Cloning and physical mapping of DNA sequences encompassing a region in N-myc amplicons of a human neuroblastoma cell line

Kiyotaka Akiyama and Yoshisuke Nishi*

Life Science Research Laboratory, Japan Tobacco, Inc., 6-2 Umegaoka, Midori-ku, Yokohama, Kanagawa 227, Japan

Received August 5, 1991; Revised and Accepted November 20, 1991

ABSTRACT

Cloning and physical mapping of DNA sequences encompassing N-myc amplicons of a human neuroblastoma cell line were done. A number of lambda phage clones within this region were isolated using the probes prepared by the phenol emulsion reassociation technique. Based on the restriction mapping, they were integrated into 8 contigs with sizes of 25 – 60 kb which, in total, encompassed a 330 kb region. Several amplicons, 100, 420, 480 and 520 kb in size as a NotI fragment, were identified using hexagonal field gel electrophoresis, and the contigs were assigned in these NotI fragments. The region encompassed by the contigs was equivalent to some 60–80% of the amplicons identified as a NotI fragment. In order to compare the amplified regions flanking the N-myc gene among the cell lines, the phage clones to cover the whole contigs were used for hybridization as a probe. The results showed that the portions of the whole contigs ranging 18 – 45% were also amplified in the cell lines examined. These results allowed us to identify the 'rearranged sites' which were rather evenly distributed, one at every 40 kb, through the contigs. These observations lead to the idea that an amplified DNA domain is constructed after the multiple rearrangements and then increases in number, finally resulting in the formation of subsets of amplicons with sequence homogeneity.

INTRODUCTION

Unscheduled DNA amplification has been found in mammalian cells. The genes exhibiting the amplification are of two types: one is drug-resistant genes and the other is cellular proto-oncogenes (1, 2, 3, 4, 5). More than 10 proto-oncogenes have been found to exhibit amplification. Among them, the N-myc gene amplified in human neuroblastomas has been studied the most extensively. It has been shown that N-myc amplification is of clinical significance in neuroblastomas (6), since the multiplication of the N-myc gene frequently correlates with the advanced stages of neuroblastomas (7, 8, 9, 10, 11).

In order to better understand the mechanism of N-myc amplification, several experimental approaches have been tried. They are classified into two categories. One is to analyze the original N-myc locus. Such studies help us to elucidate the event at the very beginning of N-myc amplification. In this regard, Hunt *et al.* (12) recently reported that excision of the N-myc locus from the original position on chromosome 2p24 occurred in some neuroblastoma cell lines. The other approach is to analyze the structural organization of N-myc amplification in the amplified regions. Such studies are valuable for elucidation of the subsequent processes in the amplification. In most cases, N-myc amplification has been found in abnormal chromosomes manifested as double minutes (DMs) and/or homogeneously staining regions (HSRs). These chromosome structures are composed of repeated units of amplified sequences termed 'N-myc amplicons', each of which contains not only the N-myc gene but also several-hundred kilobases of flanking sequences. Different approaches have been used to analyze this region. These include cloning of the amplified sequences by sorting the chromosome containing HSR (13), and other methods, such as the in-gel denaturation-renaturation technique (14, 15), the phenol emulsion reassociation technique (PERT) (16, 17), and the cloning of lambda phages and cosmids using conventional chromosome walking (15). Large-scale physical mapping by pulsed field gel electrophoresis has also been used (18, 19). These studies indicate that there are two types of amplification profiles. One contains many rearrangements (16, 18, 19, 20) and the other relatively few (15). Within individual tumors and/or cell lines, there are several N-myc amplicons which are oriented in a tandem head-to-tail array (18). Despite the evidence described above, there is still not enough information to draw a complete picture of the N-myc amplicons and the process of amplification. This is chiefly due to the fact that, in comparison with the size of the N-myc amplicons, the isolated sequences are too short to be clearly defined, even if they are ordered (14, 15, 16, 17).

Cloning full sets of ordered clones of N-myc amplicons will be of great value in elucidating the structural organization and assessing the amplification mechanism. The 'chromosome walking' achieved by isolating overlapping clones from the N-myc gene is not readily applicable to cloning the whole series

