:880–888. [PubMed: 15681534]

Figure 1. Localization of minimal common amplified region in 2p24 chromosomal region (A) and 12q13.3-q14.1 chromosomal region (B)

The gray horizontal bars indicate the extent of amplification identified in each tumor using the Affymetrix GeneChip 50K array. The stippled vertical bar indicates the minimal common area of amplification based on this initial mapping data using the 50K array. The striped horizontal bars indicate the further localization of the amplicons in a selected set of tumors using the Affymetrix GeneChip 250K array. The black vertical bar indicates the final localization of the minimal common area of amplification based on this more detailed mapping analysis. At the top, the genes present in each region are shown, based on NCBI Build 36.2 of the human genetic map, and at the bottom, a scale of the corresponding genomic position (in megabases) is shown.

Figure 2. Assessment of *MYCN* (A) and *DDX1* (B) mRNA expression in ARMS cases with and without 2p24 amplification

MYCN and *DDX1* mRNA were measured by qRT-PCR in RNA isolated from ARMS samples. Values were normalized to corresponding 18S RNA expression levels and represent the ratio of the means ($\pm$ SD) of triplicate measurements. The expression of each sample is shown individually and ordered on the X-axis by expression level. The black circles indicate the cases with 2p24 amplification. The three arrows indicate the three amplified cases with high *MYCN* expression and low *DDX1* expression.

Figure 3. Assessment of mRNA expression of genes within the 12q13-q14 amplicon in ARMS cases with and without amplification

CDK4 (A), *CENTG1* (B), *DDIT3* (C), *DTX3* (D), *MBD6* (E), and *PIP5K2C* (F) mRNA were measured by qRT-PCR as described in Figure 2. The dark horizontal bar indicates the median and the lower and upper edges of the rectangles indicate the 25th and 75th percentile cutoffs of the data distribution. The bars indicate the smallest and largest values and the points above or below the bar indicate outliers in the population. Abbreviations: P3 – *PAX3-FOXO1*, P7 – *PAX7-FOXO1*, FN – fusion-negative, A – amplified, NA – non-amplified.

Figure 4. Effect of 2p24 amplification or 12q13-q14 amplification on outcome of ARMS patients
 A. Comparison of failure-free survival between all ARMS cases with and without 2p24 amplification. Amplified, n = 10; non-amplified, n = 83. B. Comparison of failure-free survival between all ARMS cases with and without 12q13-q14 amplification. Amplified, n = 11; non-amplified, n = 88. C. Comparison of failure-free survival between *PAX3-FOXO1*-positive cases with and without 12q13-q14 amplification. Amplified, n = 10; non-amplified, n = 35.

Table 1A

Contingency Table for fusion status and 2p24 amplification status.

		Fusion Status			
Amplification Status		PAX3	PAX7	Negative	p value = 0.018
	Amplified	11	5	1	
	Non-amplified	50	18	41	

Table 1B

Contingency Table for fusion status and 12q13-q14 amplification status.

		Fusion Status			
		PAX3	PAX7	Negative	
Amplification Status	Amplified	12	1	0	p value = 0.0011
	Non-amplified	39	26	31	

Table 2A

Univariate failure-free survival analysis.

Parameter	Category	Number	HR ^{a,b}	95% CI ^c	p-value
2p24 amp	No Amplification	83	1	-	0.57
	Amplified	10	1.28	0.55 - 3.02	
	2x increase	70	0.93	0.81 - 1.06	0.29
MYCN expression	No Amplification	81	1	-	0.023
	Amplified	11	2.24	1.12 - 4.82	
CDK4 expression	2x increase	60	1.40	1.02 - 1.93	0.038

Cox proportional hazards models was fitted to estimate the ^ahazard ratio (HR) and its 95% ^cconfidence interval (CI), as well as the corresponding p value.

^bFor MYCN and CDK4 expression, HR was estimated from fitting the Cox proportion hazards models with log2-scaled gene expressions

Table 2B

Univariate overall survival analysis.

Parameter	Category	Number	HR ^{a,b}	95% CI ^c	p-value
2p24 amp	No Amplification	83	1	-	0.87
	Amplified	10	0.92	0.33 - 2.58	
MYCN expression	2x increase	70	0.87	0.77 - 0.98	0.024
	No Amplification	81	1	-	0.040
CDK4 expression	Amplified	11	2.26	1.04 - 4.91	
	2x increase	60	1.33	0.95 - 1.85	0.093

Cox proportional hazards models was fitted to estimate the ^ahazard ratio (HR) and its 95% ^cconfidence interval (CI), as well as the corresponding p value.
^bFor MYCN and CDK4 expression, HR was estimated from fitting the Cox proportion hazards models with log2-scaled gene expressions

Table 3

Multivariate analysis for testing the association of 12q13-q14 amplification with failure-free survival after adjusting for clinical variables (n=136).

	Category	HR ^a	95% CI ^b	p-value
Clinical group	I/II/III	1	-	0.0012
	IV	2.233	1.33 - 3.75	
Size	≤ 5 cm	1	-	0.11
	> 5 cm	1.58	0.90 - 2.76	
Nodal status	N-0	1	-	0.42
	N-1	1.23	0.73 - 2.06	
Primary site	Other	1	-	0.0083
	Extremity	0.54	0.33 - 0.90	
12q13-q14 status	Not Amplified	1	-	0.19
	Amplified	1.653	0.78 - 3.51	

Cox proportional hazards models was fitted to estimate the ^ahazard ratio (HR) and its 95% ^bconfidence interval (CI), as well as the corresponding p value