NEOPLASIA
www.neoplasia.com

Volume 14 Number 9 September 2012 pp. 807–812 807

Retained Heterodisomy Is Associated with High Gene Expression in Hyperhaploid Inflammatory Leiomyosarcoma[1,2]

Karolin H. Nord*, Kajsa Paulsson*, Srinivas Veerla†, Johan Wejde‡, Otte Brosjö§, Nils Mandahl* and Fredrik Mertens*

*Department of Clinical Genetics, University and Regional Laboratories, Skåne University Hospital, Lund University, Lund, Sweden; †Department of Immunotechnology, Lund University Biomedical Centre, Lund, Sweden; ‡Department of Pathology, Karolinska Institute and University Hospital, Stockholm, Sweden; §Department of Orthopedics, Karolinska Institute and University Hospital, Stockholm, Sweden

Abstract

Inflammatory leiomyosarcoma (ILMS) is a soft tissue tumor that morphologically resembles conventional leiomyosarcoma (LMS) admixed with a prominent inflammatory infiltrate. Genetic data on ILMS are still limited but have suggested that this entity is characterized by hyperhaploidy (24–34 chromosomes). This low chromosome number is otherwise uncommon in neoplasia and has been found only in 0.2% to 0.3% of cytogenetically investigated tumors. Here, three ILMS were investigated using cytogenetic, single-nucleotide polymorphism (SNP) array, and global gene expression analyses. All cases displayed a hyperhaploid origin. Combined with previously reported cases, hyperhaploidy has been found in six of seven cytogenetically investigated ILMS. The copy number distribution of individual chromosomes is clearly nonrandom; the hyperhaploid clones of all six cases displayed disomy for chromosomes 5 and 20, and two copies of chromosomes 18, 21, and 22 were also common. All chromosomes identified as disomic showed a biparental origin by SNP array analysis; whether this is of pathogenetic importance is not known. Compared with conventional LMS, ILMS had a distinct gene expression signature. Furthermore, the number of chromosome copies correlated well with gene expression levels; disomic chromosomes showed higher gene expression levels than monosomic chromosomes, a finding that has not previously been reported for hyperhaploid tumors. Taken together, our findings suggest that disomy for some chromosomes, notably 5 and 20, as well as distorted gene expression achieved through massive loss of other chromosomes are essential features of ILMS.

Neoplasia (2012) 14, 807–812

Introduction

In 1995, Merchant et al. described a fascicular spindle cell tumor of the deep soft tissues that showed morphologic features in common with leiomyosarcomas (LMS) as well as a prominent inflammatory component [1]. The new entity was termed inflammatory LMS (ILMS). Some of the described tumors had originally been included in the heterogeneous group of so-called inflammatory malignant fibrous histiocytoma. However, ILMS should be distinguished from both inflammatory and pleomorphic malignant fibrous histiocytomas, which are poorly defined undifferentiated sarcomas with poor prognosis [2,3].

Abbreviations: ALL, acute lymphoblastic leukemias; FDR, false discovery rate; GSEA, gene set enrichment analysis; HCL, hierarchical clustering analysis; ILMS, inflammatory leiomyosarcoma; LMS, leiomyosarcoma; LOH, loss of heterozygosity; PCA, principal component analysis; SNP, single-nucleotide polymorphism; UPID, uniparental isodisomy

Address all correspondence to: Karolin H. Nord, PhD, Department of Clinical Genetics, Lund University, BMC B13, SE-221 84 Lund, Sweden.
E-mail: Karolin.Hansen_Nord@med.lu.se

[1]This work was supported by the Swedish Cancer Society.
[2]This article refers to supplementary materials, which are designated by Tables W1–W3 and Figure W1 and are available online at www.neoplasia.com.
Received 11 June 2012; Revised 2 August 2012; Accepted 3 August 2012

DOI 10.1593/neo.12930

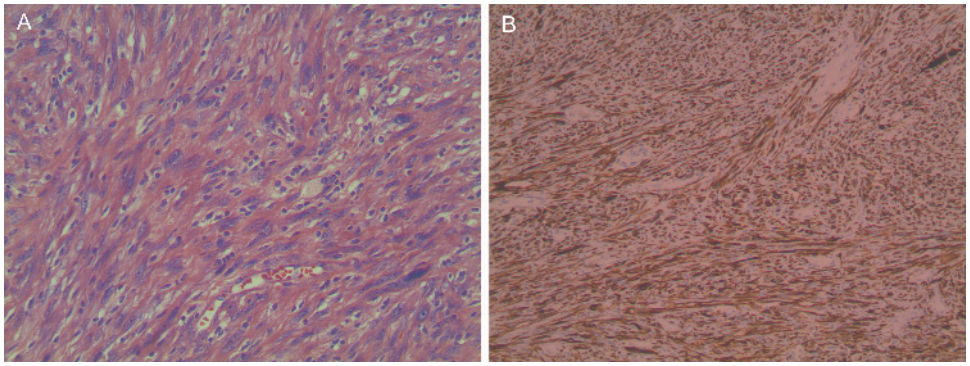

Figure 1. Morphologic and immunophenotypic appearance of an ILMS (case 2). (A) Medium-power (H&E, ×20) view showing fascicles of spindle cells admixed with inflammatory cells. (B) Strong and diffuse positivity for desmin (×10).

