

Research article

Open Access

Combined subtractive cDNA cloning and array CGH: an efficient approach for identification of overexpressed genes in DNA amplicons

Katleen De Preter¹, Filip Pattyn¹, Geert Berx², Kristin Strumane², Björn Menten¹, Frans Van Roy², Anne De Paepe¹, Frank Speleman¹ and Jo Vandesompele*

Address: ¹Center for Medical Genetics, Ghent University Hospital 1K5, De Pintelaan 185, 9000 Gent, Belgium and ²Department for Molecular Biomedical Research, Flemish Interuniversity Institute for Biotechnology (VIB), Ghent University, Technologiepark 927, 9052 Zwijnaarde, Belgium

Email: Katleen De Preter - katleen.depreter@ugent.be; Filip Pattyn - filip.pattyn@ugent.be; Geert Berx - geert.berx@dmbr.ugent.be; Kristin Strumane - kristin.strumane@dmbr.ugent.be; Björn Menten - bjorn.menten@ugent.be; Frans Van Roy - frans.vanroy@dmbr.ugent.be; Anne De Paepe - anne.depaepe@ugent.be; Frank Speleman - franki.speleman@ugent.be; Jo Vandesompele* - joke.vandesompele@ugent.be

* Corresponding author

Published: 03 February 2004

Received: 03 October 2003

BMC Genomics 2004, **5**:11

Accepted: 03 February 2004

This article is available from: <http://www.biomedcentral.com/1471-2164/5/11>

© 2004 De Preter et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Activation of proto-oncogenes by DNA amplification is an important mechanism in the development and maintenance of cancer cells. Until recently, identification of the targeted genes relied on labour intensive and time consuming positional cloning methods. In this study, we outline a straightforward and efficient strategy for fast and comprehensive cloning of amplified and overexpressed genes.

Results: As a proof of principle, we analyzed neuroblastoma cell line IMR-32, with at least two amplification sites along the short arm of chromosome 2. In a first step, overexpressed cDNA clones were isolated using a PCR based subtractive cloning method. Subsequent deposition of these clones on a custom microarray and hybridization with IMR-32 DNA, resulted in the identification of clones that were overexpressed due to gene amplification. Using this approach, amplification of all previously reported amplified genes in this cell line was detected. Furthermore, four additional clones were found to be amplified, including the *TEM8* gene on 2p13.3, two anonymous transcripts, and a fusion transcript, resulting from 2p13.3 and 2p24.3 fused sequences.

Conclusions: The combinatorial strategy of subtractive cDNA cloning and array CGH analysis allows comprehensive amplicon dissection, which opens perspectives for improved identification of hitherto unknown targeted oncogenes in cancer cells.

Background

Human cancers frequently manifest amplification of large stretches of DNA, cytogenetically detectable as homogeneously staining regions (HSR) or double minute chromatin bodies (dmin). DNA amplification is considered to be a consequence of the intrinsic genomic instability of cancer cells, and it is presumed that overexpression of a single or few amplified genes confers a selective advantage on these HSR or dmin bearing clones. Consequently, activation of proto-oncogenes by amplification is thought to play an important role in the development and maintenance of many human solid tumours [1-3]. Detection of amplification of many different chromosomal regions in various tumour types has led to the identification of the targeted oncogenes and greatly contributed to our current understanding of the genetic basis of cancer. Furthermore, amplified genes often act as markers of tumour behaviour, drug response, patient outcome, and may represent targets for future molecular cancer therapy.

In the past, various strategies have been used for the detection of amplified chromosomal regions and DNA sequences in cancer. Comparative genomic hybridization (CGH) [4] has been particularly useful for detection of amplified sequences and assignment of the chromosomal position [5]. This approach allows whole genome screening for chromosomal imbalances up to 5–10 Mb and gene amplification of sufficiently large amplicons and/or highly overrepresented regions. Due to this limited resolution, a time consuming mapping is required following amplicon identification in order to pinpoint the putative target genes. Recently, two methods were introduced that allow mapping of the genomic content of amplicons with a 10–100 fold increased resolution. Array CGH employs arrayed fragments of genomic DNA clones (with partial or complete sequence information) instead of metaphase chromosomes [6,7]. Digital karyotyping is a SAGE (serial analysis of gene expression) based method to enumerate genomic DNA tags [8]. This method allows identification of specific amplifications and deletions that were not previously detected by conventional CGH or other methods. An important limitation of both methods is the inability to directly identify the overexpressed genes that are targeted by the amplification. This limitation was overcome by another variant on the classic CGH approach in which the normal metaphase chromosomes were replaced by a large number of microarrayed cDNA clones [9]. This approach has the advantage that losses and gains are mapped by their gene position rather than chromosomal band (as with conventional CGH) or genomic position (as with array CGH and digital karyotyping). The analysis immediately provides a list of candidate genes that occur within the region of interest. Another advantage is the ability to perform expression profiling on the same slides using the cDNA microarray approach, which enables the

investigator to correlate copy number and gene expression, in order to identify candidate oncogenes that are both amplified and overexpressed. Moreover, the small size and large number of the arrayed cDNA clones provide a higher resolution in contrast to current PAC and BAC arrays for which the resolution is limited because of the relatively large size of the clones (120–200 kb). Two limitations of the cDNA array CGH approach are the confined analysis of genes that are present on the array and the analytical challenges in terms of sensitivity by the complexity of the probe and the small sizes of the arrayed target cDNAs (0.5–2 kb) (signal intensities in genomic hybridizations being proportional to the length of the target DNA).

In this paper, we propose a fast and straightforward approach to identify overexpressed genes in amplified regions, enabling direct identification of the relevant targeted oncogene(s). The approach is based on the above-mentioned cDNA array CGH method, but includes a preceding selection of differentially expressed genes. In a first step, subtractive cDNA cloning is performed on an amplified tumour sample to isolate cDNA clones that are overexpressed. Subsequent CGH analysis on a cDNA microarray containing the subtracted clones allows detection of differentially expressed genes which are amplified at the DNA level. As a proof of principle, neuroblastoma cell line IMR-32 with at least two amplification sites along the short arm of chromosome 2 (including the *MYCN* locus) was used as a model system [10,11].

In this study, the combinatorial power and efficacy of subtractive cDNA cloning and high-throughput DNA copy number determination using array CGH was demonstrated for identification of amplified and overexpressed genes. In addition to those genes which were already known to be amplified and overexpressed in IMR-32, we also detected hitherto unknown genes, which were not previously described to be amplified in neuroblastoma.

Results

Identification of differentially expressed genes by suppression subtractive hybridisation (SSH)

In a first step, overexpressed genes in neuroblastoma cell line IMR-32 were isolated (of which some have increased expression due to amplification) through a PCR select cDNA subtraction with IMR-32 as tester and SK-N-SH as driver (the latter being a neuroblastoma cell line without DNA amplification [12]). This yielded a cDNA library of 960 clones. By comparing the unsubtracted and subtracted cDNA library for the abundance of an internal control gene *GAPD*, the enrichment was estimated to be 100 fold. Upon hybridization of the subtracted and reverse subtracted probe on nylon filters containing all subtracted clones (i.e. differential screening according to the

manufacturer), false positive clones (non-differentially expressed genes) could be identified, resulting in the retention of 281 IMR-32 overexpressed clones. After sequencing, alignment, EST contig building and UniGene database search, a non-redundant list was obtained containing 126 known genes, 22 UniGene clusters, and 10 anonymous ESTs. For each unique gene, transcript or EST contig, at least one representative clone was selected for re-arraying, insert amplification and spotting on a custom microarray.