* To whom correspondence should be addressed

of sequences contained in the N-myc amplicons. The difficulties are due to the fact that each amplicon is so large that the analysis of cloned sequences requires the time-consuming repetition of stepwise walking procedures. In addition, the fact that there are integrated structures of amplicons that are not simply tandem repeats, but include complex rearrangements with several subclasses of amplicons with closely related structures, makes it much harder to clone sequences using a one-step walking method. As shown in Fig. 1, our strategy to overcome these difficulties is as follows: 1) using phenol emulsion reassociation technique (PERT) to clone a larger number of the short fragments of amplified sequences randomly from an enrichment library, 2) isolating lambda phage genomic clones hybridized with PERT clones as probes, 3) constructing restriction maps of these lambda clones through partial digestion by eight six-base-cutting enzymes and ordering these clones by comparing their patterns, and 4) combining these lambda clones into contigs and further combining them using pulsed field gel electrophoresis to make a complete physical map. In the previous work (16), we succeeded in isolating a large number of PERT clones from a N-myc amplicons of a neuroblastoma cell line showing extensive rearrangements in the amplicons that permitted us to characterize the N-myc amplification of that cell line. Our strategy to clone lambda phages and construct their restriction maps was powerful enough to permit cloning of the whole *E. coli* genome (21). In the present work, we have combined these methods and further attempted to clone ordered sets of the DNA sequences derived from the N-myc amplicons. Finally, we were able to clone the amplified regions encompassing 330 kb, which were subdivided into eight contigs. These regions covered some 70 to 80% of the N-myc amplicons identified in this cell line. With these clones, the amplification organization of N-myc amplicons was constructed in this cell line and further characterized in other neuroblastoma cell lines. Based on the evidence obtained, the process of the amplification of the N-myc gene is discussed.

MATERIALS AND METHODS

Cell sources

The human neuroblastoma cell lines utilized in this study were MC-NB-1, NB 9, NB 16, NB 19, NB 69, LA-N-1, LA-N-2, LA-N-5, CHP-126, CHP-134, IMR-32, and TNB-1. As control cell lines without N-myc amplification, a neuroblastoma cell line NB 69 and a human lymphoblastoid cell line GM3798 were also used. The sources of these cell lines and the culture conditions were previously described (16). A subline containing a large HSR on chromosome 2q but no DMs was isolated from MC-NB-1 (16).

Preparation of an EMBL3 genomic library

A phage library was constructed from this subline. Genomic DNA isolated with the method described by Aldridge *et al.* (22) were partially digested with Sau3AI (Nippon Gene, Toyama, Japan), dephosphorylated by bacterial alkaline phosphatase (Takara Shuzo, Kyoto, Japan), and ligated with lambda EMBL3 phage DNA cleaved with BamHI (Stratagene, San Diego, Calif. U.S.A.). The ligation mixture was packaged into GIGAPACK XL (Stratagene) and propagated on *E. coli* P2392 (Stratagene) following the manufacture's protocols.

Construction of restriction maps of cloned phage DNA

In order to isolate EMBL3 phage clones from the amplified sequences of the neuroblastoma cell line MC-NB-1, PERT clones which had been isolated from a PERT library of a subline of

MC-NB-1 and exhibited co-amplification with N-myc (16) were used as a probe. The isolated phage clones were transferred to 100 μ l of the phage diluent (0.1 M NaCl, 8 mM MgSO₄, 50 mM Tris-HCl at pH 7.5 and 0.01% SDS). These clones were then propagated, and DNA were extracted and digested following the 'mini-plate lysate method' as previously described (21). In this method, all steps, including preparation of phage lysates, extraction with phenol, precipitation with ethanol and digestion with restriction enzymes, were carried out using 96-well microtiterplates. For construction of the restriction maps of each phage clone, DNA were partially digested with BamHI, HindIII, EcoRI, Sall, BglII, PstI, PvuII or KpnI (Nippon Gene). Digestion was carried out for 1 hr at 37°C at 0.16 U for BglII, 0.3 U for PstI, 0.4 U for BamHI and PvuII and 3 U for HindIII, EcoRI, Sall or KpnI. The units were previously determined by sequential dilution using EMBL3 phages without inserts. The digests were then electrophoresed in 0.4% agarose gel, transferred to nylon filters (Hybond N, Amersham, Arlington Height, IL, U.S.A.) and hybridized with a 4.4 kb HindIII fragment of lambda phage DNA, which is situated at the very end of the right arm. The sizes of the hybridized bands of the phage clones were determined on a disitizer (Hitachi LTD, Tokyo, Japan) with the aid of the computer program 'DNASIS' (Hitachi Software Engineering Co., Ltd. Kanagawa, Japan), based on the sizes of the co-migrated DNA fragments of lambda partially digested with Aval.

Large-scale mapping using hexagonal field gel electrophoresis (HFGE)

HFGE, a modified method of pulsed field gel electrophoresis (23), was performed following the protocol described by Smith and Cantor (24) using an LKB Pulsaphor (Pharmacia LKB Biotechnology, Uppsala, Sweden). In brief, gel blocks containing 4×10^6 cells in 100 μ l of 0.5% low melting agarose (FMC BioProducts, Rockland, ME, U.S.A.) were lysed by lauroyl sarcosine (1%, Sigma, St Louis, MO, U.S.A.), proteinase K (1.0 mg/ml, Sigma) and EDTA at 50°C overnight. Each block was digested with 40 units of NotI, SfiI (Toyobo, Osaka, Japan) or SmaI (Nippon Gene), or digested with a combination of NotI and SfiI. Digestion was carried out for 12 hr at 37°C for NotI and SmaI, at 50°C for SfiI, and both for double digestion. Pieces of agarose blocks were loaded and electrophoresed in 1% agarose gel in a loading buffer (0.45 M boric acid, 0.45 M Tris-base at pH 8.0, and 12.5 mM EDTA) at 170 V for 40 hr at 4°C at a pulse time of 15, 20 or 45 sec. DNA were then transferred to nylon filters and hybridized with the PERT clones as probes.

