Chromosomal Assignments of Novel Genes Expressed in HL60 Granulocytes

Atsushi FUKUSHIMA,* Kenichi MATSUBARA,* Katsuji MURAKAWA,** Junji YOSHII,†† Masahiro YOKOYAMA,† and Kousaku OKUBO

Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565, Japan

(Received 23 March 1994)

Abstract

By collecting 3'-directed cDNA sequences called gene signatures (GSs) on a large scale, it is possible to make an expression profile of genes in a particular tissue, as well as discovering a number of novel genes. A total of 305 novel GSs collected from granulocytoid cells derived from HL60, a human promyelocytic leukemia cell line, by exposure to dimethyl sulfoxide were radiolabeled and used for Southern blot analyses to determine the copy number of the corresponding genes. Of these, 198 GSs identified as representing single-copy genes were then used as probes for hybridization analyses using a monochromosomal hybrid cell DNA panel. Sixty-nine of them were assigned to individual chromosomes. These results demonstrate that the chromosomal distribution of the GSs seems not to be proportional to the cytogenetic length of each chromosome.

Key words: gene signature; HL60; granulocytes; chromosome assignment

1. Introduction

Advances in sequencing technology have enabled large scale collection of partial cDNA fragments of active genes in a given tissue(s) of interest. As the fragments become comprehensive, mapping these genes, along the physical and contig maps of the genome, or at least assigning them to chromosomes, has become an important issue. 2 3 4 Comprehensive mapping of active genes can be used to identify the genes responsible for genetic disorders. Especially for positional cloning, efforts to make an expression profile of the gene in the tissues will complement other aspects of cDNA analyses which involves predicting the function of the gene by the full length cDNA sequence.

We have initiated systematic collection of 3'-directed partial cDNA sequences that cover the sequences from poly(A) to the nearest Mbol site (GATC), to construct a gene expression profile database with as many cDNA libraries as are available (Okubo et al. to be published). As these partial cDNA sequences, called gene signatures (GSs), are derived from the same region of transcripts, they can not only "tag" the expressed genes, but also allow clone-to-clone comparisons, thus making it possible to measure the number of mRNA copies in the original population. In this sense, the GSs are different from the expressed sequence tags (ESTs) obtained with the random-primed cDNA libraries of Adams et al.1 with which one cannot distinguish whether two sequences are from different genes or whether they represent different regions of the same gene. Collection of GSs for gene expression profiling has been done with a promyelocyte cell line, HL60,2 with and without induction by DMSO which converts the cells into granulocytoids.3 Nearly 1000 GSs each have been analyzed with HL60 and granulocytoids, among which 83% were of novel genes.

In a previous effort, we initiated assigning the 160 GSs obtained from human liver cell line HepG2 to human chromosomes by Southern hybridization using a monochromosomal hybrid cell panel.4 We now have extended the analyses to the novel GSs of granulocytoid cells, to provide mapping information to those researchers who are interested in studying granulocytes, and to see whether such GSs might be distributed among chromosomes in the same fashion as observed with HepG2.

2. Materials and Methods

2.1. cDNA sequences of granulocytoid cells

A 3'-directed cDNA library was constructed from mRNA of DMSO-treated HL60 cells, and the sequences of 1118 components have been determined (Okubo et al.
to be published).

### 2.2. Genomic DNA and Southern blot analysis

Human genomic DNA was prepared from a lymphoblastoid cell line GM0130b that has a normal karyotype, as described previously. A panel of monochromosomal hybrid cells whose integrity had been monitored by in situ hybridization has been described. The genomic DNA from each hybrid cell was digested overnight with EcoRI and aliquots (5 μg) from each of the digests were electrophoresed side-by-side in a 0.8% agarose gel, followed by transferring to a Hybond N+ membrane (Amersham) using 0.4N NaOH. The membrane was rinsed in 2xSSC, and stored at 4°C until use.

GSs which are novel (Okubo et al. to be published), and have more than 150 bp were selected and the cDNA inserts were amplified by polymerase chain reaction (PCR) with primers as described. The PCR products were isolated by electrophoresis in a 2% low-melting temperature agarose (Nusieve to SeaPlaque, 3:1), and then eluted from the gel by melting at 65°C and digesting with β-AgaraseI (Bio Labs) at 40°C for 1 h. They were labeled with [α-32P]dCTP by a random priming method using a commercial kit (Amersham). Hybridization was performed at 65°C in a high salt buffer containing 6xSSC, 1xDenhardt’s, and 0.5% SDS in the presence of 0.1 mg/ml sonicated, denatured salmon sperm DNA. The membranes were washed in 2xSSC containing 0.1% SDS at 65°C for 30 min and twice for 30 min in 0.1xSSC containing 0.1% SDS at 65°C, and then analyzed by a Fuji BAS-2000 imaging analyzer.