CGH on cDNA microarray and characterization of the clones

In order to determine which of the overexpressed cDNA clones were amplified in IMR-32, array CGH analysis was performed on a custom cDNA microarray using DNA of IMR-32 as test probe and DNA of a normal male lymphoblastoid cell line as normal control probe. All clones that were found to be amplified, were located on one of the two known amplified regions on the short arm of chromosome 2 (2p13-14 and 2p24) (Figures 1 and 2). In addition to the 3 known genes on 2p24 that are frequently co-amplified in neuroblastoma (i.e. *MYCN*, *DDX1* and *NAG*), ten other partial cDNA clones on the microarray were shown to be amplified in cell line IMR-32. One clone (g6f6) was part of the *MEIS1* homeobox gene (2p14) that was recently shown also to be amplified in IMR-32 [13,14]. Two other clones (g1h7 and g8f10) belonged to the *TEM8* gene on 2p13.3. A fourth clone (g4d5) is located between the *MEIS1* and the *TEM8* gene and is part of an as yet not characterized gene. RT-PCR analysis revealed that this clone is not part of the neighbouring (not yet fully annotated) *ETAA16* gene. No homology was found for g4d5 with other EST sequences or known genes. Another clone (g10d12) is located 500 kb telomeric to *NAG* and also displayed no homology to any known sequence. Two other transcripts (g9d9 and g10e3) are probably part of the *NSE1* gene as demonstrated by alignment of the clones to *NSE1* transcript variants (Figure 2) and RT-PCR assays using a forward primer in the subtracted clone and a reverse primer in the *NSE1* gene. Two clones (g1c2, g6d4) were located in the large 150 kb intron of the 4.5 kb *NAG* sequence reported by Wimmer et al. [15] (acc. no. AF056195) between exon 4 and 5. BLAST analysis of the human EST database with exon 4 and 5 of the *NAG* gene as a query sequence failed to identify an EST clone that contained both exons. Furthermore, RT-PCR with a forward primer in exon 4 and a reverse primer in exon 7 failed to yield the expected band of 341 bp in IMR-32 and SK-N-SH. In contrast, a sharp and single band of approximately 3.5 kb was amplified. Furthermore, Northern blot analysis estimated the *NAG* transcript size to be approximately 2.5 kb longer compared to the published sequence (data not shown). These data further support the recent observation that the published

NAG gene (acc. no. AF056195) is misannotated and should contain 21 more exons between former exon 4 and 5 [16]. Hence, clones g1c2 and g6d4 (present on our cDNA array) and clone g3e7 are in fact part of the newly annotated *NAG* gene (acc. no. AF388385).

The tenth amplified clone (g2h10a) is of particular interest because one part of the sequence aligns to the *TEM8* gene on 2p13.3 and the other part aligns to a sequence in band 2p24.3 (Figure 2). The fusion nature of this clone was confirmed by RT-PCR on cell line IMR-32 using a primer in the first part of the transcript on 2p13.3 and a primer in the other part of the transcript on 2p24.3. Cloning and sequencing of the PCR product revealed that IMR-32 contains at least two different splice variants of the fusion transcript, i.e. g2h10b (acc. no. CD664535) and g2h10c (acc. no. CF384614) (splice variants are detected in the part that aligns to 2p24.3). Most splicing sites are surrounded by consensus splice site sequences (data not shown).

Confirmation of amplification status

Real-time quantitative PCR on IMR-32 was performed in order to validate the amplification status of all sequences that were catalogued as amplified by array CGH analysis: five genes that were previously reported to be amplified in IMR-32 (*MYCN*, *DDX1*, *NAG*, *NSE1* and *MEIS1*), one newly amplified gene (*TEM8*), 2 anonymous expressed sequences (g10d12 and g4d5) and 1 fusion transcript. Amplification of all these genes and clones was confirmed in IMR-32 (Table 2).

Using FISH analysis, it was demonstrated that the *MYCN*, *DDX1*, *NAG*, *MEIS1*, *NSE1* and *TEM8* genes and the g10d12 clone are present as multiple copies on all 3 known HSRs in IMR-32 (Figure 3). This suggests that the 3 HSRs originate from the same complex amplicon.

To verify whether the subtracted clones that were shown to be amplified are indeed overexpressed at the mRNA level in IMR-32, real-time quantitative RT-PCR was performed and demonstrated that all genes were highly overexpressed (range 10^1 – 10^4 fold overexpression) (Table 2). The fusion transcript was only expressed in cell line IMR-32.

Three genes were shown to be amplified in the 2p13.3-14 amplicon (of which only *MEIS1* was previously reported). To our surprise, more known genes are located between amplified clone g4d5 and *TEM8*, but those were not present in our subtracted cDNA library. To test whether our approach failed to identify these genes or whether these genes were indeed not amplified in IMR-32, we randomly selected 3 genes (*PPP3R1*, *PLEK* and *BMP10*) and

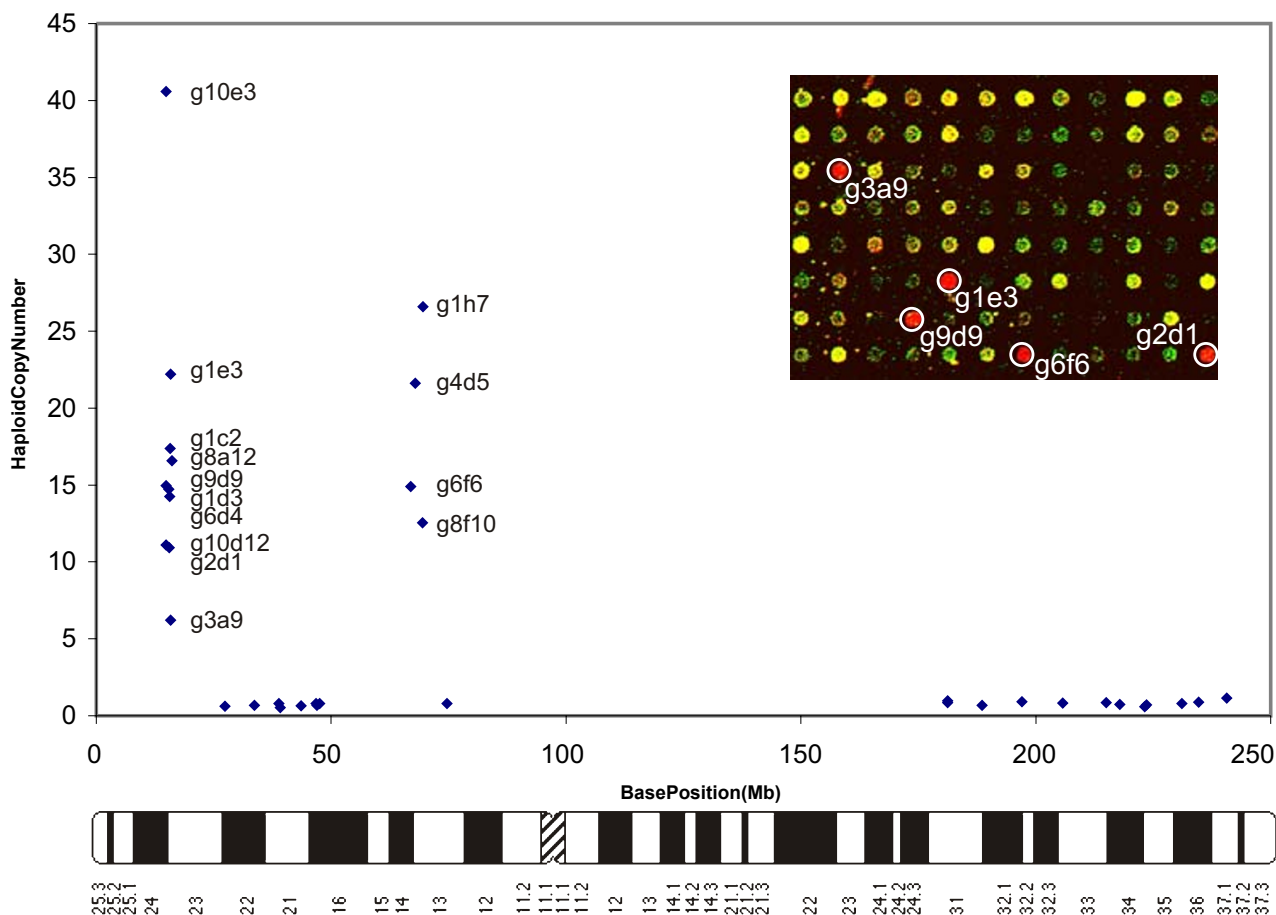


Figure 1
Array CGH based haploid copy number of SSH clones mapping on chromosome 2: Base position of the SSH clones on chromosome 2 (with exception of fusion transcript clone g2h10) was determined according to the human genome browser at UCSC (April 2003 freeze [33]). Two clear amplification sites along the short arm emerge. Insert: detail of the array CGH (IMR-32 in red and control DNA in green), amplified clones are indicated.

determined their copy number and expression level in IMR-32. Neither amplification nor overexpression could be detected for these genes, demonstrating that the 2p13.3-14 amplicon in IMR-32 is complex and discontinuous.