Experiments of amplification profiles on different neuroblastoma cell lines

DNA of the cell lines isolated following the protocol described above (22) were digested completely with HindIII (New England Biolabs., Beverly, MA, U.S.A.). The concentrations of DNA were determined fluorometrically using the fluorodye 4', 6-diamidino-2-phenylindole dihydrochloride hydrate (Sigma). Exactly 2 μ g of each digest were electrophoresed in 0.8% agarose gel, transferred to a nylon filter and hybridized with the phage clones as probes. Amplified regions were determined by comparing the intensity of the hybridization signals of the neuroblastoma cell lines with those of GM3798. In order to cover all the contigs, the minimally overlapping phage clones were used as probes. These included 1A7, 1G4, 1F2, 1A1, 1C11, 1C10, 1G6, 2B10, 2A1, 1E1, 1B1, 2B9, 1F4, 1A5, 1G2, 1G6, 2B7, 1F9, 1H6, 2C1, 1F1, 2B2, 1B10 and 1F5.

Figure 1. The procedure for the construction of the ordered clones of the N-myc amplicon in the cell line MC-NB-1 is illustrated. The process consists of four steps in which the amplified short fragments were isolated by PERT, the lambda phage recombinant clones were screened by the PERT clones, these clones were then ordered by constructing the restriction maps, and these maps were examined by PFGE.

Probes, radiolabeling and hybridization

The N-myc probe used was: NB-19-21A, a 1.0 kb BamHI-EcoRI fragment of the 3' end of NB-19-21, which includes a portion of the 3' end of exon II (8). As there is a NotI site within exon II of the N-myc gene, the NotI-linking clones isolated were: N-myc A, a 0.9 kb NotI-EcoRI fragment of the 3' downstream of the NotI site, and N-myc B, a 1.1 kb EcoRI-NotI fragment of the 5' upstream of the NotI site of the EcoRI-EcoRI fragment that includes exon II of the N-myc gene used in HFGE. All PERT clones cloned in the plasmid pBS+ were cut with PvuII, and the inserts were used for the hybridization of DNA after HFGE. 1F5-1, a HindIII-SalI fragment located on the extreme end of Contig H, was also used as a probe. EMBL3 phage clones were used without any digestion. In all experiments, radiolabeled probes were prepared by the random priming method (25, 26). Hybridization was performed in $6\times$ SSC, $1\times$ Denhardt's solution, 0.1 M sodium phosphate buffer at pH 7.0, 0.1% SDS and 100 mg/ml of sonicated and denatured salmon sperm DNA (Sigma) at 65°C overnight. In the experiments on neuroblastoma panels, 250 mg/ml of sonicated and denatured total human placental DNA was added to the hybridization mixture in order to quench the hybridization signals due to the repeated sequences from human genomes in the phage clones. After hybridization, blots were washed twice with $2\times$ SSC at room temperature for 15 min. and with $0.2\times$ SSC plus 0.1% SDS at 65°C for 30 min.

RESULTS

Isolation and ordering of recombinant phage clones

We previously reported the isolation of 41 PERT clones which exhibited amplification with the N-myc gene (16). The twenty-four PERT clones were used as probes to further isolate the phage clones. With these PERT clones, we obtained 121 recombinant

Figure 2. An example of the linking of overlapping phage clones (2B2 and 1B10) is given. In the autoradiograms, B, H, E, S, G, T, V and K denote the lane for the DNA digested by BamHI, HindIII, EcoRI, SalI, BglII, PstI, PvuII and KpnI, respectively, and m is the size of maker DNA. The boundaries of the cloned DNA in the phage vector are indicated by arrows and the letter between two arrows represents the cleavage site for the respective enzymes. The overlapping site in 2B2 and 1B10 is also indicated by brackets. In the restriction maps deduced from the autoradiograms, the cleavage sites are represented as vertical bars with the letter denoting the respective enzyme. The direction indicated by the arrows corresponds to the direction from the bottom to the top in the autoradiograms.

phage clones from the EMBL3 genomic library. The phage clones were then sorted by hybridizing to the same PERT clone(s) that were linked to one another. This process resulted in the cloning of seventy-five phage clones. These clones were then analyzed by examining the restriction map within the cloned regions.

As cleavage sites for a single six-base-recognizing enzyme appear on average once in every 4 kb in a random sequence, sequences ranging from some 1000 to 3000 kb in length corresponding to the upper limit of the N-myc amplicon would have 2000 to 6000 cleavage sites for eight enzymes. Based on the calculation as described, the number of cleavage sites necessary to identify one stretch from another would then be 3.7 to 4.2, as calculated from the equation $(1/N)^n = 1/(M \times N)$, where n is the number of cleavage sites uniquely appearing, N ($= 8$) is the number of enzymes used and M ($= 2000$ to 6000) is the number of cleavage sites throughout the genome. Therefore, at least five cleavage sites are necessary to show that one stretch is identical to another.