### 3. Results and Discussion

#### 3.1. Copy number in the genome

Among the 858 GSs species collected in the 3'-directed cDNA library prepared from granulocytoid cells, 447 had novel sequences and were larger than 150 bp (Okubo et al. to be published). These GSs were purified from the plasmids, radio-labeled, and used as probes for hybridization with human genomic DNA EcoRI-digested and blotted on a membrane to determine the number of copies. The results are shown in Table 1. Among the 305 GSs tested, 198 (65%) showed a single band, 37 (12%) double bands, and 20 (7%) 3 bands or more. The proportion of genes which produced a single band was larger than in HepG2 GSs (47%). The difference could reflect the difference in profiles of active genes in the two cells, since HepG2 are replicating cells and granulocytoids are highly differentiated non-growing cells. Of course, there must be a large number of different cell-specific genes (Okubo et al. to be published). These problems must be elucidated in future studies.

#### 3.2. Chromosomal assignments

To assign the chromosomal locations, GSs that showed a single band in genomic analyses were freed from the vector plasmid, radiolabeled and used as hybridization probes against a set of DNA from a monochromosomal hybrid cell panel. Ninety-three GSs were tested in this study. Figure 1 shows an example of hybridization. In Fig. 1a, where gs001526 was used as a probe, a human-specific band appeared only in lane 21. With gs001038, a band appeared in lane 17 (Fig. 1b). Thus, these GSs were assigned to chromosomes 21 and 17, respectively. Sixty-nine GSs were assigned to chromosomes in this manner and the results are listed in Table 2. Among the 24 of the GSs that could not be assigned to chromosomes, 4 produced bands that overlapped with cross-reacting rodent DNA. Other 20 GSs showed no band on monochromosomal hybrid cell DNA lanes. Such failures probably arose from undetected deletions in the set of human chromosomes in the panel. Among the 69 assigned GSs, 3 (gs001084, gs001134, gs001240) have been analyzed independently by Murakawa et al. by a PCR method, resulting in the consistent assignments.

#### 3.3. Distribution of GSs along chromosomes

The distribution of the 69 assigned GSs on chromosomes is shown in Fig. 2a. Their distribution is not proportional to the cytogenetic length of the chromosomes. For example, the numbers are high with chromosomes 13, 17, and 20, and low with chromosomes 15 and 18. Murakawa et al. assigned 155 GSs from granulocytoid cells prepared similarly by PCR method, and observed that a larger proportion of GSs were assigned to chromosomes 13, 17, and 20. For reference, their finding is reproduced.

### Table 1. Number of bands that appear in Southern analyses of EcoRI-digested human genomic DNA using radiolabeled GS probes from granulocytoid cells (a) and HepG2 cells (b).

| No. of bands | 1 | 2 | 3 | 4 | >5 | sm | non | total |
|--------------|---|---|---|---|----|----|-----|------|
| (a) granulocytoid cells | | | | | | | | |
| No. of GS | 198 | 37 | 9 | 2 | 9 | 17 | 33 | 305 |
| % | 65 | 12 | 3 | 1 | 3 | 5 | 11 | 100 |
| (b) HepG2 | | | | | | | | |
| No. of GS | 75 | 33 | 7 | 7 | 17 | 7 | 14 | 160 |
| % | 47 | 21 | 4 | 4 | 11 | 4 | 9 | 100 |