A recent study reported that the *DNMT3A* gene on chromosome band 2p23.3 is amplified in IMR-32 and is probably part of a third amplicon on 2p [17]. As our approach did not identify this gene, we decided to evaluate the *DNMT3A* gene copy number and expression level with real-time quantitative PCR. Neither amplification nor overexpression could be detected in cell line IMR-32.

Extended gene copy number and mRNA expression analysis of the novel amplified genes in a panel of neuroblastoma cell lines

Real-time quantitative PCR was performed in order to analyse the mRNA expression level and gene copy number of novel amplified genes *TEM8*, *g10d12*, *g10e3*, and *g4d5*, and already known amplified genes *MYCN*, *DDX1*, *NAG* and *MEIS1* in 30 NB cell lines and 9 normal human tissue samples (Table 3 and Figure 4). These analyses showed that *g10e3* and *g4d5* were only amplified and overexpressed in cell line IMR-32. Clone *g10d12* was also found to be amplified and overexpressed in cell line SJNB-6. Subsequent gene copy number determination of *g10d12* in primary tumour samples indicated a co-amplification frequency with *MYCN* of 12 % (9/75 tested *MYCN* ampli-

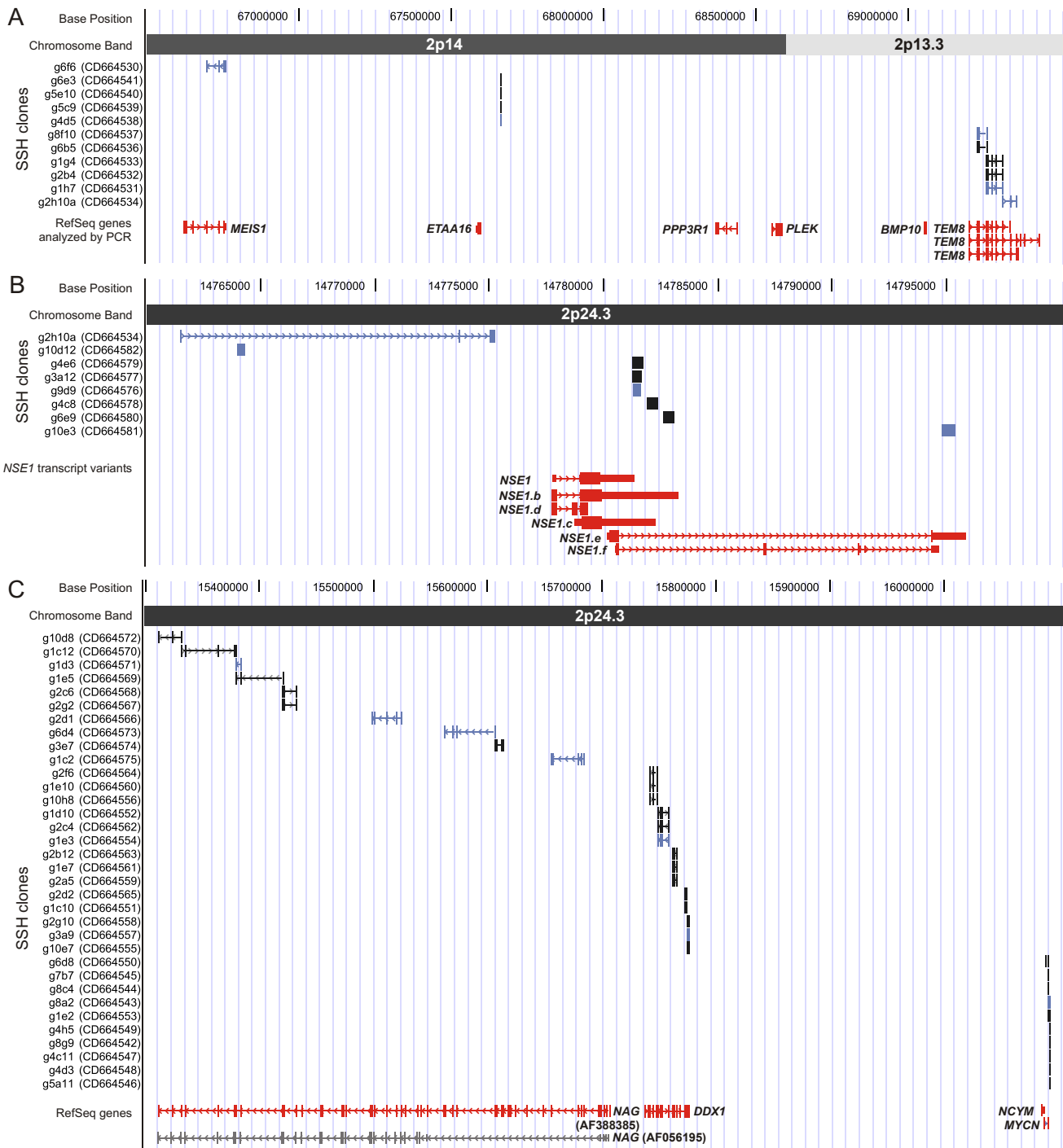


Figure 2
Genomic position of known genes and 2p amplified SSH clones: These results were obtained by a human BLAT search (UCSC genome browser, April 2003 freeze [33]) (clones that were present on the microarray are marked in blue; RefSeq genes are marked in red and the initially misannotated gene NAG in grey). A: amplicon on chromosome band 2q13.3-14; B and C: amplicon on chromosome band 2p24.3 (acc. no. of SSH clone sequences between brackets).

Table 1: Oligonucleotide sequences for real-time PCR based DNA copy number determination (GCN) and gene expression analysis (GXP): Primer sequences for known genes are submitted to RTPrimerDB [37,38]. Proper DNase treatment (see Methods) allows the use of intra-exonic primer pairs for both DNA and RNA profiling (F: forward primer; R: reverse primer).

APPLICATION	GENE/CLONE		PRIMER SEQUENCES (5' > 3')
GCN	<i>MYCN</i>		RTPrimerDB ID 11
GXP	<i>MYCN</i>		RTPrimerDB ID 180
GCN	<i>DDX1</i>		RTPrimerDB ID 12
GXP	<i>DDX1</i>		RTPrimerDB ID 93
GCN	<i>NAG</i>		RTPrimerDB ID 13
GXP	<i>NAG</i>		RTPrimerDB ID 114
GCN + GXP	<i>g10e3</i>	F	CTTTTCCTAAGAGCAAGGAAACAGA
		R	CTGTTATTTAAAGAAACCAGCATTCACT
GCN + GXP	<i>g9d9</i>	F	CTTTGGCCAGTTCACAGTTC
		R	CACACCCAGCCTTAAGTTTTTGA
GCN + GXP	<i>NSE1</i>		RTPrimerDB ID 715
GCN + GXP	<i>g10d12</i>	F	ACCTTCCATCCACACCTATGCT
		R	TTTCATTGAGTTCAGTCTTCATCGA
GCN + GXP	<i>DNMT3A</i>		RTPrimerDB ID 713
GCN + GXP	<i>TEM8</i>		RTPrimerDB ID 712
GCN + GXP	<i>BMP10</i>		RTPrimerDB ID 717
GCN + GXP	<i>PLEK</i>		RTPrimerDB ID 716
GCN + GXP	<i>PPP3R1</i>		RTPrimerDB ID 718
GCN + GXP	<i>g4d5</i>	F	AGATGCCCATTTTCATCTCTCTTG
		R	TGCTACAGGTCTTGCATTATCAAAC
GCN + GXP	<i>ETAA16 F</i>		RTPrimerDB ID 714
GCN + GXP	<i>MEIS1</i>		RTPrimerDB ID 690
GCN + GXP	<i>g2h10</i>	F	TTCTGATGCCACAAAGTTTA
		R	ACACAACTCTGAAAGCCAATAATT

Table 2: Haploid DNA copy number in IMR-32 and SK-N-SH compared to a normal human control sample and fold expression difference between IMR-32 and SK-N-SH (tester vs. driver): Real-time quantitative PCR based determination of gene copy number and fold-expression of 5 previously reported amplified genes (*MYCN*, *DDX1*, *NAG*, *NSE1*, *MEIS1*) and 6 other amplified 2p clones (rounded mean of 2 measurements) (*primers designed in 2p24.3; **no expression in SK-S-SH).

