High-throughput Production of Recombinant Antigens for Mouse KIAA Proteins in *Escherichia coli*: Computational Allocation of Possible Antigenic Regions, and Construction of Expression Plasmids of Glutathione-S-transferase-fused Antigens by an *in vitro* Recombination-assisted Method

Yasuhiro HARA,1 Kiyo SHIMADA,1 Hiroshi KOHGA,1 Osamu OHARA,2,3 and Hisashi KOGA1,2,*

Chiba Industry Advancement Center, 2-6 Nakase, Mihama-ku, Chiba 261-7126, Japan,1 Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan,2 and Laboratory of Immunogenomics, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama City, Kanagawa, 230-0045, 8, Japan

(Received 26 May 2003; revised 3 June 2003)

Abstract

Since the end of 2001, we have conducted a project to isolate and determine entire sequences of mouse cDNA clones which encode the polypeptides corresponding to human KIAA proteins. Towards the ultimate goal of this project to clarify the biological functions of KIAA genes, we have set production of antibodies against mouse KIAA gene products based on their sequence information as the next important stage. As the first step, we developed a high-throughput system utilizing shotgun clones generated during entire sequencing of mouse KIAA cDNAs. The system consists of the following three parts: (1) Shotgun clones encoding regions suitable for production of antigens were selected using a newly developed browsersystem; (2) the protein-coding sequences of the selected shotgun clones were transferred into an expression vector by *in vitro* recombination-assisted method in a 96-well format, and expressed as glutathione S-transferase fusion proteins in *Escherichia coli*; and (3) the solubility of the recombinant antigens were preliminarily assessed in a small-scale culture and then large-scale production and purification was performed using glutathione-affinity beads or retrieval from polyacrylamide gels depending on their solubility. Using these systems, we successfully produced and purified 400 antigens for production of mKIAA antibodies to date.

Key words: mKIAA; antigen; antibody; high-throughput; *in vitro* recombination-assisted method
of KIAA genes is considered to be only about 5% of the total number of human protein-coding genes, the current KIAA cDNA collection accounts for one-third of genes encoding large proteins (>1000 amino acid residues) in the Ensembl database (http://www.ensembl.org/). Furthermore, taking into consideration the fact that KIAA cDNAs were isolated mainly from the brain, we consider that the KIAA cDNA collection is comprehensive enough to characterize large proteins produced in the brain.

To better understand the physiological function of KIAA genes and KIAA proteins, we recently initiated a project to construct genomic and proteomic resources for functional analysis of KIAA genes in animal models.6,7 Experiments using animals do not carry with them the ethical limitations imposed on the use of human materials, thus enabling us to apply genetic manipulation technology to elucidate the physiological roles of the genes. Among the many animal model systems currently available, we selected the mouse system as the first choice for the following reasons: (1) the mouse is one of the most widely used animals for the study of mammalian gene functions; (2) approximately 99% of mouse genes have human homologues;9 (3) the mouse draft genome sequence is already available (http://www.ncbi.nlm.nih.gov/genome/guide/mouse/index.html, http://www.ensembl.org/Mus_musculus/); (4) mutant mouse resources are being well developed, and many lines of mutant mice are also available9,10 (http://imsm.har.mrc.ac.uk, http://www.jax.org, http://www.emma.rm.cnr.it/); (5) it is possible to investigate the genetic linkage for quantitative differences in complex behaviors;11 and (6) the mouse cDNA resource is also widely available to the public (http://www.ncbi.nlm.nih.gov/, http://www.gsc.riken.go.jp/e/FANTOM/, http://lgsun.grc.nia.nih.gov/cDNA/cDNA.html). We thus believe that exploration of physiological functions of mouse KIAA genes will give us insight into the physiological functions of KIAA genes in human.