Cytogenetic analyses have suggested that ILMS has distinct genomic features characterized by a hyperhaploid chromosome complement, i.e., 24 to 34 chromosomes [4]. Karyotypes from four ILMS have been reported, three of which showed fewer than 35 chromosomes [4,5]. In general, this genetic pattern is uncommon in neoplasia; only 0.2% to 0.3% of cytogenetically investigated tumors have shown a chromosome number in the near-haploid range [6]. However, in certain tumor types, distinct subgroups with hyperhaploid chromosome numbers can be discerned. About 0.5% to 1% of acute lymphoblastic leukemias (ALL) and about 5% of chondrosarcomas have displayed hyperhaploid chromosome numbers. Moreover, in an unselected series of chondrosarcomas, global DNA copy number analyses using single-nucleotide polymorphism (SNP) arrays have indicated that an additional subset of tumors have gone through a hyperhaploid stage [7]. To identify genomic aberrations and determine their impact on gene transcription in ILMS, we have here combined information from cytogenetic, SNP array, and global gene expression analyses.

Materials and Methods

Patients and Tumor Material

Three tumors with histologic features compatible with ILMS were included (Figure 1 and Table 1; cases 1–3) [1]. All of them affected adult men (age range, 24–52 years). The tumors ranged in size from 6 to 9 cm and were located in the thigh (two patients) or arm. Fifteen conventional LMS (5 spindle cell and 10 not otherwise specified) were included for comparison in the global gene expression analysis (Table W1). All patients with LMS were adults (age range, 46–86 years; eight men, seven women). All samples were obtained after informed consent, and the study has been approved by the Lund University Ethics Review Board.

DNA and RNA Extractions

DNA and RNA were extracted from fresh frozen tumor biopsies using the DNeasy tissue kit including the optional RNaseH treatment and the RNeasy lipid tissue kit, according to the manufacturer's instructions (Qiagen, Hilden, Germany). Quality and concentration of the extracted material were measured with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA).

Cytogenetic Analyses

Cytogenetic analyses were performed according to standard procedures, and karyotypes were written following the recommendations of the International System for Human Cytogenetic Nomenclature (ISCN) (Tables 1 and W1) [8]. None of the 15 control LMS displayed a near-haploid karyotype; all cases showed abnormal karyotypes with near-diploid (11 cases), near-triploid (3 cases), or near-tetraploid (1 case) chromosome numbers (Table W1). The complexity of the cytogenetic changes ranged from supernumerary ring and marker chromosomes as the sole anomalies to highly aberrant karyotypes with multiple gains, losses, and structural rearrangements.

Genomic Copy Number and Loss of Heterozygosity Analyses

SNP array analysis was used for combined DNA copy number and loss of heterozygosity (LOH) investigation. Tumor DNA was hybridized

Table 1. Clinical and Cytogenetic Features of ILMS.

Case	Age/Sex	Location	Grade (1–3)	Size (cm)	Follow-up (months)*	Karyotype†	Reference‡
1	52/M	Lower arm	2	8	36 NED	51-53,X,−Y,+X,+5,+5,+18,+18,+20,+20,+21,+21,+22,+22	Present study
2	41/M	Thigh	?	9	LTF	26,X,+5,+20,+22/26,idem,r(?15)/52,idemx2/51-52,idemx2,r(?15)x2/102-104,idemx4,inc	Present study
3	24/M	Thigh	2	6	120 NED	27,X,+Y,+5,+der(20)t(X;20)(p2?1;q13.3),+22/27-28,idem,+r/54,idemx2	Present study
4	53/M	Thigh	III§	3	18 NED	28,X,+5,+18,+20,+21,+22/56,idemx2	7758; case 4
5	20/M	Thigh	NA	4	NA	28,X,+5,+18,+20,+21,+22/29,idem,+r	7758; case 5
6	64/F	Thigh	NA	12	22 M 41 NED	25-28,add(X)(q26),add(3)(q29),+5,+8,+20,+21,+1-2r/47-53,add(X)x2,+4,+5,+5,+20,+20,+21,+21	11143; case 1
7	32/M	Lung	NA	3.5	NA	46,XY,del(8)(p21p23)/46,XY,del(9)(q11)	11143; case 3

*NED, no evidence of disease; LTF, lost to follow-up; NA, not available; M, metastasis.
†The karyotypes are based on information from G-banding and, in case 3, fluorescence *in situ* hybridization analyses.
‡Reference numbers correspond to publications in the Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (2012; Mitelman F, Johansson B and Mertens F (Eds.), http://cgap.nci.nih.gov/Chromosomes/Mitelman).
§III, grade 3 in a four-grade malignancy scale; NA, not available.

onto Illumina Human CNV370 version 1.0 and Illumina Human CNV370-Quad version 3.0 BeadChips, containing ~370,000 markers, following standard protocols supplied by the manufacturer (Illumina, San Diego, CA). Data analysis was performed using the GenomeStudio software (Illumina), detecting imbalances by visual inspection. Constitutional copy number variations were excluded through comparison with the Database of Genomic Variants (http://projects.tcag.ca/variation/) [9].