GENE	CLONE	ACCESSION NUMBER	LOCATION	HAPLOID COPY NUMBER IN IMR-32	HAPLOID COPY NUMBER IN SK-N-SH	FOLD OVER-EXPRESSION
<i>MYCN</i>		NM_005378	2p24.3	53	1.26	513
<i>DDX1</i>		NM_004939	2p24.3	42	1.28	40
<i>NAG</i>		AF388385	2p24.3	41	1.31	37
<i>NSE1</i>	<i>g10e3</i>	CD664581	2p24.3	70	1.27	280
	<i>g9d9</i>	CD664576		44	1.05	193
		NM_145175		34	0.93	236
-	<i>g10d12</i>	CD664582	2p24.3	27	1.17	1616
<i>TEM8</i>	<i>g1h7</i>	CD664531	2p13.3	26	1.02	8
-	<i>g4d5</i>	CD664538	2p14	30	1.04	1819
<i>MEIS1</i>	<i>g6f6</i>	CD664530	2p14	42	1.27	10
-	<i>g2h10a</i>	CD664534	2p24.3*	42	1.07	**

fied tumour samples). The mRNA expression and gene amplification pattern for *TEM8* resembles that of *MEIS1* ([13] and this study): high expression in a number of cell lines, independent of DNA amplification.

Discussion

In this study, we demonstrate that subtractive cDNA cloning followed by CGH on cDNA microarrays containing the subtracted clones is a powerful strategy for rapid and efficient isolation of amplified genes that are overex-

Table 3: Relative expression levels obtained by real-time quantitative RT-PCR: Quantitative RT-PCR results in 30 NB cell lines and 9 normal human tissue samples (- : not tested; samples with gene amplification are marked in bold-italics).

Sample	MYCN	TEM8	DDX1	NAG	g10e3	g10d12	g4d5	MEIS1
CHP134	1.79E+00	2.40E-01	5.13E+00	2.13E+00	1.21E-01	1.60E-03	4.08E-03	2.39E+00
CHP901	9.72E-01	1.04E-01	1.71E-02	8.94E-02	9.57E-02	6.24E-03	0.00E+00	4.28E-01
CLB-GA	5.35E-02	3.08E-01	1.10E-01	4.19E-01	2.08E-01	2.69E-02	9.98E-03	5.42E-01
GI-ME-N	9.37E-05	4.53E-01	4.95E-02	2.09E-01	2.71E-01	9.83E-03	0.00E+00	1.27E-01
IMR-32	2.00E+00	1.48E+01	2.94E+00	8.78E+00	2.84E+01	1.81E+01	4.06E+01	3.66E+00
LA-N-1	2.31E+00	4.21E-01	4.64E+00	1.39E-01	1.06E-01	2.84E-03	5.01E-02	3.09E-01
LA-N-5	1.16E+00	3.17E-01	2.15E+00	3.95E-01	1.38E-01	2.56E-02	0.00E+00	5.62E-01
LA-N-6	6.93E-03	1.71E-01	1.09E-01	4.72E-01	8.31E-02	2.94E-02	1.99E-02	1.53E-01
N206	2.79E+00	2.22E-01	1.49E-01	4.91E-01	7.35E-01	1.09E-01	1.21E-01	6.48E-01
NBL-S	1.32E-01	3.67E-01	7.35E-02	3.63E-01	4.96E-01	5.28E-02	7.37E-02	6.96E-01
NGP	5.86E+00	1.14E-01	1.25E-01	1.03E+00	2.91E-01	1.10E-02	0.00E+00	1.85E+00
NLF	3.99E-01	1.21E+00	2.91E+00	3.10E-01	2.20E-02	1.38E-02	0.00E+00	2.38E+00
NMB	1.65E+00	8.94E-01	9.47E-02	2.78E-01	5.41E-01	2.15E-02	3.00E-02	1.23E+00
SJNB-1	5.14E-02	4.14E-01	6.76E-02	1.98E-01	3.83E-01	3.26E-02	5.34E-02	3.16E-01
SJNB-10	2.16E+00	2.56E-01	4.89E-02	2.63E-01	1.16E-01	1.92E-02	9.66E-02	1.94E-01
SJNB-12	4.70E-06	3.11E-01	8.89E-02	2.46E-01	3.96E-01	1.60E-02	5.48E-03	1.14E+00
SJNB-6	1.44E+00	2.19E-01	4.61E-02	4.91E-01	4.12E-01	1.44E+00	4.37E-02	7.26E-01
SJNB-8	2.99E+00	6.71E-01	6.26E+00	3.40E-01	3.57E-01	0.00E+00	5.35E-02	3.47E+00
SK-N-AS	2.39E-03	7.89E-01	6.05E-02	1.76E-01	4.54E-03	7.46E-03	2.39E-01	2.63E+00
SK-N-BE	5.31E-01	7.13E-02	8.12E-02	3.30E-01	1.50E-01	3.60E-02	7.36E-02	1.49E-01
SK-N-FI	9.22E-02	8.44E-01	3.69E-02	1.24E-01	7.22E-03	0.00E+00	3.18E-02	2.73E-01
SK-N-SH	3.90E-03	1.88E+00	7.43E-02	2.36E-01	1.01E-01	1.12E-02	2.23E-02	3.78E-01
SMS-KAN	4.59E+00	3.10E-01	5.93E-02	1.55E-01	1.39E-01	6.69E-03	3.03E-02	2.41E+00
SMS-KCNR	2.96E+00	2.61E-01	1.42E-01	2.34E-01	2.42E-01	1.41E-02	0.00E+00	1.03E+00
STA-NB-10	7.28E-01	1.37E+00	3.05E+00	1.90E-01	1.21E-01	7.30E-03	0.00E+00	2.64E-01
STA-NB-12	5.60E-01	-	7.45E-02	3.49E-01	8.12E-02	0.00E+00	6.47E-02	3.14E-01
STA-NB-3	9.82E-01	2.39E-01	8.67E+00	1.59E+01	4.86E-01	0.00E+00	2.22E-01	7.43E-01
STA-NB-8	2.30E-01	5.44E+00	1.04E+00	6.26E-01	6.99E-02	1.38E-02	9.05E-02	5.20E-01
TR-14	1.61E+00	3.59E-01	4.11E-02	1.63E-01	4.47E-02	0.00E+00	1.75E-02	2.00E-01
UHG-NP	2.09E+00	1.91E-01	6.96E-02	6.09E-01	2.04E-01	4.86E-02	8.88E-02	3.58E-01
Human brain	9.74E-03	2.11E-01	-	-	-	0.00E+00	5.68E-02	1.16E-01
Human fetal brain	-	-	-	-	-	6.95E-02	0.00E+00	2.20E-01
Human heart	1.67E-03	1.78E-01	1.20E-01	5.13E-01	5.18E-02	0.00E+00	3.44E-02	3.75E-01
Human kidney	3.53E-02	1.37E-01	9.06E-02	4.21E-01	2.06E-01	3.27E-02	6.82E-02	6.12E-01
Human liver	1.91E-03	1.26E-01	1.20E-01	3.89E-01	1.73E-01	0.00E+00	7.45E-02	5.94E-01
Human lung	6.61E-03	1.94E+00	6.95E-02	3.82E-01	7.77E-01	2.18E-02	1.05E-01	1.10E+00
Human mammary gland	-	-	-	-	-	4.63E-02	4.10E-02	9.39E-01
Human trachea	1.19E-02	3.81E-01	1.90E-02	2.06E-01	1.11E+00	0.00E+00	6.02E-02	1.09E+00
Human uterus	-	-	-	-	-	4.34E-02	2.55E-01	6.70E+00

pressed. As a proof of principle, we analysed neuroblastoma cell line IMR-32 which contains at least two distinct amplification sites on the short arm of chromosome 2 [10,11].

Upon subtractive cDNA cloning and array CGH analysis, fifteen partial cDNA clones located on these sites on 2p were found to be amplified in IMR-32, representing 9 different transcripts. Five of these constitute genes that were previously reported to be amplified in IMR-32 (Table 4), i.e. *MYCN* [18], *DDX1* [19], *NAG* [15] and *NSE1* [17] on chromosome band 2p24, and *MEIS1* [13,14] on 2p14, demonstrating the validity and success of our approach. We not only confirmed *NAG* amplification, but also isolated two partial cDNA clones located within a large intron of the *NAG* gene. Subsequent analyses demonstrated that these clones are part of the *NAG* gene that was

initially misannotated and should in fact contain 21 additional exons, as recently confirmed in another study [16]. We also identified 4 newly amplified transcripts, including the tumour endothelial marker gene *TEM8* (2 partial cDNA clones), encoding a protein highly expressed in tumour endothelial cells but not in normal endothelial cells [20]. Two other transcripts show no homology to any known sequence. The detailed characterization of these anonymous transcripts was beyond the scope of this study.