Besides providing mRNA sequence information, cDNA clones serve as an indispensable reagent for the functional analyses of genes. Thus, we first attempted to accumulate mouse KIAA-homologous cDNA (mKIAA) clones as a genomic resource. We have already established a high-throughput system to select mouse KIAA-homologous (mKIAA) cDNA clones based on the end sequence information from size-fractionated cDNA libraries, entirely sequenced mKIAA cDNA clones, and released the structural features of 500 mKIAA cDNA clones to the public.6,7 We have launched the ROUGE database (http://www.kazusa.or.jp/rouge/) for sharing these lines of information with the research community. We are also convinced of the necessity to prepare a set of antibodies against mKIAA proteins as a proteomic resource for further extension of our project. This is because we consider that specific detection and capture of mKIAA proteins play a key role in the investigation of the physiological functions of the KIAA proteins in vivo. In this context, we started to establish a high-throughput system for the generation of anti-mKIAA antibodies in parallel with the mKIAA cDNA sequencing project. In this paper, we present our system to systematically produce recombinant antigens for mKIAA proteins for the generation of anti-mKIAA antibodies. By using this system, we have so far prepared more than 400 recombinant antigens and have already used them for the immunization of rabbits.

1. Selection of Shotgun Clones Encoding a Region Suitable for Production of Recombinant Antigens for mKIAA Proteins

To select appropriate shotgun clones for antigen production, we built a browser system in which various lines of information necessary for allocation of a region are used for antigen production (Fig. 1). Since antigenicity and solubility of proteins are generally difficult to predict only from their primary structures, multiple characteristics of protein structures should be carefully checked for each mKIAA protein.12–17 This is a time-consuming front-end step in the production of recombinant antigens. In this context, the browser system shown in Fig. 1 was developed to enable us to readily integrate many different lines of information regarding an mKIAA protein structure. In practice, we allocated a region to be used for recombinant antigen production under the following criteria in this project: (1) a C-terminal portion of the predicted mKIAA protein was selected as the antigen region whenever possible because of its surface exposure, negative charge of carboxyl-terminus, and low occurrence of modifications,18 (2) the averaged hydrophilicity of the selected region was higher than 0.05, which was calculated from the entire amino acid residues encoded by shotgun clones by the Hopp and Woods formula, the most commonly used formula for predicting protein hydrophilicity,14 and (3) the region should not be involved in any membrane-spanning region predicted by SOSUI19 and was designed to contain at least 80 amino acid residues. The averaged values of hydrophilicity and length (in terms of the number of amino acid residues) of the actually selected regions were 0.18 and 164, respectively. Once the region to be produced as the antigen was specified, the browser system allowed us to allocate shotgun clones simultaneously. Because we could directly check sequence chromatograms of respective shotgun clones on the same browser, it is straightforward to move to construction of expression plasmids for the recombinant antigen (Fig. 1).

In addition, the browser system also allowed us to check whether or not the selected region was conserved between mouse and human (Fig. 1). The use of this browser system thus offered many lines of information critical for allocation of a region to be used as an antigen...
for each mKIAA protein and greatly expedite selection of a shotgun clone to be used for construction of the expression plasmid of the recombinant antigen.