Global Gene Expression Profiling

Global gene expression analyses were performed using Affymetrix Human Gene 1.0 ST Arrays according to the manufacturer's instructions (Affymetrix, Santa Clara, CA). Included in the analyses were 3 ILMS and 15 control LMS. The control tumors were selected on the basis of their histologic resemblance to ILMS and the fact that well-differentiated LMS previously have shown a distinct gene expression signature, separate from the group of poorly differentiated pleomorphic sarcomas formerly known as "malignant fibrous histiocytomas" [10]. In addition, the control tumors did not show any cytogenetic sign of a hyperhaploid origin (Table W1). Gene expression data were normalized, background-corrected, and summarized by using the Robust Multichip Analysis (RMA) algorithm implemented in the Expression Console version 1.1 software (Affymetrix). To reduce batch effects, we adjusted normalized data using the ComBat algorithm (Figure W1) [11]. Correlation-based principal component analysis (PCA) and hierarchical clustering analysis were performed using the Qlucore Omics Explorer version 2.2 (Qlucore AB, Lund, Sweden). Differences between tumor groups in $\log_2$ transformed expression data were calculated using a *t* test, and corrections for multiple testing were based on the Benjamini-Hochberg method (Qlucore AB). Genes with $P < .05$ and a false discovery rate (FDR) < 0.2 were considered significantly altered. Microarray data have been deposited in the Gene Expression Omnibus database (www.ncbi.nlm.nih.gov/geo, Accession No. GSE38104).

Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) can be used to calculate whether a defined set of genes shows statistically significant differences between two groups [12,13]. Here, the method was applied to the RMA-normalized and ComBat-adjusted gene expression data described above using gene sets corresponding to each chromosome, except the sex chromosomes. The chromosomal location of each gene was extracted from the annotation file HuGene-1_0-st-v1_na27_hg18 (Affymetrix). GSEA was performed using default settings for adjustable parameters, except for the size of the gene sets, which was adjusted to include gene sets for all autosomes, and the permutation type; *P* values were estimated by permuting the genes.

Correlation between Gene Expression Levels and Number of Chromosome Copies

The mean expression value for each gene was calculated in the groups of ILMS and LMS, respectively. The normalized intensity value in ILMS was subsequently divided with the corresponding value in LMS. Genes with a ratio >1 thus showed a higher expression in ILMS than in LMS. Each chromosome contained between 268 and 2242 genes, and for each chromosome, the mean and median gene expression ratios were extracted. The mean gene expression values were subsequently correlated with the relative number of chromosome copies.

Results

Both Parental Copies of Chromosomes 5, 20, and 22 Are Retained, whereas the Remaining Autosomes Most Often Display LOH

Case 1 showed a hyperdiploid karyotype with two copies of all autosomes except 5, 18, 20, 21, and 22, which were all present in four copies (Table 1). These five chromosomes displayed relative gain and retained heterozygosity by SNP array analysis and all other autosomes showed LOH (Table 2). The karyotype and the SNP array results strongly suggest that the hyperdiploid clone detected by G-banding analysis originated from a hyperhaploid clone. Case 2 showed relative gains of chromosomes 5, 20, and 22 in clones at different ploidy levels ranging from hyperhaploidy to hypertetraploidy (Table 1). No other gains or losses were detected by SNP array analysis (Table 2). By G-banding analysis, some clones displayed a ring chromosome possibly consisting of chromosome 15 material. Case 3 displayed a karyotype that agreed well with the relative gain and retained heterozygosity for chromosomes 5, 20, and 22 revealed by SNP array analysis (Tables 1 and 2). The SNP array data also showed relative gain and retained heterozygosity for chromosome 21 in a subclone of the investigated material.

In summary, all three cases presented hyperhaploid clones and/or duplicates of such clones. Numerical imbalances involved losses and/or gains of whole chromosomes; no small imbalance was detected. Structural rearrangements were confined to a translocation in case 3 and ring chromosome formation in subclones of cases 2 and 3. Both parental copies of chromosomes 5, 20, and 22 were present in all three cases. Retained heterozygosity was also detected for chromosome 18 in case 1 and chromosome 21 in cases 1 and 3. All other autosomes displayed LOH as a result of chromosome loss and, in some cases, subsequent duplication of the remaining homologue giving rise to uniparental isodisomy.

ILMS Has a Distinct Gene Expression Profile

Unsupervised PCA demonstrated clearly distinct gene expression profiles for the ILMS compared with the LMS controls (Figure 2). On

Table 2. SNP Array Findings in Three ILMS*.