Our amplicon dissection strategy clearly provides a comprehensive view on the gene content and complex structure of the HSRs (homogeneously staining regions) present in cell line IMR-32. All three HSRs appear to contain the same genes as visualized in a series of FISH mappings, and presumably arise from a single non-synthetic

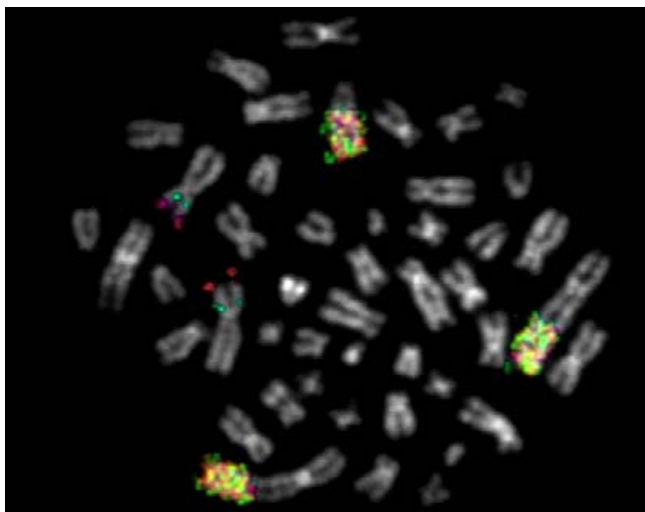


Figure 3
FISH based visualisation of *MYCN* co-amplification with other genes on 2p in neuroblastoma cell line IMR-32: Amplification is present under the form of homogeneously staining regions. *MYCN* (in red) in combination with BAC clone RP11-85D18 (*TEM8*) (in green). Similar results (data not shown) were obtained with clone RP11-444B4 (*MEIS1*), clone RP11-314E10 (*NSE1* and *g10d12*), clone RP11-422A6 (*DDX1*) and clone RP11-516B14 (*NAG*).

amplification event. The amplified genes are located on two different regions, of which the 2p24 region appears to be amplified contiguously, while the other region to be amplified in a discontinuous manner as demonstrated by the presence of a single copy region on 2p13-14, flanked on both sides by amplified sequences. The non-synthetic amplification and inherent fusion of the 2 amplification sites may have caused the formation of a fusion transcript, with activation of cryptic exons, which are not transcribed under normal circumstances. This amplified and highly expressed fusion transcript contains part of the *TEM8* gene on 2p13.3, fused to anonymous spliced sequences located in BAC clone RP11-314E10 on band 2p24.3. The occurrence of a fusion transcript as a result of amplicon formation has been described in a breast cancer cell line MCF7 (caused by non-synthetic co-amplification of two common amplification sites in breast cancer, i.e. 17q23 and 20q13) [21]. However, the significance of these fusion transcripts is at present unclear as no similar fusion transcripts have been detected in other neuroblastoma or breast cancer cells.

Our study clearly demonstrates the power, speed and efficacy of combined subtractive cDNA cloning and DNA copy number determination using array CGH for the identification of clones that are overexpressed and part of

the amplicon, within 4 weeks time. Moreover, the procedure results in the infinite availability of the subtracted cDNA clones, suitable for downstream analyses, such as Northern blot, in situ hybridization or RNA interference using diced double stranded RNA. As a further improvement and simplification of the proposed strategy, we recommend to sequence only the amplified genes detected on the array, instead of sequencing all subtracted clones as performed in this proof-of-principle study. The proposed strategy will not allow isolation of genes which are amplified but not overexpressed. However, one can question the relevance of these genes, as these will most probably not have any biological effect. To our knowledge, such amplification events have not been reported yet.

Oncogene identification consisting of prior selection of differentially expressed genes has already been reported in other cancer cell lines, but -unlike our strategy- was severely hampered by a rate-limiting step for the verification of amplification by radiation hybrid mapping of the subtracted clones [22]. Table 4 summarizes the different strategies used in the past for the identification of amplified genes in neuroblastoma cells. Some of these reports employed a laborious and/or technically challenging method to identify or clone only one single amplified gene. In contrast, a recent study provided a global gene content analysis of the observed amplicons in IMR-32 cells, using CGH on cDNA microarrays [17]. However, this approach was restricted to the identification of genes that were present on the microarray and consequently missed some genes as compared to our strategy (such as the known amplified *DDX1* gene, previously unannotated *NAG* exons, the *TEM8* gene and fusion transcript). Amplification of *DNMT3A* located at 2p23.3, was also reported in above referenced CGH on cDNA microarray study. As the subtractive cDNA cloning procedure did not yield a clone for this gene we performed real-time quantitative PCR analyses which clearly showed that no *DNMT3A* amplification nor overexpression was present in the investigated IMR-32 cells in this study. An explanation for this discrepancy may be cell heterogeneity, as it has been reported that a third amplification site was only present in a minor portion of IMR-32 cells [10].

Investigation of the amplification status of IMR-32 amplified genes in other NB cell lines revealed that three of the nine genes were also amplified in other samples, albeit always co-amplified with *MYCN*. However, it remains an unsolved question whether co-amplified genes represent silent passengers, or co-determinants of phenotype [23]. The frequently co-amplified gene *DDX1* is a nice example, as no correlation between amplification and patient outcome could be established [24], but nevertheless, the gene appears to have oncogenic properties [23]. Six of the nine identified genes, were only amplified in cell line

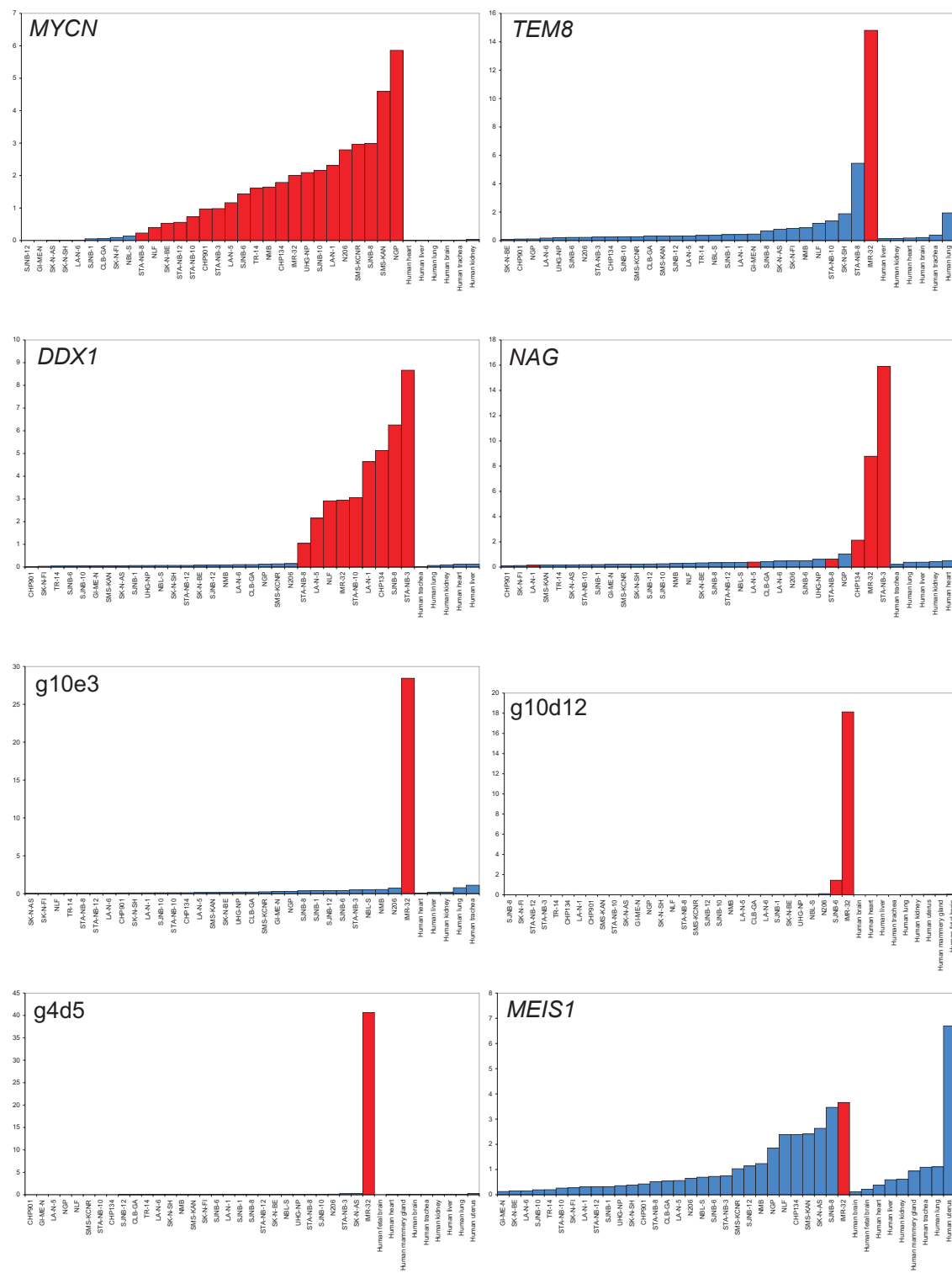


Figure 4
Relative expression levels obtained by real-time quantitative RT-PCR: Relative mRNA expression levels obtained by quantitative PCR in 30 neuroblastoma cell lines and 9 normal human tissue samples (samples with gene amplification are marked in red) (relative scale, rescaled to an average expression level of 1).