2. Overall Strategy of Construction of Expression Plasmids for Glutathione S-transferase (GST)-fused mKIAA Proteins

Another time-consuming and labor-intensive process in the production of recombinant antigens is the construction of expression plasmids for recombinant antigens of multiple proteins. In a conventional method, ligation-assisted cloning with a product of polymerase chain reaction (PCR) or a restriction fragment was used for this purpose. However, we did not consider it reasonable to apply this conventional method for high-throughput production of recombinant antigens, since parallel construction of expression plasmids for many recombinant antigens is almost impossible in practice. Thus, we had to develop a new method which enabled us to generate multiple expression plasmids in parallel. Figure 2 shows a flowchart of this construction method of expression plasmids for recombinant antigens of mKIAA proteins. Shotgun clones had been accumulated during the process of entire mKIAA cDNA sequencing as described in our previous reports and harbored 0.8- to 1.0-kb fragments of each mKIAA cDNA in a modified pBluescript II SK vector (STRATAGENE, La Jolla, CA), called pBS2SK-EXPSB, in which the in vitro recombination sites (attB1 and attB2 sites) were placed adjacent to an SmaI site in the multiple cloning sites (Fig. 1). In our project, glutathione S-transferase (GST) was selected as a fusion partner to mKIAA proteins simply because GST is a widely used tag for soluble recombinant antigen production in Escherichia coli. Thus, we constructed a set...
Figure 2. Outline of the strategy to construct GST-mKIAA fusion protein expression plasmid by the in vitro recombination-assisted method. A) 0.8- to 1.0-kb shotgun fragments derived from whole mKIAA plasmid DNA were ligated into the dephosphorylated Smal-digested pBS2SK-EXPSB vector described previously. Since templates for sequencing were obtained by direct colony PCR, the remaining PCR products were recycled for recombination reactions to construct GST-mKIAA fusion protein expression plasmids. To minimize the chance of mis-incorporation during PCR reaction, PCR amplification was performed using LA Taq polymerase (TAKARA, Shiga, Japan) since this is known to have higher fidelity than conventional Taq polymerase, and the number of PCR cycles was kept 20. First, the BP reaction was performed to create an entry clone from the PCR products. Subsequently, the resultant entry clone containing the PCR fragment was used for LR reaction. These reactions were performed in a single-tube format in principle according to the instructions provided by Invitrogen. Destination vector pGEX-4TDES for LR reaction was constructed by inserting an attR1-ccdB-Cmr^r-attR2 fragment (Invitrogen Corp., Carlsbad, CA) into an SmaI site of a pGEX-4T-1 vector (Amersham, Piscataway, NJ). Am^r, Gm^r, and Cm^r are abbreviation for ampicillin-, gentamicin-, and chloramphenicol-resistance, respectively. The ccdB gene encoding a toxin targeting the essential DNA gyrase of E. coli and the phage lambda recombination sites (attB, attP, attL, and attR) are also indicated in this figure. B) Six different reading frames of the destination vector were prepared for adaptation to every shotgun clone. Most recently we also constructed pGEX-6PDES which derived from pGEX-6P-1 (Amersham, Piscataway, NJ) encoding PreScission™ protease recognition sequence, and applied to the latest 27 mKIAA shotgun clones. The GST-fused mKIAA antigens produced from these new constructs enabled us to elute mKIAA antigen from glutathione Sepharose-column by protease cleavage leaving the GST moiety on the column (data not shown). We decided to apply this strategy to the next stage of mKIAA protein purification since the lack of a GST moiety is preferable for the production of antibodies. We are currently preparing the MOUSE database, the details of each GST-mKIAA antigen will become available through our web site in the near future.

of in vitro recombination-compatible vectors which were designed to direct the synthesis of polypeptides encoded by shotgun clones as GST-fusion proteins in E. coli. These vectors, conventionally termed the destination vectors by Invitrogen, were constructed by the insertion of an attR1-ccdB-Cmr^r-attR2 cassette (Invitrogen Corp., Carlsbad, CA) into the Smal I site of a pGEX-4T-1 vector (Amersham, Piscataway, NJ). In practice, 6 destination vectors were prepared so as to accept shotgun clones in any orientation and any reading frame (Fig. 2). In combination with this set of the destination vectors, mKIAA shogun clones were readily recycled for GST-fusion protein production using BP and LR recombination reactions, as shown in Fig. 2. To reduce the number of the rounds of bacterial transformation, we performed these recombination reactions in a single-tube format in principle according to the instructions provided by Invitrogen. This strategy has two major advantages over the conven-
3. Small-scale Production of GST-fused mKIAA Proteins in a 96-well Format

Because the expression plasmids for GST-fused mKIAA proteins were generated in a 96-well plate as described above, they were first introduced into E. coli DH5α cells (Novagen, Madison, WI) also in a 96-well plate, and then recovered with a MAGNA plasmid preparation robot (TOYOBO, Osaka, Japan). For actual production of recombinant antigens, the recovered expression plasmids were then subjected to transformation of E. coli Rosetta (pLysS) cells (Novagen, Madison, WI), a BL21 derivative designated to enhance the expression of eukaryotic proteins containing codons rarely used in E. coli, in a 96-well plate.