(b) from Fukushima et al.(1994). sm: smears, non: no band formation.
Figure 1. Chromosome distribution of single copy GSs. Granulocytoid cells induced by DMSO (a). For comparison, data modified from Murakawa et al. with similarly induced granulocytoid (b) and from Fukushima et al. with HepG2 are reproduced (c).
in Fig. 2b. It is too premature to discuss this bias. However, there is a possibility that it may come from biased distribution of genes along the chromosomes in general. Alternatively, distribution of active genes in the granulocytoid cells might be highly biased. Results of similar works with 160 GSs from HepG2, the derivative of liver cells, showed a relatively high number of GSs on chromosome 17, but very few or none on chromosomes 15 and 18 (Fig. 2c). Even though the number of analysed samples is small, these findings seem to support the idea that gene density is higher on some chromosomes, such as chromosome 17, whereas it is lower on chromosomes 15 and 18. Furthermore, Polymeropoulos et al. have observed that the number of ESTs assigned to chromosome 13 is lower than anticipated by its length. With HepG2 cells, the number of assigned GSs was low, but with the granulocytoid cells, this was not true. Similarly, the number of assigned GSs on chromosome 20 seems to differ between...
Table 2. Chromosome assignments of single copy GSs from granulocytoid cells as obtained by inducing HL60 with DMSO.

| GS no. (a) | acc. no. (b) | clone | size (c) | b.m. (d) | Chr. no. (e) | Mo | Ha |
|------------|--------------|-------|----------|----------|--------------|----|----|
| gs001519   | D20544       | pm0106| 236      | h        | 2            | 0  | 0  |
| gs001236   | D20289       | pm0117| 233      | h        | 8            | 0  | 0  |
| gs001505   | D20530       | pm0220| 282      | s        | 3?           | 0  | 0  |
| gs001433   | D20459       | pm0256| 217      | s        | 9            | 0  | 0  |
| junk       | D19729       | pm0264| 226      | —        | 13           | 0  | 0  |
| gs000910   | D19948       | pm0302| 342      | h        | 3?           | 0  | 0  |
| gs001311   | D20337       | pm0311| 207      | h        | 20           | 0  | 0  |
| gs001084   | D20111       | pm0315| 375      | s        | 14,16        | 0  | 0  |
| gs001538   | D20563       | pm0333| 611      | s        | 13           | 4  | 4  |
| gs001086   | D20113       | pm0349| 378      | h        | 10           | 0  | 0  |
| gs001385   | D20411       | pm0358| 176      | s        | 17           | 0  | 0  |
| gs001141   | D20167       | pm0377| 315      | h        | 11           | 0  | 0  |
| gs001359   | D20385       | pm0410| 320      | s        | 17           | 0  | 0  |
| gs001501   | D20526       | pm0420| 153      | s        | 20           | 0  | 0  |
| gs000991   | D20020       | pm0434| 571      | s        | 4            | 1  | 1  |
| gs001067   | D20094       | pm0516| 692      | s        | 3            | 1  | 1  |
| gs001524   | D20549       | pm0577| 276      | s        | 17           | 0  | 0  |
| junk       | D20028       | pm0605| 294      | —        | 13           | 0  | 1  |
| gs001526   | D20551       | pm0614| 319      | s        | 21           | 0  | 0  |
| junk       | unreg.       | pm0649| 373      | —        | 4            | 0  | 0  |
| gs001070   | D20097       | pm0652| 442      | s        | 3            | 0  | 0  |
| junk       | unreg.       | pm0661| 423      | —        | 4            | 0  | 0  |
| gs001534   | D20559       | pm0775| 155      | s        | 19           | 0  | 0  |
| gs001075   | D20124       | pm0883| 385      | h        | 1            | 0  | 0  |
| gs001097   | D20124       | pm0903| 356      | s        | 10           | 1  | 1  |
| gs001419   | D20445       | pm0906| 152      | s        | 1            | 0  | 0  |
| gs001079   | D20106       | pm0907| 380      | s        | 5            | 1  | 1  |
| gs000191   | D11848       | pm0908| 278      | h        | 17           | 0  | 0  |
| gs001537   | D20562       | pm1040| 319      | s        | 3            | 0  | 0  |
| gs001394   | D20420       | pm1057| 154      | h        | 5            | 0  | 0  |
| gs001179   | D20205       | pm1104| 287      | s        | 6            | 0  | 0  |
| gs001397   | D20423       | pm1132| 159      | s        | 11           | 0  | 0  |
| gs001309   | D20335       | pm1136| 291      | s        | 7            | 0  | 0  |
| gs001338   | D20364       | pm1137| 289      | s        | 22           | 0  | 0  |
| junk       | unreg.       | pm1187| 247      | —        | 4            | 4  | 0  |
| gs001208   | D20234       | pm1232| 310      | h        | 14           | 0  | 0  |
| gs001210   | D20236       | pm1235| 283      | s        | 20           | 0  | 0  |
| gs001253   | D20279       | pm1240| 238      | s        | X            | 2  | 1  |
| junk       | unreg.       | pm1250| 333      | —        | 13           | 3  | 1  |
| gs001134   | D20160       | pm1447| 326      | h        | 1,5          | 0  | 0  |
| junk       | unreg.       | pm1523| 255      | —        | 6            | 0  | 0  |
| gs001378   | D20404       | pm1604| 168      | h        | 13           | 0  | 0  |
| gs001232   | D20258       | pm1312| 252      | s        | 1            | 0  | 0  |
| gs000053   | D11797       | pm1614| 412      | h        | 6            | 0  | 0  |
| junk       | D20130       | pm1615| 355      | —        | 17           | 1  | 1  |
| gs001112   | D20139       | pm1645| 385      | s        | 4?           | 0  | 0  |
| gs001262   | D20288       | pm1682| 233      | s        | 6            | 0  | 0  |
| gs001377   | D20403       | pm1701| 169      | h        | 20           | 0  | 0  |
| gs001083   | D20110       | pm1711| 376      | s        | 2            | 1  | 1  |
| gs001201   | D20227       | pm1726| 276      | s        | 7            | 0  | 0  |
| gs001122   | D20149       | pm1756| 377      | s        | 3            | 0  | 0  |
| gs001382   | D20408       | pm1741| 167      | h        | 2            | 0  | 0  |
| gs001221   | D20247       | pm1814| 260      | s        | 16           | 0  | 0  |
There are a significant number of human GSs that cross
hybridize with rodent genomic DNA. In this study, 57
GSs among the 255 GSs that gave clear hybridization
bands belonged to this category. Figure 1b shows an
example of such cross hybridization. With gs001038, a
single band (2.2 kb) appeared with human DNA (lane
H), whereas a 4.6-kb cross-reacting band appeared with
Chinese hamster DNA. Two bands (4.8 kb and 1.0 kb)
appeared with the same probe with mouse genomic DNA.
The number of cross-hybridizing bands with rodent DNA
appeared with the same probe with mouse genomic DNA.