Chromosome	Case 1	Case 2	Case 3
1	LOH	LOH	LOH
2	LOH	LOH	LOH
3	LOH	LOH	LOH
4	LOH	LOH	LOH
5	RH	RH	RH
6	LOH	LOH	LOH
7	LOH	LOH	LOH
8	LOH	LOH	LOH
9	LOH	LOH	LOH
10	LOH	LOH	LOH
11	LOH	LOH	LOH
12	LOH	LOH	LOH
13	LOH	LOH	LOH
14	LOH	LOH	LOH
15	LOH	LOH	LOH
16	LOH	LOH	LOH
17	LOH	LOH	LOH
18	RH	LOH	LOH
19	LOH	LOH	LOH
20	RH	RH	RH
21	RH	LOH	RH†
22	RH	RH	RH

*RH, relative gain of chromosomes with retained heterozygosity.
†Relative gain and retained heterozygosity for chromosome 21 were detected in a subclone.

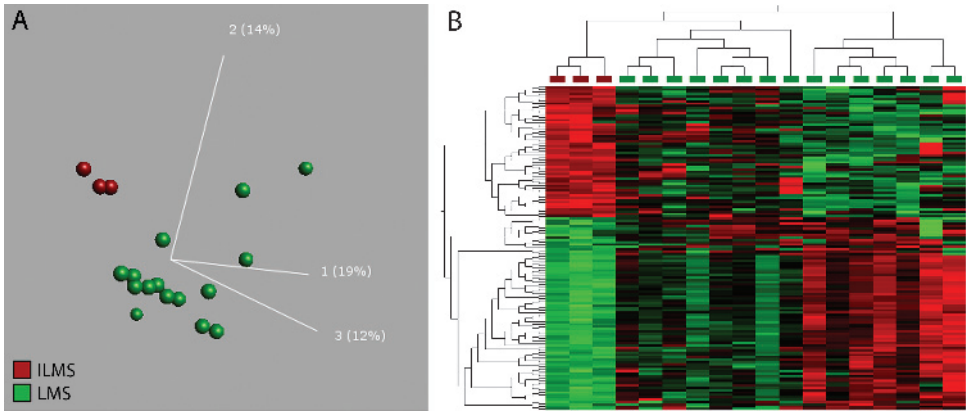

Figure 2. ILMS shows distinct gene expression features. (A) Unsupervised PCA based on the expression of the 989 most variable genes ($\sigma/\sigma_{max} = 0.4$) shows that the three ILMS (red) form a group that has an expression profile separate from the 15 LMS (green). The first three principal components, representing 19%, 14%, and 12% of the variance, are displayed. Of the 989 most variable genes, 138 genes showed a significantly different expression between the two tumor groups ($P < .05$, FDR < 0.2; Tables W2 and W3). (B) The differentially expressed genes are displayed in a supervised hierarchical clustering analysis. Genes with high and low expression values are labeled in red and green, respectively.

the basis of the expression of the 989 most variable genes ($\sigma/\sigma_{max} = 0.4$), 138 genes showed significantly different expression values between the two groups; 56 genes showed higher and 82 genes showed lower expression levels in ILMS compared with LMS (Figure 2 and Tables W2 and W3).

Genes on the Disomic Chromosomes 5, 20, and 22 Show High Expression Levels in ILMS

Using gene sets representing each autosome, GSEA showed a good correlation between transcriptional levels and gene copy numbers. Transcripts from chromosomes 5, 20, 21, and 22 were found to be