Expression of GST-fused mKIAA antigens was tested by culturing E. coli cells in 300 µl of LB medium in a 96-well plate format. The first preliminary expression experiment was done to evaluate the production level of each GST-fused mKIAA antigen. In this experiment, we compared staining patterns on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of all E. coli proteins harboring expression plasmids for their respective GST-fused mKIAA antigens with that of vehicle E. coli proteins under the same culture conditions. If an unusually strong band was seen in the staining pattern, it was taken as an indication that the production level of the GST-fused mKIAA antigens was acceptable. Furthermore, just in case the apparent molecular mass of the protein in the unusually strong band deviated considerably from the predicted one, the produced protein was excluded from further experiments. Subsequently, the second preliminary expression experiment was done to evaluate whether or not the produced GST-fused mKIAA antigens could be subjected to glutathione-based affinity purification. For this purpose, we established an easy and rapid assay, hereafter simply designated as “solubility assay,” which enabled us to assess not only the solubility but also the glutathione-binding capability of the GST-fused mKIAA antigens in a high-throughput manner. A conventional method based on protein blot analysis using anti-GST antibody includes multiple steps and is thus time-consuming. In addition, detection using anti-GST antibody does not necessarily reflect the binding specificity to glutathione. To faithfully check the feasibility of the glutathione-based affinity purification in practice, we devised this solubility assay based on the capture of the GST-fused mKIAA antigens by glutathione immobilized on a solid phase followed by detection of peroxidase-conjugated anti-GST antibody. The experimental details of this assay are described in the legend of Fig. 3. The amounts of the soluble recombinant antigens with glutathione-binding ability were semi-quantitatively expressed as absorbance at 405 nm (A_{405nm}), which were row outputs of the solubility assay.

4. Purification of GST-fused mKIAA Antigens

Based on the results of the solubility assay, we took two different approaches for preparation of GST-fused mKIAA antigens, depending on the amount of available soluble recombinant antigens. When the solubility assay gave an A_{405nm} value larger than 0.25, E. coli Rosetta (pLysS) cells harboring the expression plasmid were cultured in 10 ml of LB medium containing 200 µg/ml ampicillin and 5 µg/ml chloramphenicol for 15 hr at 37°C, and then transferred to 500 ml of the same medium containing 0.1 mM isopropyl β-D-thiogalactoside (IPTG) and cultured for further 15 hr at 18°C. The GST-fused mKIAA antigens extracted in the soluble fraction were purified on glutathione-Sepharose beads as described. Two hundred and twenty out of 400 GST-fused mKIAA antigens processed so far (55%) were produced as soluble proteins and thus purified as described above. The recovered amounts of the GST-fused mKIAA antigens were 0.4 mg to 16 mg (3.0 mg on average) by this method.

On the other hand, when the solubility assay gave an A_{405nm} value smaller than 0.25, E. coli Rosetta (pLysS) cells harboring the expression plasmid were cultured in 4 ml of LB medium containing 200 µg/ml ampicillin and 5 µg/ml chloramphenicol for 15 hr at 37°C, and then transferred to 100 ml of same medium containing 1 mM IPTG and cultured for further 3 hr at 37°C. After removal of soluble proteins as described in the legend of Fig. 3, the insoluble proteins were solubilized with a sample buffer for SDS-PAGE and analyzed by 12% SDS-PAGE. After visualization of proteins by copper negative staining (Copper Stain; Bio-Rad, Hercules, CA), the protein band was excised and then subjected to electro-elution with Model 422 apparatus (Bio-Rad). We allocated protein bands to be excised as described in the first preliminary expression experiment. One hun-
High-throughput Production of GST-fused Mouse KIAA Antigens

A. Selection of shotgun clone appropriate for antigen production

- Transfer of shotgun fragments into expression vector using in vitro recombination reactions in a 96-well plate
- Transformation of E.coli DH5α cells and plasmid purification
- Transformation of E.coli Rosetta (pLysS) cells and confirmation of the expression in a 96-well plate
- Solubility assay using immobilized glutathione in a 96-well plate

B. Soluble Insoluble

- Purification with glutathione beads
- Purification by SDS-PAGE

C. A405

mKIAA numbers

D. mKIAA numbers

Soluble fusion proteins

Insoluble fusion proteins
dried and eighty GST-fused mKIAA antigens were prepared and purified in this way. Using this method, we could recover 0.5 mg to 8.4 mg (1.7 mg on average) of the recombinant antigens. Typical SDS-PAGE patterns of the recovered soluble and insoluble GST fused mKIAA antigens are shown in Fig. 3D.

