Article

Construction of cDNA Library from NPC Tissue and Screening of Antigenic Genes

Jun Shu[1, 2], Xiaojuan He[1] and Guancheng Li[1, 3]

To construct cDNA library of nasopharyngeal carcinoma (NPC) and obtain the NPC associated or specific antigens from it, we used a powerful new method to identify the antigens eliciting humoral immune response, which is SEREX (serological identification of antigen by recombinant cDNA expression library). Autologous serum of NPC patient was used to screen the reactive clones in the human NPC tissue cDNA library consisted of 3.64×10^6 recombinants. The 23 exact positive clones were subcloned to monoclonality and the size of cDNA inserts was identified by PCR. Then the nucleotide sequence of cDNA inserts was determined, and the sequence alignments were performed with BLAST software on GenBank database. They represented 16 different antigens. A detailed sequence analysis showed that 10 of 16 genes were high homologous to genes known in GenBank, such as RPL31, S100 A2, MT2A, etc. However, there were also 6 genes with low homology to genes in GenBank. Furthermore, 3 of 6 genes may be novel genes. The associations of these genes to NPC and the roles that they played in the occurrence and development of NPC should be further revealed. *Cellular & Molecular Immunology*. 2006;3(1):53-57.

Key Words: NPC, cDNA library, SEREX, antigenic gene

Introduction

A prerequisite for the successful application of recombinant tumor vaccines and other immunotherapeutic interventions in cancer patients is the recognition by the immune system of tumor-specific and tumor-associated antigens (i.e., of molecules that are overexpressed or specifically expressed in the tumor cells or tissues). Serological analysis of recombinant cDNA expression library (SEREX) is an approach for the identification of antigenic proteins in cancer, and was first used by Sahin et al. in 1995 (1). Since then, it has been applied to a variety of human tumor types resulting in the detection of more than 2,000 potential serologically active tumor antigens (2). However, up to now there was not a report about screening and identification of NPC antigens with immunological method. As reported previously, monoclonal antibodies against NPC cell line CNE1 and HNE2 were generated respectively, which had high specification to NPC (3, 4). These results proved the existence of NPC antigen, and also allowed identifing the antigenic proteins. In present research some antigenic genes of NPC were obtained by screening cDNA library derived from NPC tissue. Some of these antigenic genes might be provided as new molecules for immunological research about NPC.

Materials and Methods

Tissue specimens and patient sera

NPC tissue specimen was resected from a patient undergoing no therapy at Hunan Tumor Hospital and frozen in liquid nitrogen immediately. In addition, serum sample was obtained from the same patient bearing cancer and stored at -20°C.

Isolation and purification of total RNA

Total RNA was isolated from tumor tissue sample, using Trizol according to the manufacturer's instructions (Invitrogen). Then the total RNA was digested by DNase-I to clear out remnant DNA and its concentration and quantity were determined by detection of OD value and electrophoresis.

Construction of the cDNA library

A cDNA expression library was constructed from a tumor specimen of a low differentiated nasopharyngeal carcinoma, using SMART™cDNA library construction kit according to

[1]Cancer Research Institute, XiangYa School of Medicine, Central South University, Changsha 410078, China;

[2]Clinic Medical Institute, China-Japan Friendship Hospital, Beijing 100029, China;

[3]Corresponding to: Dr. Guancheng Li, Cancer Research Institute, XiangYa School of Medicine, Central South University, Changsha 410078, China. Tel: +86-731-480-5445, Fax:+86-731-235-5042, E-mail: libsun@public.cs.hn.cn.

Received Nov 22, 2005. Accepted Jan 20, 2006.

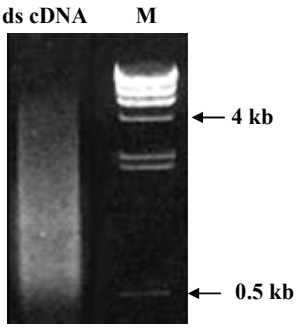

Figure 1. Ds-cDNA of NPC tissue was analyzed by electrophoresis on 1% agarose gel. With the SMART technology, a modified oligo (dT) primer was used to prime the first-strand synthesis reaction, and the SMART IV oligo was served as a short, extended template at the 5' end of the mRNA. cDNA synthesis employed LD-PCR for generating full-length cDNA. Electrophoresis showed that this dsDNA ranged from 0.3 kb to 4.0 kb. M, λ hind III marker.