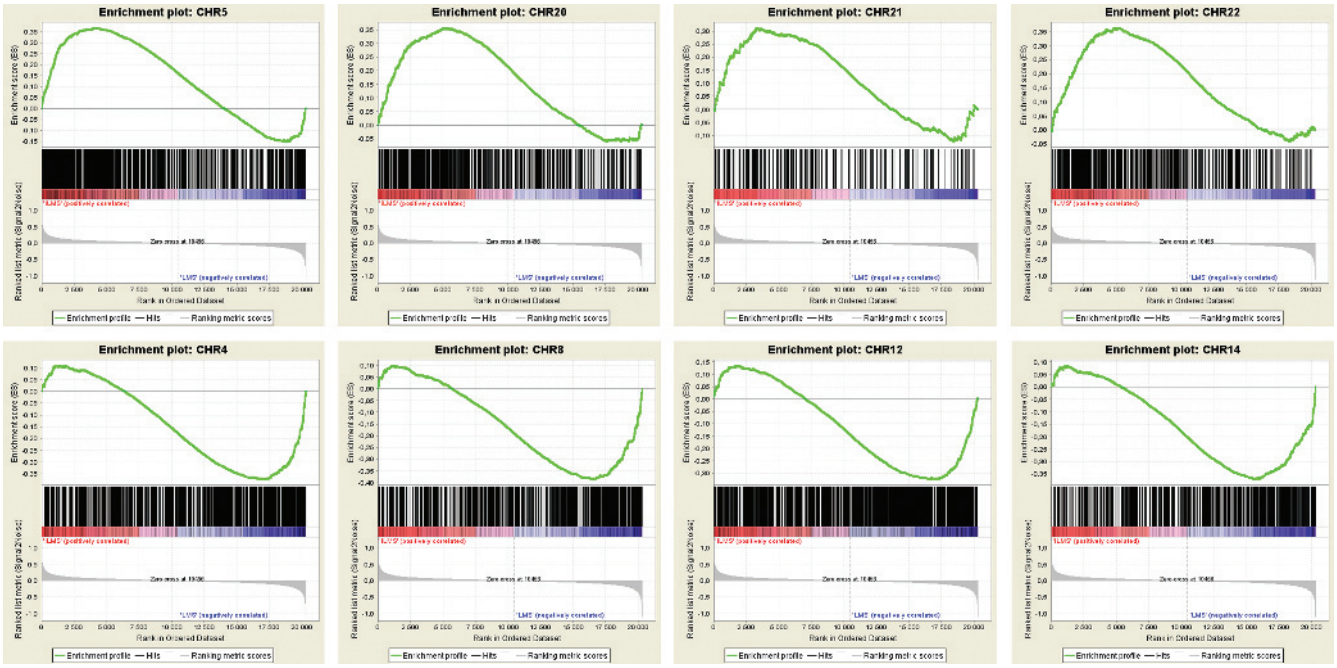

Figure 3. Gene sets for chromosomes 5, 20, and 22 are enriched in ILMS. GSEA was performed using gene sets corresponding to each autosome. Genes on chromosomes 5, 20, 21, and 22 show enrichment in ILMS, whereas genes on chromosomes 4, 6, 8, 12, 14, and 15 show the opposite pattern, i.e., enrichment in the control group of conventional LMS ($P < .01$ and FDR < 0.01). A ranked gene list, represented by the red and blue heat maps, is created on the basis of the gene expression values in the two groups. The black lines correspond to the location of the genes from the respective gene set in the sorted lists. The enrichment profiles show the running sum (green line) for the particular gene set. This value is calculated by walking down the ranked gene list and increasing the score when encountering a gene in the gene set and decreasing when encountering genes not in the gene set [12]. The enrichment profiles thus show if the members of a gene set are randomly distributed throughout the ranked gene list or primarily found at the top or bottom. A gene set with the latter distribution is expected to be related to the phenotype. The images were extracted from the GSEA software [12,13].

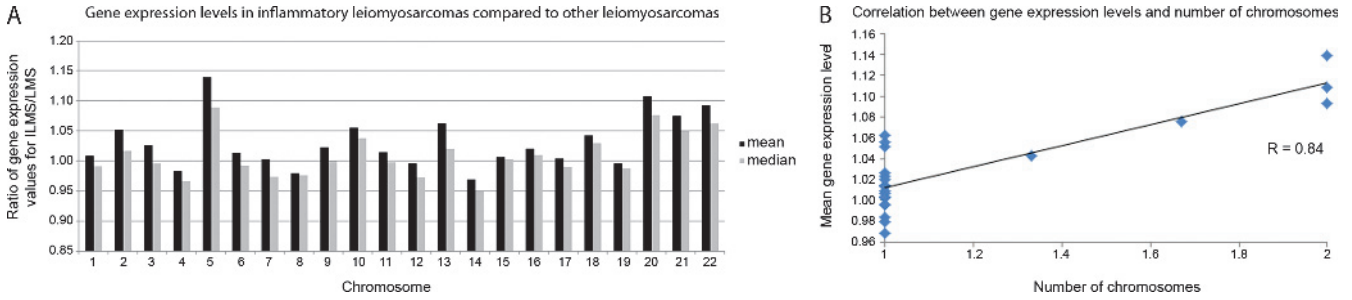

Figure 4. Heterodisomic chromosomes 5, 20, and 22 display high average gene expression levels. (A) The mean expression value was calculated for each gene in ILMS and conventional LMS, respectively. The normalized intensity value in ILMS was divided by the corresponding value in LMS. Genes with a ratio >1 thus showed a higher expression in ILMS. For each chromosome, containing between 268 and 2242 genes, the average gene expression level was estimated by calculating the mean (black) and median (gray) gene ratios. The highest average gene expression ratios were detected for chromosomes 5, 20, and 22. (B) The relative number of chromosome copies (2 for chromosomes 5, 20, and 22, 1.67 for chromosome 21, 1.33 for chromosome 18, and 1 for the remaining autosomes) correlated well with the average gene expression level of the respective chromosomes ($R = 0.84$).

enriched in ILMS, whereas transcripts from chromosomes 4, 6, 8, 12, 14, and 15 were enriched in LMS ($P < .01$ and FDR < 0.01; Figure 3). The average gene expression levels per chromosome, calculated using normalized intensity values, were compared between the tumor groups, and the highest ratios (ILMS/LMS) were detected for chromosomes 5, 20, and 22 (Figure 4*A*). The relative number of chromosome copies, compared to the haploid state, was 2 for chromosomes 5, 20, and 22, 1.67 for chromosome 21, 1.33 for chromosome 18, and 1 for the remaining autosomes (Table 2). Correlation analysis showed a good correlation between average gene expression level and number of chromosome copies ($R = 0.84$; Figure 4*B*).

Discussion

Five of seven cytogenetically investigated ILMS have shown a hyperhaploid chromosome number (Table 1). In addition, a sixth case (case 1 of the present study) showed a hyperdiploid karyotype containing exclusively disomic and tetrasomic chromosomes. The tetrasomic chromosomes consisted of two maternal and two paternal homologues, and all disomic chromosomes were uniparental isodisomies. This pattern of aberrations strongly indicates a hyperhaploid origin where the original clone has doubled the entire set of chromosomes, most likely trough a single event. Thus, we can confirm the previously suggested high frequency of hyperhaploidy in ILMS.