From the experience of production of more than 400 GST-fused mKIAA antigens, we have realized that some GST-fused mKIAA antigens could not be produced in sufficient amount in E. coli cells in either soluble or insoluble form. In particular, we found that the abundance of serine residues in mKIAA antigen was well correlated with the success rate of production of soluble GST-fused mKIAA antigens in E. coli; if the content of serine residues in mKIAA antigen is higher than 15%, the GST-fused mKIAA antigen is unlikely to be overproduced in E. coli. We also noted that the presence of amino acid regions revealing more than 80% occupancy of either glutamate, leucine, or proline residues was unfavorable for overproduction of GST-fused proteins. Although these were only empirical laws, we considered it safe to select a region without any of these risk factors for the production of GST-fused protein in E. coli.

In this manuscript, we introduced our system for the preparation a large numbers of GST-fused mKIAA antigens utilizing E. coli as an expression host. We already inoculated these 400 GST-fused mKIAA antigens to rabbits, purified the rabbit anti-mKIAA antibodies, and are now evaluating the affinity and specificity of the antibodies. In addition, we are currently preparing the MOUSE database to share the information regarding mKIAA proteins and their recombinant antigens with the research community.

Acknowledgements: We thank Yukiko Kamatani, Tomomi Tajino, Tomomi Kato, and Kiyoe Sumi for their technical assistance. We also thank Dr. Takagi (Protein Express) for valuable support to Y. Hara, and Drs. Nagase and Okazaki (Kazusa DNA Research Institute) for their useful suggestions and continuous encouragement. This study was supported by the CREATE Program (Collaboration of Regional Entities for the Advancement of Technological Excellence) from JST Express.}