IMR-32. However, the amplification of a gene in only a single sample does not preclude in advance its possible role in tumour biology. An interesting example is the *MEIS1* oncogene, with proven oncogenic properties (reviewed in [25]). Albeit amplified in only one neuroblastoma sample, the gene is overexpressed in about one quarter of other tested neuroblastoma tumour samples ([13] and this study). A similar situation occurs for *TEM8*, a tumour-specific endothelial marker that has been implicated in colorectal cancer [26]. Besides amplification and overexpression in IMR-32, high *TEM8* expression independent of gene amplification is observed, suggesting alternative pathways for gene activation and a possible role in neuroblastoma pathogenesis. Further evidence that one or more genes in the 2p13-14 amplicon plays a role in neuroblastoma comes from the observation of genomic amplification at chromosome bands 2p13-14 in 3 primary tumour samples, from a large European multi-centre CGH study of 204 cases [27]. Unfortunately, no material was available for further investigation of these samples. Clearly, more detailed analyses of the amplified genes (amongst others in a large cohort of uniformly treated primary tumour samples) and functional studies are required to establish a possible role of one of the new genes in tumourigenesis.

Conclusions

The present study shows that the combinatorial method of subtractive cDNA cloning followed by array CGH allows straightforward and efficient isolation of overexpressed genes located in amplification sites. The validity of our approach is clearly illustrated by the detection of all genes that were previously found to be amplified in neuroblastoma cell line IMR-32; the identification of 3 newly amplified genes and a fusion transcript and the generation of new data on gene content and structure of the amplicon.

Methods

DNA and RNA isolation

DNA from cultured neuroblastoma cells and 75 *MYCN* amplified DNA tumours was extracted using the Easy DNA kit following the instructions of the manufacturer (Invitrogen). Total RNA of cultured cell lines was isolated using the RNeasy Midi kit (Qiagen), and mRNA was extracted from SK-N-SH and IMR-32 with the FastTrack kit (Invitrogen), both according to the manufacturer's instructions.

RNA and DNA concentration was determined using the Picogreen and Ribogreen reagent, respectively (Molecular Probes) on a TD-360 fluorometer (Turner Designs).

Suppression subtractive hybridization (SSH)

Starting from 2 µg of mRNA from cell lines SK-N-SH (driver) and IMR-32 (tester), SSH was performed with the PCR-Select cDNA Subtraction kit (BD Biosciences, Clontech) as described by the manufacturer. The PCR product mixture of putative differentially expressed genes was subcloned into the pGEM-T Easy vector (Promega) and propagated in DH5α *E. coli*. 960 clones were picked, grown in 96-well plates and stored as glycerol stocks at -80°C for further analysis. Differential screening was performed to eliminate possible false positive clones according to the guidelines described in the Differential Screening kit (BD Biosciences, Clontech).

DNA sequencing and analysis

SSH clones were PCR amplified using SP6 and T7 vector specific sequences flanking the cloning site. PCR products were exonuclease and phosphatase treated and cycle sequenced using BigDyeTerminator chemistry on an ABI377 (Applied Biosystems) with primers that annealed to the SP6 or T7 sequences. Similarity searches were performed using the BLAST algorithm [28] after removing vector and masking repeat sequences using RepeatMasker [29]. Sequence alignment and EST contig building were performed using the freely available BioEdit package [30].

Microarray slide production

From selected SSH clones, plasmid DNA was prepared using the Montage Plasmid Miniprep₉₆ Kit (Millipore) according to the manufacturer. The plasmid insert was amplified on a PTC-200 DNA engine (MJ Research) in a total volume of 100 µl containing 1 µl of 1/100 diluted plasmid DNA (1–2 ng), 1 × PCR Gold buffer (Applied Biosystems), 5 mM MgCl₂, 400 µmol of each dNTP, 5 U AmpliTaq Gold DNA polymerase (Applied Biosystems) and 1 µmol of each primer (amino-linked SP6 and T7 primers). The cycling conditions comprised 5 min polymerase activation at 95°C, 40 cycles with denaturation at 94°C for 15 sec, annealing at 55°C for 15 sec and extension at 72°C for 2 min, and a final extension for 5 min at 72°C. PCR products were run on a 1.5% TBE-agarose gel. After vacuum centrifugation, dried PCR products were dissolved in 20 µl spotting buffer (200 mM sodium phosphate buffer pH 8.5, 0.2% sarcosyl) and rearrayed in a 384 well plate. The PCR products were then arrayed in triplicate on CodeLink Activated Slides (Amersham Biosciences) using a GMS417 spotter (MWG Biotech). After 48 hour incubation in a NaCl humidified chamber, slides were transferred to 1% ammonium hydroxide solution for 5 min, rinsed in Milli-Q ddH₂O (Millipore) at room temperature and then placed in 95°C Milli-Q ddH₂O for 2 min to completely denature the bound DNA molecules. After transfer to ice-cold Milli-Q ddH₂O, slides were briefly rinsed twice in room temperature Milli-Q ddH₂O.

and dried by spinning in a centrifuge for 5 min at 1000 rpm.

CGH on cDNA microarray

This protocol is based on a previously published CGH on cDNA microarray protocol [31] and a CGH on BAC microarray protocol [32].

Approximately 10 µg of genomic DNA of neuroblastoma cell line IMR-32 and a normal male lymphoblastoid cell line was digested overnight at 37°C with 25 units AluI and RsaI in 100 µl React1 buffer (Invitrogen). Digested DNA was purified using QIAquick PCR purification columns (Qiagen) according to the manufacturer.

Purified DNA was labelled using the BioPrime random-priming labelling kit (Invitrogen) substituting the biotin labelled nucleotide with Cy3-dCTP and Cy5-dCTP for tumour and normal DNA, respectively. A total of 4 labelling reactions were set up (2 reactions for each Cy dye), each containing 2 µg of digested DNA. Twenty µl 2.5 × random primer buffer mix was added to 2 µg DNA (diluted in 21 µl) and then boiled for 10 min. On ice, 5 µl 10 × dNTP mix (2 mM each dATP, dGTP, and dTTP and 0.5 mM dCTP in TE), 3 µl Cy5 or Cy3-dCTP (Amersham Biosciences, 1 mM) and 1 µl Klenow Fragment were added to each tube. This mixture was incubated at 37°C for 2 hours and stopped by adding 5 µl stopping buffer. The DNA probes were purified on a Microspin G50 column (Amersham Biosciences) as described by the manufacturer. Cy3 and Cy5 labelled probes were subsequently mixed and combined with 50 µg human Cot-1 DNA (Invitrogen), 100 µg yeast tRNA and 20 µg poly dA (Sigma). After ethanol-sodium acetate precipitation, the probe was dissolved in 70 µl hybridisation buffer (50% formamide, 10% dextrane sulphate, 0.1% Tween20, 2 × SSC, 10 mM Tris/HCl pH 7.4). The hybridisation mixture was then denatured at 100°C for 2 min and incubated for 30 min at 37°C in a PTC-200 thermocycler (MJ Research). The probe was applied to a microarray slide that had been pre-hybridised for 2 hours with hybridisation buffer. An open hybridisation (without cover slip) was performed for 2 nights at 37°C in a sealed, humidified chamber on a rocking table. Washes were performed in three steps: PBS/0.05% Tween20 for 10 min at room temperature, 50% formamide/2 × SSC for 30 min at 42°C and PBS/0.05% Tween20 for 10 min at room temperature. Slides were dried by spinning for 5 min at 500 rpm.