Table 2. Continued

| GS no. (a) | acc. no. (b) | clone size (c) | b.m. (d) | Chr. no. (e) | background (f) |
|-----------|-------------|---------------|--------|-------------|----------------|
| gs001502  | D20527      | pm1890 413    | s      | 20          | 0 0            |
| gs001038  | D20096      | pm2056 433    | s      | 17          | 2 1            |
| gs001072  | D20099      | pm2071 388    | h      | 17          | 0 0            |
| gs001231  | D20257      | pm2081 253    | s      | 9           | 0 0            |
| gs001349  | D20375      | pm2412 183    | s      | 19          | 0 0            |
| gs001222  | D20248      | pm2446 209    | h      | 9           | 0 0            |
| gs001040  | D20068      | pm2489 431    | s      | 7           | 0 0            |
| gs001037  | D20065      | pm2504 435    | s      | 16          | 0 0            |
| gs001240  | D20266      | pm2512 245    | h      | 6,16        | 0 0            |
| gs001019  | D2048      | pm2612 468    | h      | 11          | 0 1            |
| gs001177  | D20203      | pm2672 288    | s      | 22          | 1 1            |
| gs001028  | D20057      | pm2716 452    | s      | 13          | 0 0            |
| gs001205  | D20231      | pm2792 274    | s      | 12          | 1 1            |
| gs001243  | D20269      | pm2793 244    | s      | 1           | 1 1            |
| gs001156  | D20182      | pm0207 437    | s      | 17          | 0 0            |
| junk      | unreg.      | pm1676 221    | h      | 5           | 0 0            |

Radiolabeled GS DNA's were used as probes in Southern analyses using total human genomic DNA or DNA's from monocho-
mosome hybrid cell panel 1 (Fukushima et al., 1993). (a) Gene Signatures. Junk.; not categorized in GSs, as having more
than 5 unidentified nucleotide within the sequence of 100 bp 5' end region. (b) Accession Number (DDBJ/GenBank/EMBL).
unreg.; unregistered. (c) Size, in nt, of the GS. (d) Body Map. s.; GS found only in granulocytoid cell library, h.; GS found
in more than two libraries after analyzing 11 different libraries (HepG2, HL60, DMSO-differentiated HL60, TPA-differentiated
HL60, colon mucosa, adipose, lung, osteoblast, small cell carcinoma, 19-week fetal liver, & 40-week fetal liver: Matsubara &
Okubo, 1993). (e) Number of cross-hybridizing bands with mouse (Mo) and Chinese hamster (Ha) genomic DNAs.