Hyperhaploidy is thought to be a relatively rare phenomenon in neoplasia. However, the incidence estimates of this extreme chromosome number are primarily based on cytogenetic analyses of cultured tumor cells [6]. Because this technique does not reveal the parental origin of the chromosomes, it is possible that a subset of tumors with higher chromosome numbers has passed through an undetected hyperhaploid stage. In line with this, still limited SNP array data on chondrosarcomas suggest that the majority of cases with more than 46 chromosomes originate from a hyperhaploid or hypodiploid precursor [7]. Thus, hyperhaploidy may be more common than suggested by cytogenetic studies.

Although the number of investigated cases remains low, it can be concluded that the copy number distribution of individual chromosomes is clearly nonrandom in ILMS; all six cases with a hyperhaploid origin have shown disomy for chromosomes 5 and 20, and two copies of chromosomes 18, 21, and 22 are also common (Table 1). In a recent review of the karyotypic patterns in hyperhaploid neoplasms, it was found that bone and soft tissue tumors displayed nonrandom patterns with respect to retained disomies; chromosomes 5 and 20 are most often disomic in soft tissue tumors and chromosomes 5, 7, and 20 are rarely monosomic in chondrosarcoma [6]. In hyperhaploid ALL, however, there is preferential retention of chromosomes 14, 18, 21, and X [14], and also other tumor types, that less frequently show hyperhaploidy seem to have nonrandom distribution of monosomies and disomies [6].

The pathogenetic impact of hyperhaploidy is not easily explained. Many hyperhaploid tumors show subpopulations of hyperdiploid to hypertetraploid clones, which sometimes constitute the majority of investigated cells. It thus seems unlikely that it is the low chromosome number as such that is crucial for tumor development. More likely, the combination of lost and retained chromosomes is important. One potential pathogenetic mechanism is that the widespread LOH found in hyperhaploid tumors is vital for tumorigenesis by unmasking recessive mutations. The explanation for the fact that both parental copies of some chromosomes are consistently retained could then be that loss of any of these chromosomes would be lethal for the tumor cells. However, it seems unlikely that the pathogenetic consequences of deleting almost half of the genome would be to reveal a few potentially deleterious mutations. In this context, it could be of interest to note that none of the present tumors showed microdeletions, which is in contrast to hyperhaploid chondrosarcoma and ALL [7,15]. Therefore, perhaps a more likely possibility would be that hyperhaploidy results in multiple gene dosage imbalances that distort the global gene expression profile. In support of this hypothesis, there was a clear correlation between gene copy number and level of expression; genes on disomic chromosome were enriched in ILMS and disomic chromosomes showed higher average gene expression levels than monosomic chromosomes. Thus, extensive allelic loss through whole chromosome deletions coupled with retention of other chromosomes may favor the impact of dosage-sensitive genes [16,17].

It is conspicuous that chromosomes 5, 20, and 22 were heterodisomic in all cases investigated and that nonrandom patterns of retained heterodisomy have been reported also in chondrosarcoma and ALL [7,15]. Obviously, disomy for certain tumor-specific chromosomes is of importance for hyperhaploid tumor development. However, it is not known whether the preservation of both parental homologues, as in contrast to isodisomy, is a consequence of the mechanism by

which the hyperhaploid cells originate or if it is of pathogenetic significance. In normal diploid cells, some genes are functionally hemizygous because of parental imprinting, a silencing mechanism that results in directed transcription of an allele from only one parent [18]. Although the exact prevalence is not known, it has been estimated that around 1% of human genes is affected by parental imprinting, and by computational analysis, these genes have been predicted to be present in all autosomes [19]. Furthermore, monoallelic gene expression has also been shown to occur because of random inactivation of maternal or paternal alleles, i.e., some cells express the maternal allele and others express the paternal allele; a pattern of expression that is maintained in daughter cells [20]. It is thus possible that the nonrandom pattern of chromosome aberrations found in hyperhaploid tumors is influenced by monoallelic gene transcription.

In conclusion, we confirm that hyperhaploidy is a distinguishing characteristic of ILMS and show that disomy for some chromosomes, notably 5 and 20, as well as distorted gene expression achieved through massive loss of other chromosomes are essential features of ILMS.

Acknowledgments

We thank Jenny Nilsson and Anna Johnson for technical assistance and Swegene Centre for Integrative Biology, Lund University for the help with the microarray analyses.

References

[1] Merchant W, Calonje E, and Fletcher CDM (1995). Inflammatory leiomyosarcoma: a morphological subgroup within the heterogeneous family of so-called inflammatory malignant fibrous histiocytoma. *Histopathology* **27**, 525–532.

[2] Fletcher CDM, van den Berg E, and Molenaar WM (2002). Pleomorphic malignant fibrous histiocytoma/undifferentiated high grade pleomorphic sarcoma. In *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Soft Tissue and Bone.* CDM Fletcher, KK Unni, and F Mertens (Eds). IARC Press, Lyon, France. pp. 120–122.