Figure 3. Flow diagram and representative results of the expression and the purification of the GST-mKIAA fusion proteins. A) Flow diagram of the steps for expression and purification of the GST-mKIAA fusion proteins. Nearly 96 selected shotgun fragments were simultaneously transferred into an expression vector using BP and LR reactions in a 96-well plate, followed by transformation into E. coli DH5αE. coli DH5α cells (Novagen, Madison, WI). After single colonies were picked up, their insert sizes were compared to the original PCR products on agarose gel electrophoresis to eliminate the risk of accidental selection of wrong clones during these steps. These PCR products were also used for sequence analysis to confirm their sequence integrity. Robotically isolated plasmid DNAs were then used for transformation into E. coli Rosetta (pLysS) (Novagen, Madison, WI) in the same 96-well format. Expression of the fusion proteins was tested in the 96-well format as described in (B). The solubility of the fusion proteins were also tested by solubility assay in the same format as described in (C). GST-mKIAA fusion proteins were differently purified according to the results of the solubility assay. B) A typical result of first preliminary expression experiment. E. coli Rosetta (pLysS) cells harboring the expression plasmid were cultured in 300 µl of LB medium containing 200 µg/ml ampicillin and 5 µg/ml chloramphenicol at 16 hr at 37°C, subsequently diluted to one-tenth in 300 µl of the same medium containing 0.2 mM isopropyl β-D-thiogalactoside (IPTG) and cultured for further 1.5 hr at 37°C. Following centrifugation, cell pellets were resuspended in 50 µl of TBS (25 mM Tris-HCl, pH 8.0, 0.15 M NaCl), then added 25 µl of 3 × SDS sample buffer [150 mM Tris-HCl, pH 6.8, 6% SDS, 30% glycerol, 3% β-mercaptoethanol (β-ME), and 0.06% BPB]. One-tenth of the resuspended cell pellets (7.5 µl) was resolved on 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and the gels were then stained with Coomassie Brilliant Blue (R-250; Wako Pure Chemicals, Osaka, Japan). The positions of the molecular mass markers (in kDa) are indicated on the right of the each panel. Corresponding mKIAA numbers are indicated on the top of the each lane. C) Representative results of the solubility assay. E. coli Rosetta (pLysS) cells harboring the expression plasmid were cultured in 300 µl of LB medium containing 200 µg/ml ampicillin and 5 µg/ml chloramphenicol for 16 hr at 30°C, subsequently diluted to one-tenth in 300 µl of the same medium containing 0.1 mM IPTG and cultured for an additional 1.5 hr at 25°C. Following centrifugation, cell pellets were lysed in 100 µl Cellytic B H (Sigma, St. Louis, MO) and soluble proteins were extracted from the cell pellets. After dilution with TBS, one-hundredth of the diluted soluble proteins were then captured by immobilized glutathione on a 96-well plate (EXIQON, Vedbaek, Denmark). The captured soluble proteins were detected by peroxidase conjugated anti-GST antibody (Amersham, Piscataway, NJ), followed by color development (blue-green) using 2,2′-azinobis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS; KPL, Inc., Gaithersburg, MD). A_{405nm} was determined with a microplate reader after a 6-min peroxidase reaction. A_{405nm} is indicated on the left. The corresponding mKIAA numbers are indicated on the bottom of the each lane. A_{405nm} = 0.25, limited threshold of the solubility, is indicated by a dash line. D) Representative results of the purified proteins. After judgment by the solubility assay, E. coli Rosetta (pLysS) cells revealed A_{405nm} ≥ 0.25 were cultured in 10 ml of LB medium containing 200 µg/ml ampicillin and 5 µg/ml chloramphenicol for 15 hr at 37°C, subsequently transferred to 500 ml of the same medium containing 0.1 mM IPTG and cultured for an additional 15 hr at 18°C. Following sonication of the cell pellet in TBST-E-β buffer containing 0.02% Tween 20, 1 mM EDTA, and 0.2% β-ME, GST fusion protein was captured to 1 ml of glutathione sepharose beads (50% slurry) (Amersham, Piscataway, NJ) then eluted by 6 ml of glutathione buffer containing 10 mM glutathione, 0.1 M Tris (pH 8.0), and 0.2 M NaCl. On the other hand, E. coli Rosetta (pLysS) cells where A_{405nm} < 0.25 were cultured in 4 ml of LB medium containing 200 µg/ml ampicillin and 5 µg/ml chloramphenicol for 15 hr at 37°C, subsequently transferred to 100 ml of the same medium containing 1 mM IPTG and cultured for an additional 3 hr at 37°C. Following sonication of the cell pellet, the cell pellet was extensively washed with modified TBST-E-β buffer containing 4% TX-100. Residual pellet was resuspended in 1 × SDS sample buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-ME, and 0.02% BPB) and electrophoresed on 12% SDS-PAGE. After negative staining (Copper Stain; Bio-Rad, Hercules, CA), the protein band was excised then the protein was eluted by electrolutor (Model 422; Bio-Rad, Hercules, CA). Small amounts of purified proteins were resolved on 12% SDS-PAGE to confirm purity. Molecular mass markers (in kDa) are indicated on the left of the panel. The corresponding mKIAA numbers are indicated on the top of the each lane.
(Japan Science and Technology Corporation) and a grant from the Kazusa DNA Research Institute.

References

1. Kawai, J., Shinagawa, A., Shibata, K. et al. 2001, Functional annotation of a full-length mouse cDNA collection, Nature, 409, 685–690.

2. Kargul, G. J., Dudekula, D. B., Qian, Y. et al. 2001, Verification and initial annotation of the NIA mouse 15K cDNA clone set, Nat. Genet., 28, 17–18.

3. Strausberg, R. L., Feingold, E. A., Grouse, L. et al. 2002, Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences, Proc. Natl. Acad. Sci. USA, 99, 16899–16903.