3.4. Conserved GSs

There are a significant number of human GSs that cross
hybridize with rodent genomic DNA. In this study, 57
GSs among the 255 GSs that gave clear hybridization bands belonged to this category. Figure 1b shows an
example of such cross hybridization. With gs001038, a
single band (2.2 kb) appeared with human DNA (lane
H), whereas a 4.6-kb cross-reacting band appeared with
Chinese hamster DNA. Two bands (4.8 kb and 1.0 kb)
appeared with the same probe with mouse genomic DNA.
The number of cross-hybridizing bands with rodent DNA
appears with the same probe with mouse genomic DNA.

 Needless to say, in the near future fine mapping of the
genes reported here along chromosomes is necessary. It
will be extremely interesting, then, to identify the loci of
tissue- or differentiation-specific genes that can be dis-
covered by comparing expression profiles obtained with
different cells or tissues.9,10 At this stage, among the 69
GSs in Table 2, 41 were observed only in the granulo-
cytoid cell library. Thus, the combination of expression
profiling using GSs and comprehensive mapping supple-
ments the analyses of genome structure, and also helps
to locate cell-specific genes of interest along the chromo-
some.

Acknowledgments: We are grateful to Yuko Kojima,
Hisae Yoshinari, Junko Arimoto, Kumiko Takagi, and
Yasuko Nouni for their excellent technical and secretarial
assistance. This work was supported in part by a Grant-
in-Aid for Creative Basic Research (Human Genome Pro-
gram) from the Ministry of Education, Science and Cultu-
re, Japan.

References

1. Adams, M. D., Kelly, J. M., Gocayne, J. D., Dubnik, M.,
Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A.,
Olde, B., Moreno, R. F., Kerlavage, A. R., McCombie,
W. R., and Venter, J. C. 1991, Automated partial se-
quencing of cDNAs: “Expressed sequence tags” and the human genome, *Science*, **252**, 1651–1656.

2. Collins, S. J., Gallo, R. C., and Gallagher, R. E. 1977, Continuous growth and differentiation of human myeloid leukemic cells in suspension culture, *Nature*, **270**, 347–349.

3. Collins, S. J., Ruscetti, F. W., Gallagher, R. E., and Gallo, R. C. 1978, Terminal differentiation of human promyelocytic leukemia cells induced by dimethyl sulfoxide and other polar compounds, *Proc. Natl. Acad. Sci. USA*, **75**, 2458–2462.

4. Fukushima, A., Okubo, K., Sugino, H., Hori, N., Matoba, R., Niiyama, T., Murakawa, K., Yoshii, J., Yokoyama, M., and Matsubara, K. 1994, Chromosomal Assignments of HepG2 3'-directed Partial cDNA Sequences by Southern Blot Hybridization using Monochromosomal Hybrid Cell Panel, *Genomics*, in press.

5. Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y., and Matsubara, K. 1992, Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression, *Nature Genet.*, **2**, 173–179.

6. Murakawa, K., Matsubara, K., Fukushima, A., Yoshii, J., and Okubo, K. 1994, Chromosomal assignments of 3'-directed partial cDNA sequences representing novel genes expressed in granulocytic cell, *Genomics*, in press.

7. Polymeropoulos, M. H., Xiao, H., Sikela, J. M., Adams, M., Venter, J. C. and Merrill, C. R. 1993, Chromosomal distribution of 320 genes from a brain cDNA library, *Nature Genet.*, **4**, 381–386.

8. Warden, C. H., Mehrabian, M., He, K., Yoon, M., Diep, A., Xia, Y., Wen, P., Svenson, K. L., Sparkes, R. S., and Lusis, A. J. 1993, Linkage mapping of 40 randomly isolated liver cDNA clones in the mouse, *Genomics*, **18**, 295–307.

9. Matsubara, K. and Okubo, K. 1993, Identification of new genes by systematic analysis of cDNAs and database construction, *Current Opinion*, **4**, 672–677.

10. Matsubara, K. and Okubo, K. 1993, cDNA analyses in the human genome project, *Gene*, **135**, 265–274.