[3] Coindre JM (2002). Inflammatory malignant fibrous histiocytoma/undifferentiated pleomorphic sarcoma with prominent inflammation. In *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Soft Tissue and Bone.* CDM Fletcher, KK Unni, and F Mertens (Eds). IARC Press, Lyon, France. p. 125.

[4] Dal Cin P, Sciot R, Fletcher CDM, Samson I, De Vos R, Mandahl N, Willén H, Larsson O, and Van den Berghe H (1998). Inflammatory leiomyosarcoma may be characterized by specific near-haploid chromosome changes. *J Pathol* **185**, 112–115.

[5] Chang A, Schuetze SM, Conrad EU III, Swisshelm KL, Norwood TH, and Rubin BP (2005). So-called "inflammatory leiomyosarcoma": a series of 3 cases providing additional insights into a rare entity. *Int J Surg Pathol* **13**, 185–195.

[6] Mandahl N, Johansson B, Mertens F, and Mitelman F (2012). Disease-associated patterns of disomic chromosomes in hyperhaploid neoplasms. *Genes Chromosomes Cancer* **51**, 536–544.

[7] Olsson L, Paulsson K, Bovée JVMG, and Nord KH (2011). Clonal evolution through loss of chromosomes and subsequent polyploidization in chondrosarcoma. *PLoS One* **6**, e24977.

[8] Shaffer LG, Slovak ML, and Campbell LJ (Eds.) (2009). ISCN (2009): an International System for Human Cytogenetic Nomenclature. S. Karger, Basel.

[9] Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, and Lee C (2004). Detection of large-scale variation in the human genome. *Nat Genet* **36**, 949–951.

[10] Gibault L, Pérot G, Chibon F, Bonnin S, Lagarde P, Terrier P, Coindre JM, and Aurias A (2011). New insights in sarcoma oncogenesis: a comprehensive analysis of a large series of 160 soft tissue sarcomas with complex genomics. *J Pathol* **223**, 64–71.

[11] Johnson WE, Li C, and Rabinovic A (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127.

[12] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550.

[13] Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E, et al. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267–273.

[14] Harrison CJ, Moorman AV, Broadfield ZJ, Cheung KL, Harris RL, Reza Jalali G, Robinson HM, Barber KE, Richards SM, Mitchell CD, et al. (2004). Three distinct subgroups of hypodiploidy in acute lymphoblastic leukaemia. *Br J Haematol* **125**, 552–559.

[15] Safavi S, Forestier E, Golovleva I, Barbany G, Nord KH, Moorman AV, Harrison CJ, Johansson B, and Paulsson K (2012). Loss of chromosomes is the primary event in near-haploid and low hypodiploid acute lymphoblastic leukemia. *Leukemia.* DOI: 10.1038/leu.2012.227, E-pub ahead of print.

[16] Veitia RA, Bottani S, and Birchler JA (2008). Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. *Trends Genet* **24**, 390–397.

[17] Edger PP and Pires JC (2009). Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. *Chromosome Res* **17**, 699–717.

[18] Reik W and Walter J (2001). Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2**, 21–32.

[19] Luedi PP, Dietrich FS, Weidman JR, Bosko JM, Jirtle RL, and Hartemink AJ (2007). Computational and experimental identification of novel human imprinted genes. *Genome Res* **17**, 1723–1730.

[20] Gimelbrant A, Hutchinson JN, Thompson BR, and Chess A (2007). Widespread monoallelic expression on human autosomes. *Science* **318**, 1136–1140.

Table W1. Cytogenetic Features of LMS.