The slides were scanned in a GMS418 scanner (MWG Biotech) and images were analyzed using ImaGene v5.5 software (BioDiscovery). After background subtraction, spots (background signal < signal, for the 2 colours) were normalized with the geometric mean of selected data points (signal > background signal + 3 × standard deviation of all

background signals, for the 2 colours). Ratios were calculated using these normalized data and put in a graph against the base position of the clone according to the human genome browser at UCSC (April 2003 freeze [33]).

Real-time quantitative PCR based copy number determination and gene expression analysis

The gene copy number of known genes *MYCN*, *DDX1*, *NAG*, *MEIS1*, *TEM8*, *BPM10*, *PLEK*, *PPP3R1* and *DNMT3A* and anonymous SSH clones g10e3, g9d9, g10d12, g4d5 and g2h10a was determined in 32 other neuroblastoma cell lines with listed primers (Table 1) according to a previously described protocol with *BCMA* and *SDC4* as normalizing control genes and normal human genomic DNA (Roche) as calibrator sample [24]. Clones that were found to be amplified in cell lines other than IMR-32 were also tested in 75 *MYCN* amplified tumours. PCR reactions were performed on an ABI 5700 SDS (Applied Biosystems). Amplification mixtures (25 µl) contained template DNA (approximately 10 ng), 1 × qPCR MasterMix for SYBR Green I (Eurogentec) and 300 nM of each primer. The cycling conditions comprised 10 min polymerase activation at 95°C, 40 cycles at 95°C for 15 sec and 60°C for 1 min. A dissociation curve was run after each PCR reaction in order to verify amplification specificity.

The relative expression levels of the clones were determined in the neuroblastoma cell line panel and on 9 normal tissue samples (RNA obtained from BD Biosciences, Clontech) using the above listed primer pairs according to an optimized two-step real-time SYBR Green I RT-PCR assay [34]. The gene expression levels were normalized using the geometric mean of 4 stable housekeeping genes in neuroblastoma (*SDHA*, *UBC*, *GAPD* and *HPRT1*) as described previously [35].

Validation of amplification with FISH

FISH was performed using the LSI *MYCN* SpectrumOrange probe (Vysis) in combination with BAC clone RP11-422A6 containing *DDX1*, BAC clone RP11-516B14 containing *NAG*, BAC clone RP11-444B4 containing *MEIS1*, BAC clone RP11-314E10 containing *NSE1* and SSH clone g10d12 and BAC clone RP11-85D18 containing *TEM8*. Labelling and FISH was performed as described [36].

Further characterization of anonymous SSH clones

RT-PCR assays on IMR-32 cDNA were designed to test whether an anonymous SSH clone and a neighbouring (putatively not yet fully annotated) transcript are part of the same gene. Taking into account the orientation of the sequences, a forward primer was designed in the SSH clone and a reverse primer in the known transcript: forward primer in clone g10e3

Table 4: Different approaches used for identification of amplified genes in (IMR-32) neuroblastoma cell lines

Methodology	Publication	MYCN	DDX1	NAG	NSE1	g10d12	DNMT3A	TEM8	g4d5	MEIS1	g2h10
cloning after MYC Southern blot hybridization	Schwab et al. (1983) [18]	X									
subtractive cDNA cloning and Southern blot	Manohar et al. (1995) [39]		X								
differential cDNA library screening	Godbout and Squire (1993); Squire et al. (1995) [19,40]		X								
2D separation of genomic restriction fragments	Wimmer et al. (1999) [15]			X							
screening PAC/BAC libraries after microdissection	Jones et al. (2000) [14]									X	
serendipitous cloning after multiprobe Southern blot	Spieker et al. (2001) [13]									X	
CGH on cDNA microarrays	Beheshti et al. (2003) [17]	X		X	X		X			X	
combined subtractive cDNA cloning and array CGH	this study	X	X	X	X	X		X	X	X	X

5'AGTCACTGAGACAGAAAAGAGGTGGAATGC3' and reverse primer in gene *NSE1* 5'GGAGGAAGATGGCGCTGCGAATTC3', forward primer in clone g9d9 5'CCACAGAAGGTGTTTCACACCCAGCCT3' and reverse primer in *NSE1* 5'GGAGGAAGATGGCGCTGCGAATTC3'; forward primer in clone g10d12 5'GACAGGCTT-GCCAAATTTTCACAGTGTGG 3' and reverse primer in gene *NSE1* 5'CCGACCCGAGTTCGTCCTTTT3'; forward primer in clone g4d5 5'AGCTAGGCTCGCAAACAACGTTTCCAGA3' and reverse primer in gene *ETAA16* 5'GCCAAGAAGTCCAGAGGCTTTTGGGA3'. To determine the *NAG* transcript length between exon 4 and 7 (acc. no. AF056195), RT-PCR with a forward primer in exon 4 and a reverse primer in exon 7 was performed (F 5'GCTCCCTGATGGACTGGTTCGCTTGGT3' and R 5'CCGGCCAGTGTGCCTCGTCAATCTA3'). Examination of the fusion transcript was done with a forward primer in the first part of the transcript (F 5'CACTGTTCTGACGGTTCCA3') and a reverse primer in the other part (R 5'CAAAGTAGAATATAGTTGTCCAAAACACAA3'). RT-PCR amplification on random hexamer primed IMR-32 cDNA was performed with the Advantage 2 PCR Kit (Clontech, BD Biosciences) according to the manufacturer.

PCR fragments run on a 1.5% TBE-agarose gel were excised and purified on a GenElute Minus EtBr Spin Column (Sigma-Aldrich). Cycle sequencing was performed using purified amplicons (3–10 ng) using the above-mentioned primers at a concentration of 80 nM and the ABI PRISM BigDye Terminators v3.0 Cycle Sequencing Kit (Applied Biosystems) according to the manufacturer, with the following thermocycling conditions: 25 cycles at 92° for 10 sec, 55° C for 5 sec and 60° C

for 3.5 min. Sequencing of the fusion transcript was preceded by cloning of the PCR product with the TOPO TA cloning kit for sequencing (Invitrogen). After ethanol precipitation, the products were run on an automated sequencer ABI3100 and analyzed with the Sequencing Analysis software v3.7 (Applied Biosystems).

List of Abbreviations

- SSH = suppression subtractive hybridisation
- CGH = comparative genomic hybridisation
- FISH = fluorescence in situ hybridisation
- SAGE = serial analysis of gene expression
- HSR = homogeneously staining region
- dmin = double minute chromatin bodies
- RT-PCR = reverse transcriptase polymerase chain reaction

Authors' contributions

JV oversaw the project and performed SSH, differential screening and sequencing, in collaboration with GB and KS in the lab of FVR. KDP and FP were involved in the microarray production, array CGH analysis and further characterization of SSH clones by quantitative RT-PCR. BM helped with fine-tuning of the array CGH protocol. KDP and JV performed further analysis on the amplified SSH clones and drafted the manuscript; all other authors have reviewed the manuscript and FS and ADP were the final editors of the manuscript.

Acknowledgements

We would like to thank Els De Smet for the FISH mapping, Geert De Vos and Peter De Graeve for culturing the cell lines, Katrien Staes for help with the subtractive cDNA cloning, and Lieven Thorrez for the sequencing.

This work was supported by BOF-grant 01IF1200 and 01IB4300, GOA-grant I205I203, FWO-grant G.0028.00 and VEO-grant 01IV1302. Kathleen De Preter and Filip Pattyn are aspirants with the Fund for Scientific Research Flanders (FWO-Vlaanderen). Jo Vandesompele is supported by a post-doctoral grant from the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT).