Based on this assumption, we searched for pairs matching at more than five consecutive cleavage sites. An example of overlapping clones found in this way is given in Fig. 2. Two phage clones, 2B2 and 1B10, overlapped each other in an inverted manner on the portion where there are 10 unique cleavage sites, including one each for BamHI, EcoRI, PstI and KpnI, two PvuII and four HindIII sites.

Figure 3. The overall result deduced from the results obtained by the linking of overlapping phage clones given as the restriction maps of 8 different enzymes. They were incorporated into 8 different contigs denoted as Contigs A to H. In these maps, the restriction sites are represented as shown in Fig. 2. The positions of the PERT clones and N-myc and the individual phage clones are also indicated.

Figure 4. Examples of large-scale mapping using the PERT clones. Southern blot hybridizations of the DNA from the cell line MC-NB-1 digested with NotI (a), SfiI (b), NotI and SfiI (c) and SmaI (d) with the PERT clones and N-myc on HFGE are shown in these figures. The PERT clones are given below the autoradiograms. Lambda phage DNA concatemers used as a size marker are also indicated.

This process enabled the identification of eight independent contigs, named Contig A to H, in which 2 (Contig B) to 23 (Contig H) phage clones were incorporated (Fig. 3). We further determined the exact position(s) of each PERT clone within the contigs by digesting individual phage clones completely with HindIII and hybridizing them with the PERT clones as probes. The size of these contigs ranged from 25 kb to 60 kb and the total size identified as contigs was 330 kb. The N-myc gene and most of the PERT clones appeared once among these contigs, while a few (PERT clone Nos. 150 and 277) appeared twice on different contigs. The phage clones were frequently recovered in the region surrounding PERT clone No. 229 in Contig A, Nos. 150, 264 and 176 in Contig C, Nos. 212, 224 and 191 in Contig E, and Nos. 196, 201, 100, 79, 96 and 116 in Contig H, while they were much less frequent in the region surrounding N-myc and PERT clone No. 129.

Large-scale mapping with HFGE using PERT clones

To determine the range over which DNA sequences 5' and 3' of the N-myc gene are amplified, the NotI-linking probes, N-

Table 1. Summary of hybridization of NotI-, SfiI-, NotI and SfiI-, and SmaI-cut DNA of MC-NB-1 with PERT clones on HFGE

Contig	Clone number	NotI (kb)	SfiI (kb)	SfiI/NotI (kb)	SmaI (kb)
A	60	100, 420, 480, 520	110, 130	ND ^{a)}	ND ^{a)}
	53	ND ^{a)}	110, 130	70	ND ^{a)}
	229	ND ^{a)}	110, 130	70, 130	ND ^{a)}
B	N-myc A	100, 420, 520	110, 130	70	ND ^{a)}
	N-myc B	100, 420, 480, 520	110, 130	30, 60	ND ^{a)}
C	277	420, 480, 520	90, 110, 130	90, 110, 130	35
	14	ND ^{a)}	90	ND ^{a)}	ND ^{a)}
	118	ND ^{a)}	30, 90	ND ^{a)}	ND ^{a)}
	150	420, 480, 520	30, 110, 130	ND ^{a)}	15
	264	420, 480, 520	(50) ^{b)} 110, 130	30, (70) ^{b)} 110, 130	30
D	176	420, 480, 520	110, 130	60, 110, 130	35
	150	420, 480, 520	30, 110, 130	ND ^{a)}	15
E	263	420, 480, 520	110, 130	110, 130	35
	191	ND ^{a)}	130, 170	130, 170	75
F	277	420, 480, 520	90, 110, 130	90, 110, 130	35
G	129	420, 480, 520	90	ND ^{a)}	40
H	116	420, 480, 520	130	130	10
	96	ND ^{a)}	130	ND ^{a)}	50
	100	420, 480, 520	130, 170	ND ^{a)}	50
	260	ND ^{a)}	(110) ^{b)} 130, 170	130, 170	50
	1F5-1	ND ^{a)}	30	ND ^{a)}	ND ^{a)}

a) ND, not determined

b) The bands were weakly hybridized.

myc A, a 0.9 kb fragment of the 3' end, and N-myc B, a 1.1 kb fragment of the 5' end of a NotI site in exon II were used for hybridization experiments on NotI digests of MC-NB-1 DNA. N-myc A revealed three NotI fragments with sizes of 100, 420 and 520 kb, and N-myc B revealed ones those with sizes of 100, 420, 480 and 520 kb (Fig. 4 a) and Table 1). In order to determine whether the contigs identified in this study are located on these NotI fragments or not, several PERT clones located in different contigs were further examined by means of hybridization experiments. PERT clone No. 60 in Contig A identified the 100, 420, 480 and 520 kb NotI fragments. All other PERT clones, including Nos. 277, 264 and 176 in Contig C, No. 150 in Contigs C and D, No. 263 in Contig E, No. 277 in Contig F, No. 129 in Contig G and Nos. 116 and 100 in Contig H, identified three NotI fragments with sizes of 420, 480 and 520 kb, but not the fragment with a size of 100 kb (Fig. 4 a) and Table 1).