Case	Diagnosis*	Publication†	Karyotype‡
8	LMS, NOS	Ref. 11719, case 13	47-49,XX,+1-2r,+mar
9	LMS, NOS	Ref. 10805, case 35	43-47,X?,+1-4r,inc/83-90,idemx2
10	LMS, NOS		45-47,X,−Y/45-47,idem,+r/47-50,XY,+1-3r
11	LMS, NOS		41-43,XY,del(1)(q11),−2,−6,−9,−10,−12,−13,−17,−18,+r,+3mar/79-85,idemx2,inc
12	LMS, NOS		43-47,XY,−5,+2mar/92-93,idemx2
13	LMS, NOS		39-49,XY,−3,+2-6r/85-92,XXYY,+5-9r,inc
14	LMS, NOS		86-99,X?,add(1)(q21),?del(1)(p11),add(11)(q23), add(19)(q13),+2-6r,inc
15	LMS, NOS		44-48,XX,t(9;13)(q33;q21),+0-3r,+0-2mar/89-92,idemx2
16	LMS, NOS		46,XY,del(3)(q25),−7,add(14)(q32),+r
17	LMS, NOS		61-73,X?,add(1)(p11),ins(1;?)(p13;?),del(7)(q22), del(7)(q31),+3-4r,inc
18	LMS, spindle cell	Ref. 8300, case 3	44,X,add(X)(p22),−10,−13,add(15)(q15),add(16)(q24),der(16)t(15;16)(q15;p13), add(17)(p11),−22, +r/43,idem,−2/45,idem,+8
19	LMS, spindle cell	Ref. 8300, case 2	67-75,X,add(X)(q28),del(X)(q23),der(1;?)inv(1)(p12p36)dic(1;?)(p12;?)x2,add (3)(p14),del(3)(p14), del(3)(q12),add(4)(p16)x2,add(5)(q35),del(6)(q15)x2, der(7)del(7)(p13)ins(7;?)(q32;?),ins(7;?)(q32;?), inv(7)(p15p22),del(8)(p21), add(10)(p15),del(11)(p12)x2,del(11)(q11),add(12)(q24),add(13)(p11), der (14)add(14)(p11)add(14)(q32)x3,add(16)(p13)x2,add(17)(q25),der(18)add(18) (p11)add(18)(q23), add(19)(p13)x2,add(22)(p11),+hsr(?),inc
20	LMS, spindle cell	Ref. 5609, case 40; 8300, case 1	45-51,XY,der(1)del(1)(p36)add(1)(q32),add(3)(p25),add(7)(p11),del(7) (q11q22),add(12)(q24), add(17)(q25),add(19)(q13),der(19)t(5;19) (q13;p11)/45-49,idem,−add(3),+del(3)(q11)/81-97,idemx2, −add(3)x2,+del (3)x2/42-44,idem,add(1)(q32),del(4)(p14),add(5)(p15),−del(7),−add(19),−der (19)
21	LMS, spindle cell	Ref. 8300, case 5	40-42,XY,der(1)del(1)(p32)add(1)(q44),add(3)(q11),der(3)del(3)(p13p23)t (3;12)(q29;q13),−4,−5,−5, −9,−9,del(12)(q13),−13,−14,−14,der(15)t(15;15) (p11;q13),der(16)t(12;16)(q13;q24)ins(16;?)(q24;?), −17,−17,−18,+der(?)t (?;11)(?;q14)hsr(11)(q14),+5mar/70-83,X?,del(1)(p12),der(1)add(1)(p11)add(1) (q42-44)x2,add(3)(p21),add(12) (p13),del(12)(q13),der(15)t(15;15),inc
22	LMS, spindle cell	Ref. 8300, case 6	75-78,XXX,+X,+1,der(1;6)(q10;p10)x2,−2,+3,−4,+5,+6,+7,+8,+9,−10,+11,+12,−13, add(15)(p13)x2, +16,+18,+19,+20,der(20)t(4;20)(q11;q13)ins(20;?)(q13;?)x2, +21,+22/76-77,idem,add(17)(q25)

*NOS, not otherwise specified.
†Reference numbers are from the Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (2012).
‡Karyotypes are described according to ISCN (2009) [8].

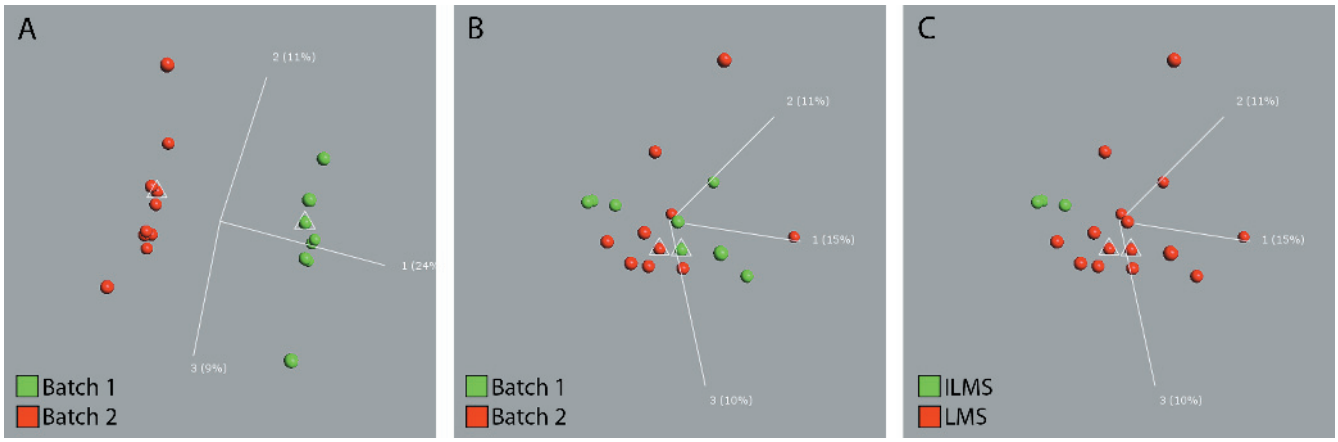

Figure W1. The ComBat algorithm was applied to adjust for batch effects when combining the two batches of global gene expression data. (A) Unsupervised PCA of RMA-normalized global gene expression data showed that the samples clustered according to batch rather than in line with any biologic feature. One sample (marked by a triangle) was analyzed in both batches. This case clustered more closely together with samples of the same batch than to its replica in the other batch. (B) After data adjustments using the ComBat algorithm, the obvious batch effects were eliminated. The duplicates of the same sample now clustered close to each other. (C) An apparent difference in global gene expression patterns between ILMS and other LMS could be discerned. The first three principal components are shown in each plot, and the proportions of the variance represented by each component are displayed within parentheses.