4. Nomura, N., Miyajima, N., Sazuka, T. et al. 1994, Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001–KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1 (supplement), DNA Res., 1, 47–56.

5. Ohara, O., Nagase, T., Mitsui, G. et al. 2002, Characterization of size-fractionated cDNA libraries generated by the in vitro recombination-assisted method, DNA Res., 9, 47–57.

6. Okazaki, N., Kikuno, R., Ohara, R. et al. 2002, Prediction of the coding sequences of mouse homologues of KIAA gene: I. The complete nucleotide sequences of 100 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries, DNA Res., 9, 179–188.

7. Okazaki, N., Kikuno, R., Ohara, R. et al. 2003, Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries, DNA Res., 10, 35–48.

8. Waterston, R. H., Lindblad-Toh, K., Birney, E. et al. 2002, Initial sequencing and comparative analysis of the mouse genome, Nature, 420, 520–562.

9. Nolan, P. M., Peters, J., Strivens, M. et al. 2000, A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse, Nat. Genet., 25, 440–443.

10. Hrabe de Angelis, M., and Strivens, M. 2001, Large-scale production of mouse phenotypes: the search for animal models for inherited diseases in humans, Brief Bioinform, 2, 170–180.

11. Boyd, Y., Blair, H. J., Cunliffe, P., Masson, W. K., and Reed, V. 2000, A phenotype map of the mouse X chromosome: models for human X-linked disease, Genome Res., 10, 277–292.

12. Chou, P. Y. and Fasman, G. D. 1978, Empirical predictions of protein conformation, Annu. Rev. Biochem., 47, 251–276.

13. Garnier, J., Osguthorpe, D. J., and Robson, B. 1978, Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins, J. Mol. Biol., 120, 97–120.

14. Hopp, T. P. and Woods, K. R. 1981, Prediction of protein antigenic determinants from amino acid sequences, Proc. Natl. Acad. Sci. USA, 78, 3824–3828.

15. Emin, E. A., Hughes, J. V., Perlow, D. S., and Boger, J. 1985, Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide, J. Virol., 55, 836–839.

16. Karplus, P. A. and Schulz, G. E. 1985, Prediction of chain flexibility in proteins: a tool for the selection of peptide antigen, Naturwissenschaften, 72, 212–213.

17. Jameson, B. A. and Wolf, H. 1988, The antigenic index: a novel algorithm for predicting antigenic determinants, Comput. Appl. Biosci., 4, 181–186.

18. Harlow, E. D. and Lane, D. 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

19. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. 1998, SOSUI: classification and secondary structure prediction system for membrane proteins, Bioinformatics, 14, 378–379.

20. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

21. Smith, D. B. and Johnson, K. S. 1988, Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase, Gene, 67, 31–40.

22. Coligan, J. E., Dunn, B. M., Speicher, D. W., and Wingfield, P. J., 2002, Current Protocols in Protein, John Wiley & Sons, Inc. New Jersey, pp. 5.17.1–5.17.10.

23. Braun, P., Hu, Y., Shen, B. et al. 2002, Proteome-scale purification of human proteins from bacteria, Proc. Natl. Acad. Sci. USA, 99, 2654–2659.

24. Murray, A. M., Kelly, C. D., Nussey, S. S., and Johnstone, A. P. 1998, Production of glutathione-coated microtitre plates for capturing recombinant glutathione S-transferase fusion proteins as antigens in immunoassays, J. Immunol. Methods, 218, 133–139.

25. Rabin, D. U., Palmer-Crocker, R., Mierz, D. V., and Yeung, K. K. 1992, An ELISA sandwich capture assay for recombinant fusion proteins containing glutathione-S-transferase, J. Immunol. Methods, 156, 101–105.

26. Sehr, P., Zumbach, K., and Pawlita, M. 2001, A generic programme for gene function studies in the mouse, Genom. Inform., 2, 153–162.

27. Coligan, J. E., Dunn, B. M., Speicher, D. W., and Wingfield, P. J. 1997, Current Protocols in Protein, John Wiley & Sons, Inc. New Jersey, pp. 6.6.1–6.6.21.