Further hybridization experiments were performed on SfiI, SfiI and NotI and SmaI digests of MC-NB-1 DNA using N-myc A, N-myc B and PERT clones located in different contigs as probes (Fig. 4 b)-d) and Table I). Both N-myc probes and the PERT clones identified 5 different SfiI fragments, 30, 90, 110, 130 and 170 kb in size. Among these, both the 110 kb and 130 kb fragments that comigrated were subdivided into two different kinds: one was split by NotI and the other was not. And the 130 kb SfiI fragment was further subdivided into two different kinds that comigrated: one was split by SmaI and the other was not. In summary, we found 9 different SfiI fragments which include three different 130 kb fragments, two different 30 kb fragments, two different 110 kb fragments, one 90 kb fragment and one 170 kb fragment. More detailed results will be given in the section on the overall relations among the contigs and the large fragments identified with HFGE.

Figure 5. An example of a DNA panel of neuroblastoma cell lines showing the different amplification of one of the phage clones, 1F7 within Contig A. Note that three bands, 2.4, 2.6 and 4.2 kb in size, were identified as commonly amplified fragments among these cell lines except for CHP134, in which two of them, 2.6 and 4.2 kb in size, were missing. The additional rearranged bands appeared on MC-NB-1 (3.8, 5.2 and 6.4 kb), TNB-1 (5.2 and 6.4 kb) and CHP126 (5.2 kb).

Analysis of amplified regions among different neuroblastoma cell lines

Using the 24 phage clones to cover the entire 330 kb region of the contigs, we examined the amplified regions among different cell lines. These experiments were done not only to find out the conservation of the amplification surrounding the N-myc region in this cell line relative to that in other cell lines, but also to identify the rearranged sites within the amplicon in this cell line and other cell lines, as well. If there is a rearranged site within the regions of the contigs, a discontinuity of the amplification in the continuous stretches within the contigs, either unamplified or amplified but different in size, will be observed. If such a discontinuity is observed in the same region among at least two cell lines, it is likely that the region is rearranged.

Figure 6. The overall results of the analysis of amplified regions among different cell lines. The fragments are indicated as either amplified (solid regions) or unamplified (open regions). When the fragment was amplified, but the size was different from the size identified in MC-NB-1, the fragments are shown as hatched regions. The fragments involved in the rearrangements are shown at the top of the restriction maps with the name of the cell lines responsible for such rearrangements. The cleavage sites for HindIII are also indicated below. The results for the cell line NB69 without amplification of the N-myc gene or any phage clones is not included. The minimally overlapping phage clones are indicated below.

Following the criteria given above, the comparisons of amplified regions between MC-NB-1 and 10 different cell lines was done and the rearranged sites within the contigs were identified. One such example is shown in Fig. 5, and the overall results are given in Fig. 6. All of the cloned regions in the contigs were amplified in MC-NB-1 and most of the regions were amplified in at least one other cell line. The regions of Contigs B and G and a part of Contig A were commonly amplified in all the cell lines examined. The amplified region within all the contigs calculated for each cell line varied from a maximum of 45% in the cell lines TNB-1 and NB19 down to a minimum of 18% in the cell line CHP134. The eight rearranged sites that seemed to occur in MC-NB-1 were found in Contigs A, C, F and H. In addition, rearranged sites that seemed to occur in other cell lines were also found. The length in which at least one rearranged site of the amplified sequences occurs in MC-NB-1 was calculated to be 41.3 kb.

Figure 7. The overall relations between the contigs and the fragments identified by HFGE. In 'Rearrangements and Amplified Regions', the closed triangles indicate the rearranged sites occurring in the cell line MC-NB-1; the hatched triangles, those in other cell lines; the open triangles, those unidentified; the closed boxes, the amplified regions observed in more than 10 cell lines (the highly conserved region in the amplification); the hatched boxes, those observed 3-9 cell lines (the moderately conserved region); and the open boxes, only 1-2 cell lines (the slightly conserved region). In 'Contigs' and 'PERT Clones', the alignment of the contigs determined using HFGE is given with the positions of the PERT clones and N-myc. The horizontal lines with double arrows indicate that both directions of the Contigs are possible. In 'PERT Clones', 1F5-1, a phage clone used to clone the region joining the region cloned by PERT No. 116, is also indicated. In 'SmaI', the only SmaI fragment that is informative is given. In 'SfiI' the possible groupings of SfiI fragments into 6 different large segments deduced from the results of HFGE are listed. The round brackets represent the regions within the contigs that are not included in the SfiI fragments. In 'NotI', any possible combination of the NotI fragments deduced from the results of HFGE is listed. The round brackets represent the regions within the contigs that are not included in the NotI fragments. Note that the NotI fragments with the head-to-tail orientation (direct repeat) are possible in Types I, II, IV and V, but not in Type III, while those with the tail-to-head orientation (inverted repeat) are possible in Types III to V, but not in Types I and II. In Types III to V, NotI fragments with a NotI site that is not linked to the site in the N-myc locus are also possible. The broken lines in Type IV indicate that the three regions encompassing Contig A, Contigs H and D, and Contigs E to G in Type IV or the two regions encompassing Contigs H and D and Contigs E to G in Type V are not localized in the same NotI fragments at the same time.