References

- Savelyeva L, Schwab M: **Amplification of oncogenes revisited: from expression profiling to clinical application.** *Cancer Lett* 2001, **167**:115-123.
- Schwab M: **Oncogene amplification in solid tumors.** *Semin Cancer Biol* 1999, **9**:319-325.
- Schwab M: **Amplification of oncogenes in human cancer cells.** *Bioessays* 1998, **20**:473-479.
- Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D: **Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors.** *Science* 1992, **258**:818-821.
- Knuutila S, Aalto Y, Autio K, Bjorkqvist AM, El-Rifai W, Hemmer S, Huhta T, Kettunen E, Kiuru-Kuhlefelt S, Larramendy ML, Lushnikov T, Monni O, Pere H, Tapper J, Tarkkanen M, Varis A, Wasenius VM, Wolf M, Zhu Y: **DNA copy number losses in human neoplasms.** *Am J Pathol* 1999, **155**:683-694.
- Pinkel D, Segreaves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG: **High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays.** *Nat Genet* 1998, **20**:207-211.
- Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P: **Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances.** *Genes Chromosomes Cancer* 1997, **20**:399-407.
- Wang TL, Maierhofer C, Speicher MR, Lengauer C, Vogelstein B, Kinzler KW, Velculescu VE: **Digital karyotyping.** *Proc Natl Acad Sci U S A* 2002, **99**:16156-16161.
- Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO: **Genome-wide analysis of DNA copy-number changes using cDNA microarrays.** *Nat Genet* 1999, **23**:41-46.
- Van Roy N, Jauch A, Van Gele M, Laureys G, Versteeg R, De Paepe A, Cremer T, Speleman F: **Comparative genomic hybridization analysis of human neuroblastomas: detection of distal 1p deletions and further molecular genetic characterization of neuroblastoma cell lines.** *Cancer Genet Cytogenet* 1997, **97**:135-142.
- Shiloh Y, Shipley J, Brodeur GM, Bruns G, Korf B, Donlon T, Schreck RR, Seeger R, Sakai K, Latt SA: **Differential amplification, assembly, and relocation of multiple DNA sequences in human neuroblastomas and neuroblastoma cell lines.** *Proc Natl Acad Sci U S A* 1985, **82**:3761-3765.
- Van Roy N, Van Limbergen H, Vandesompele J, Van Gele M, Poppe B, Salwen H, Laureys G, Manoel N, De Paepe A, Speleman F: **Combined M-FISH and CGH analysis allows comprehensive description of genetic alterations in neuroblastoma cell lines.** *Genes Chromosomes Cancer* 2001, **32**:126-135.
- Spieker N, van Sluis P, Beitsma M, Boon K, van Schaik BD, van Kampen AH, Caron H, Versteeg R: **The MEIS1 oncogene is highly expressed in neuroblastoma and amplified in cell line IMR32.** *Genomics* 2001, **71**:214-221.
- Jones TA, Flomen RH, Senger G, Nizetic D, Sheer D: **The homeobox gene MEIS1 is amplified in IMR-32 and highly expressed in other neuroblastoma cell lines.** *Eur J Cancer* 2000, **36**:2368-2374.
- Wimmer K, Zhu XX, Lamb BJ, Quick R, Ambros PF, Kovar H, Thoraval D, Motyka S, Alberts JR, Hanash SM: **Co-amplification of a novel gene, NAG, with the N-myc gene in neuroblastoma.** *Oncogene* 1999, **18**:233-238.
- Scott DK, Board JR, Lu X, Pearson AD, Kenyon RM, Lunec J: **The neuroblastoma amplified gene, NAG: genomic structure and characterisation of the 7.3 kb transcript predominantly expressed in neuroblastoma.** *Gene* 2003, **307**:1-11.
- Beheshti B, Braude I, Marrano P, Thorner P, Zielenska M, Squire JA: **Chromosomal localization of DNA amplifications in neuroblastoma tumors using cDNA microarray comparative genomic hybridization.** *Neoplasia* 2003, **5**:53-62.
- Schwab M, Alitalo K, Klempnauer KH, Varmus HE, Bishop JM, Gilbert F, Brodeur G, Goldstein M, Trent J: **Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour.** *Nature* 1983, **305**:245-248.
- Squire JA, Thorner PS, Weitzman S, Maggi JD, Dirks P, Doyle J, Hale M, Godbout R: **Co-amplification of MYCN and a DEAD box gene (DDX1) in primary neuroblastoma.** *Oncogene* 1995, **10**:1417-1422.
- Carson-Walter EB, Watkins DN, Nanda A, Vogelstein B, Kinzler KW, St Croix B: **Cell surface tumor endothelial markers are conserved in mice and humans.** *Cancer Res* 2001, **61**:6649-6655.
- Barlund M, Monni O, Weaver JD, Kauraniemi P, Sauter G, Heiskanen M, Kallioniemi OP, Kallioniemi A: **Cloning of BCAS3 (17q23) and BCAS4 (20q13) genes that undergo amplification, overexpression, and fusion in breast cancer.** *Genes Chromosomes Cancer* 2002, **35**:311-317.
- Weggen S, Preuss U, Pietsch T, Hilger N, Klawitz I, Scheidtman KH, Wiestler OD, Bayer TA: **Identification of amplified genes from SV40 large T antigen-induced rat PNET cell lines by subtractive cDNA analysis and radiation hybrid mapping.** *Oncogene* 2001, **20**:2023-2031.
- Scott D, Elsdon J, Pearson A, Lunec J: **Genes co-amplified with MYCN in neuroblastoma: silent passengers or co-determinants of phenotype?** *Cancer Lett* 2003, **197**:81-86.
- De Preter K, Speleman F, Combaret V, Lunec J, Laureys G, Eussen BH, Francotte N, Board J, Pearson AD, De Paepe A, Van Roy N, Vandesompele J: **Quantification of MYCN, DDX1, and NAG gene copy number in neuroblastoma using a real-time quantitative PCR assay.** *Mod Pathol* 2002, **15**:159-166.
- Geerts D, Schilderink N, Jorritsma G, Versteeg R: **The role of the MEIS homeobox genes in neuroblastoma.** *Cancer Lett* 2003, **197**:87-92.
- St Croix B, Rago C, Velculescu V, Traverso G, Romans KE, Montgomery E, Lal A, Riggins GJ, Lengauer C, Vogelstein B, Kinzler KW: **Genes expressed in human tumor endothelium.** *Science* 2000, **289**:1197-1202.
- Vandesompele J, Speleman F, Van Roy N, Laureys G, Brinskchmidt C, Christiansen H, Lampert F, Lastowska M, Bown N, Pearson A, Nicholson JC, Ross F, Combaret V, Delattre O, Feuerstein BG, Plantaz D: **Multicentre analysis of patterns of DNA gains and losses in 204 neuroblastoma tumors: how many genetic subgroups are there?** *Med Pediatr Oncol* 2001, **36**:5-10.
- NCBI BLAST** [<http://www.ncbi.nlm.nih.gov/BLAST/>]
- RepeatMasker** [<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker/>]
- BioEdit** [<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>]
- Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousset S, Rozenblum E, Ringner M, Sauter G, Monni O, Elkahoul A, Kallioniemi OP, Kallioniemi A: **Impact of DNA amplification on gene expression patterns in breast cancer.** *Cancer Res* 2002, **62**:6240-6245.
- Fiegler H, Carr P, Douglas EJ, Burford DC, Hunt S, Scott CE, Smith J, Vetrie D, Gorman P, Tomlinson IP, Carter NP: **DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones.** *Genes Chromosomes Cancer* 2003, **36**:361-374.
- UCSC human genome browser** [<http://genome.ucsc.edu>]
- Vandesompele J, De Paepe A, Speleman F: **Elimination of primer-dimer artifacts and genomic coamplification using a two-step SYBR green I real-time RT-PCR.** *Anal Biochem* 2002, **303**:95-98.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.** *Genome Biol* 2002, **3**:RESEARCH0034.
- Van Roy N, Laureys G, Cheng NC, Willem P, Opendakker G, Versteeg R, Speleman F: **1;17 translocations and other chromosome 17 rearrangements in human primary neuroblastoma tumors and cell lines.** *Genes Chromosomes Cancer* 1994, **10**:103-114.
- RTPrimerDB: the real-time PCR primer and probe database** [<http://medgen.ugent.be/rtprikerdb/>]

38. Pattyn F, Speleman F, De Paepe A, Vandesompele J: **RTPrimerDB: the real-time PCR primer and probe database.** *Nucleic Acids Res* 2003, **31**:122-123 [<http://medgen.ugent.be/rtprimerdb>].
39. Manohar CF, Salwen HR, Brodeur GM, Cohn SL: **Co-amplification and concomitant high levels of expression of a DEAD box gene with MYCN in human neuroblastoma.** *Genes Chromosomes Cancer* 1995, **14**:196-203.
40. Godbout R, Squire J: **Amplification of a DEAD box protein gene in retinoblastoma cell lines.** *Proc Natl Acad Sci U S A* 1993, **90**:7578-7582.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