The overall relations among the contigs and the large fragments identified with HFGE

The overall relations among contigs and the large fragments identified with HFGE are shown in Fig. 7. The orders of the contigs have partly been determined based on the fragments identified with HFGE using the PERT clones as probes within the contigs. For example, Nos. 53 and 229 in Contig A identified the 70 kb NotI-SfiI fragment. This fragment was also identified by N-myc A, but not by N-myc B, while N-myc B identified the 30 and 60 kb NotI-SfiI fragments; and all these probes identified the 110 and 130 kb SfiI fragments. These results lead to the conclusion that Contig A is the neighboring segment of the 3' end of Contig B, and there are two different structures spanning the N-myc region: one has a size of 110 kb and the other, a size of 130 kb, which are the sizes of the SfiI fragments. In Contig C, Nos. 277, 14, 118, 150, 264 and 176 are lined up in this order. Among them, the first three clones, Nos. 277, 14 and 118, identified the 90 kb SfiI fragment, while No. 118 and the next clone, No. 150, identified the 30 kb fragment; and the clones next to No. 150, Nos. 264 and 176, identified the 110 and 130 kb fragments. As the 110 kb and 130 kb fragments

identified by these PERT clones were not split by NotI, they are different from those identified by the PERT clones in Contig A and N-myc A and N-myc B. There was a SfiI site within No. 118 (data not shown). Consequently, it is likely that there are two types of structures for the SfiI fragments spanning the region of Contig C: one is composed of the 130 kb, 30 kb and 90 kb SfiI fragments and the other of the 110 kb, 30 kb and 90 kb SfiI fragments which are lined up one after another in this order. It should be noted that Nos. 150 and 277 appear twice in different contigs. On the basis of this order, the 110 and 130 kb fragments identified by No. 150 are not localized within the regions of Contig C, but rather in the regions of Contig D. As we consider that the 90 kb SfiI fragment detected by No. 129 to be identical to the fragment identified by Nos. 277, 14 and 118, Contig G must be the neighbor of Contig C. The 110 kb and 130 kb SfiI fragments detected by No. 277 are not localized in the region of Contig C, but in the region of Contig F. As all the clones, including No. 263 in Contig E, No. 277 in Contig F, and No. 176 in Contig C, detected the 35 kb SmaI fragment, and all these identified 110 kb and 130 kb SfiI fragments, which were not split by NotI, it is likely that Contigs E and F are lined up in this order and adjacent to Contig C on the side opposite Contig G. The 130 kb SfiI fragment detected by Nos. 116, 96, 100 and 260 in Contig H seems to be the same as the one detected by No. 150 in Contig D. It was not split by NotI, but based on the size of the SmaI fragments, it is different from the one detected by the PERT clones in Contigs E and F. The 170 kb SfiI fragment detected by No. 260 in Contig H seems to be the same as the one detected by No. 191 in Contig E.

Based on the observations described in the previous section, there must essentially be five types of NotI fragments spanning neighboring N-myc genes. These were referred to as Types I to V (Fig. 7). Type I is the shortest fragment (truncated?), 100 kb in size, in which only Contigs A and B are localized. Type II is the longest fragment, 520 kb in size, in which all the contigs are localized. As both N-myc probes detected these fragments, both Types I and II seem to orient the N-myc gene in such a way as to form a direct array. Type III is the 480 kb fragment. As only N-myc A did not identify this fragment, the 3' end portion next to the NotI site in exon II is not located within this fragment. This means that Type III seems to orient the 5' portion of the N-myc gene in such a way as to form an inverted array, or orient the 5' portion of the N-myc gene at one end and a NotI site that does not link to the site in exon II at the other end. As represented in Fig. 7, the 420 kb fragment must be a doublet, since the total size of a SfiI fragment was estimated to be a minimum of 520 kb [110 kb (Contigs A and B) + 30 kb (Contig H) + 130 kb (Contigs H and D) + 130 kb (Contigs E, F and C) + 30 kb (Contig C) + 90 kb (Contigs C and G)]. It is therefore most likely that there are various kinds of 420 kb NotI fragments. Following this assumption, three kinds of 420 kb NotI fragments possibly exist: one orients N-myc in such a way that there is a direct repeat (Types IV-1 and V-1), another in such a way that there is an inverted repeat (Types IV-2 and V-2) and the third in such a way that either the 5' portion (Type IV-3) or the 3' portion (Type V-3) of the N-myc gene is at one end and a NotI site that does not link to the site in exon II is at the other end.

DISCUSSION

Our present work aims to elucidate the complex process of N-myc amplification. To achieve this, it will eventually be necessary to determine the structural organization of the N-myc amplicons.

There are several groups who have been trying to clone and identify the N-myc amplicon(s) (13, 15, 17). Unfortunately, none of the groups has so far succeeded in cloning and identifying all of them. The difficulties associated with these studies are that tremendously large amplicons undergo complex rearrangements and are tandemly repeated many times. These cannot readily be overcome using the conventional one-step chromosome walking methods.

One approach to solving the problem involves the use of YAC technology (27). Alternatively, our strategy was to begin with cloning from multiple sites. In order to do so, it was necessary to have a sufficient number of amplified DNA sequences evenly distributed throughout the amplicon(s). Fortunately, we have succeeded in cloning a number of amplified fragments in the neuroblastoma cell line MC-NB-1 following the PERT strategy, an efficient means of cloning amplified sequences (16). These clones could be used in the cloning method to bridge the gaps between these amplified sequences. This method has permitted cloning of the whole *E. coli* chromosome (21). Finally, we were able to obtain a contiguous clone set composed of 8 independent contigs ranging from 25 to 60 kb in size that encompassed a total of 330 kb in the amplicon(s), which covered some 60–80% of the entire amplicon.

The sizes of the amplicons identified with HFGE using both the linking N-myc clones and many other PERT clones scattered in the contigs were similar to those estimated by means of Roninson's gel (28) or by pulsed field gel electrophoresis (18). Our approach is more appropriate for defining the distance between neighboring N-myc genes than using the N-myc probes only (18, 19), if there is another NotI site(s) between the neighboring N-myc genes. Our interpretation that the 100 kb, 520 kb and one of the 420 kb NotI fragments are tandemly ordered in a head-to-tail array (direct repeat) is consistent with that given by Amler and Schwab (18). The data presented here provide further evidence that in this cell line there are a 480 kb and some 420 kb fragments that are ordered in tail-to-head (inverted repeats) arrangements, or with a NotI site that is not linked to N-myc.

The important observation made here is that there are only a few amplicons in this cell line and either their constituents mostly overlap one another, or their structures are closely related. This strongly suggests that these amplicons have the same origin.

The commonly amplified regions (a part of Contig A, Contig B and Contig G) defined using overlapping clones were relatively few compared with the total number of amplified regions cloned, although most of them were amplified with at least one other cell line. In this cell line, the number of rearrangements occurring once in every 40 kb, which is equivalent to some ten for each amplicon with a size in the range of 400–500 kb, was comparable with that estimated previously in this cell line by simply dividing the number of rearranged bands by the sum of the amplified sequences examined as a HindIII band, the value being 19.2 rearrangements/1000 kb amplified fragment (16). The distribution of the sites suggests that the rearrangements occurred rather evenly throughout the amplicon(s).

These results are somewhat comparable with those of Shiloh *et al.* (20), but strikingly different from those of Zehnauer *et al.* (15). Shiloh *et al.* found that 6 clones isolated from the cell line IMR-32 and NB9 were amplified in an average of 37% of the regions identified by 6 clones in the tumors and 53% in the cell lines examined where N-myc was amplified, and their clones detected 5.1 rearrangements per 1000 kb amplified fragment

among the tumors, and 14.1 in the cell lines. However, Zehnbaauer *et al.* found that 48 clones isolated from the cell line NGP were amplified in an average of 81% of the regions identified by 48 clones in 12 tumors, and their clones detected only 0.4 rearrangements per 1000 kb amplified region. The difference in the values of the amplified regions probably comes from the fact that some clones were derived from amplicons with a low degree of sequence homogeneity and some from amplicons with a high degree. In this context, our cell line, as well as other cell lines, have amplicons that are different from one another to some extent. The difference in the number of rearrangements seems to be partly due to the data reported by Zehnbaauer *et al.* (15). Those numbers must be underestimated, since they neglected some minor fragments which appeared after overexposure. Furthermore, at least six different amplicons were identified in the cell line NGP (18), so the amplicons must undergo rearrangement to some extent.

Despite the low degree of sequence homogeneity among amplicons from different cell lines and the significant number of rearrangements in the amplicons within a cell line, there are only several types of amplicons with a high degree of sequence homogeneity. These observations lead to the idea of the evolution of N-myc amplicons, that is an original amplified DNA domain is first constructed, then it increases in number, and finally gives rise to several amplicons which are structurally closely related, but not identical to one another. The rearrangements are involved both in the formation of the original amplified DNA domain in the initial stage and in the increase in their numbers in the later stage. Shiloh *et al.* (20) proposed the so-called 'nested subset model' to describe the heterogeneous organization of the amplicons observed among different cell lines. This model permits an amplified DNA domain with N-myc in the most central position and the other DNA fragments at increasingly peripheral positions. Our results suggest that more complex rearrangements, such as a joining and an integration of different fragments of several tens of kb in size, which are either amplified or unamplified in other cell lines, are involved in the amplification process.

In summary, there have been no reports so far on isolating the ordered clone set which encompasses most of the entire N-myc amplicon. This clone set will therefore be of great value in elucidating the mechanism of N-myc amplification.

ACKNOWLEDGEMENT

We greatly appreciate Dr. Bruce Korf (Genetics Division, The Children's Hospital, Boston, MA, USA) for his encouragement of our research. We are very grateful to Professor Hikoya Hayatsu (Faculty of Pharmaceutical Sciences, Okayama University, Okayama, Japan) for his critical reading of this manuscript. We also thank Miss Hiromi Chiba for her excellent technical assistance.

REFERENCES

- Alitalo, K. and Schwab, M. (1989) *Adv. Cancer Res.*, **47**, 235–281.
- Schimke, R. T. (1984) *Cell*, **37**, 705–713.
- Schwab, M. (1986) In Kahn P, Graf T (eds.), *Oncogene and Growth Control: Amplification of Proto-Oncogenes and Tumor Progression*. Springer Verlag, Heidelberg, pp332–339.
- Stark, G. R., Dabatisse, M., Guilotto, E. and Wahl, G. M. (1989) *Cell*, **57**, 901–908.
- Stark, G. R. and Wahl, G. M. (1984) *Ann. Rev. Biochem.*, **53**, 447–491.
- Schwab, M. (1985) *Trends Genet.*, **1**, 271–275.
- Bartram, C. R. and Bertold, F. (1987) *Eur. J. Pediatr.*, **146**, 162–165.
- Brodeur, G., Seeger, R. C., Schwab, M., Varmus, H. E. and Bishop, J. W. (1984) *Science*, **224**, 1121–1124.
- Cohl, S. L., Herst, C. V., Maurer, H. S. and Rosen, S. T. (1987) *J. Clin. Oncol.*, **5**, 1141–1144.
- Nakagawara, A., Ikeda, K., Tsuda, T. and Higashi, K. (1988) in Evans A, D'Angio G, Seeger R (eds.) *Advance in Neuroblastoma Research: Biological Characterization of NMYC Amplified Neuroblastoma in Patients over One Year of Age*. Alan R. Liss, Inc., New York, Vol. 2, pp31–39.
- Seeger, R. C., Brodeur, G. M., Sather, H., Dalton, S., Siegel, E., Wong, K. Y. and Hammond, D. (1985) *N. Engl. J. Med.*, **313**, 1111–1116.
- Hunt, J. D., Valentine, M. and Teraba, A. (1990) *Mol. Cell. Biol.*, **10**, 823–829.
- Kanda, N., Shreck, R., Alt, F., Bruns, G., Baltimore, D. and Latt, S. A. (1983) *Proc. Natl. Acad. Sci. USA.*, **80**, 4069–4073.
- Roninson, I. B. (1983) *Nucl. Acid. Res.*, **11**, 5413–5431.
- Zehnbaauer, B. A., Small, D., Brodeur, G. M., Seeger, R. and Vogelstein, B. (1988) *Mol. Cell. Biol.*, **8**, 522–530.
- Nishii, Y., Akiyama, K. and Korf, B. (1991) *Mammalian Genome*, in press.
- Shiloh, Y., Rose, E., Colletti-Freener, C., Korf, B., Kunkel, L. M. and Latt, S. A. (1987) *Gene*, **51**, 53–59.
- Amler, L. C. and Schwab, M. (1989) *Mol. Cell Biol.*, **9**, 4903–4913.
- Kato, H., Okamura, K., Kurosawa, Y., Kishikawa, T. and Hashimoto, K. (1989) *FEBS*, **2**, 529–535.
- Shilo, Y., Korf, B., Kohl, N. E., Brodeur, G. M., Harris, P., Kanda, N., Seeger, R. C., Alt, F. and Latt, S. A. (1986) *Cancer Res.*, **46**, 5297–5301.
- Kohara, Y., Akiyama, K. and Isono, K. (1987) *Cell*, **50**, 495–508.
- Aldridge, J., Kunkel, L., Burns, G., Tantravahi, U., Lalande, M., Brewster, T., Moreau, E., Wilson, M., Bromley, W., Roderick, T. and Latt, S. A. (1984) *Am. J. Hum. Genet.*, **36**, 546–564.
- Schwartz, D. C. and Cantor, R. C. (1984) *Cell*, **37**, 67–75.
- Smith, C. L. and Cantor, R. C. In Davis, K. (ed.), *Genomic Analysis-Pulsed field gel electrophoresis and the technology of large DNA molecules*. IRL press.
- Feinberg, A. P. and Vogelstein, B. (1984) *Anal. Biochem.*, **132**, 6–13.
- Feinberg, A. P. and Vogelstein, B. (1984) *Anal. Biochem.*, **137**, 266–267.
- Burke, D. V., Carle, G. F. and Olson, M. V. (1987) *Science*, **236**, 806–811.
- Kinzler, K. W., Zehnbaauer, B. A., Brodeur, G. M., Seeger, R. C., Trent, J. M., Meltzer, P. S. and Vogelstein, B. (1986) *Proc. Natl. Acad. Sci. USA.*, **83**, 1031–1035.