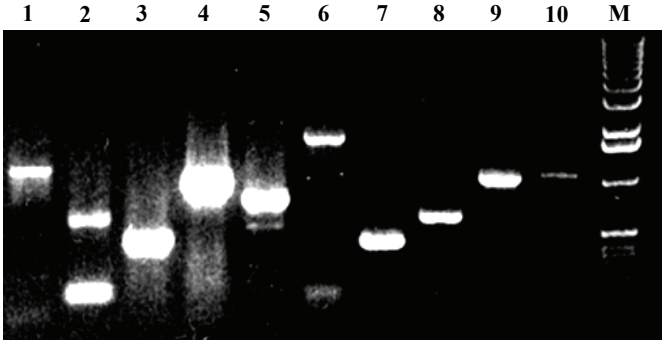

Figure 2. PCR products of random clones from cDNA amplified library were analyzed by electrophoresis on 1% agarose gel. After construction of cDNA library, ten monoclonal recombinant phages were picked up randomly and PCR was performed to verify the size of cDNA inserts in the phages. PCR condition was: 95°C, 5 min; 95°C, 30 sec; 65°C, 4 min; 35 cycles; 65°C, 8 min. The PCR products were electrophoresed in 1% agarose gel, showing that ranged from 0.5 kb to 3.0 kb. Lanes 1-10, PCR products of random clones; M, 1 kb DNA ladder.

the manufacturer's instructions (Clontech). With the SMART (switching mechanism at 5' end of RNA transcript) technology, a modified oligo (dT) primer was used to prime the first-strand synthesis reaction, and the SMART IV oligo was served as a short, extended template at the 5' end of the mRNA. cDNA synthesis employs long-distance PCR (LD-PCR) for generating full-length cDNA. cDNA was digested with *Sfi I*, then the digested cDNA fragments were cloned directionally into the bacteriophage expression vector λ TriplEx2 (*Sfi I* A&B digested) and packaged into phage particles using Packagene Lambda DNA Packaging System (Promega) to construct the primary cDNA library. Phages were used to transfect *XL1-Blue* to amplify the primary cDNA library.

Amplification of primary cDNA library

To increase the stability of cDNA library, it was amplified as described in the manufacturer's instructions. The amplified library can be stored at 4°C for 6-7 months or at -70°C (in 7% DMSO) for at least one year. Ten monoclonal recombinant phages were picked up randomly and PCR was performed to verify the size of cDNA inserts in the phages.

Autologous immunoscreening of transfectants

Immunoscreening for the detection of reactive clones was performed by taking "one versus one" means with autologous serum as previously (1, 5, 6). Briefly, *Escherichia coli* XL1-Blue were transfected with the recombinant phages, and then poured onto the LB/$MgSO_4$ agar plates. Plates were incubated at 37°C until plaques were visible. Next, expression of recombinant proteins was induced by nitro-cellulose membranes that had been soaked in isopropyl-β-D-thiogalactoside (IPTG) for 10 minutes. After 4 hours, the membranes were blocked with 5% (w/v) low-fat milk in Tris-buffered saline and incubated with 1:100 diluted patient's serum, which had been preabsorbed with transfected *E. coli* lysate. Serum antibodies binding to recombinant proteins expressed in lytic plaques were detected by incubation with a peroxidase-conjugated goat anti-human IgG and visualized by staining with 3, 3'-diaminobenzidine (DAB). The reactive phage clones were immunoscreened secondly and thirdly in order to exclude false positive clones and subclone to monoclonality.

PCR of cDNA insert in positive clones

Positive clones were performed LD-PCR to amplify the cDNA inserts with λ TriplEx LD-insert screening primers, using Advantage 2 PCR kit (Clontech) according to the manufacturer's instructions.

Sequence analysis of identified antigens

The amplified cDNA products were purified and the nucleotide sequence of cDNA inserts was determined by Bio-corporation. The sequence alignments were performed with BLAST software on GenBank database.

Results

Construction of cDNA library from NPC tissue

Double-strand cDNA was generated by LD-PCR and electrophoresed on 1% agarose gel (Figure 1). After construction of cDNA library, the titer was determined by using phages to transfect *XL1-Blue*. The result showed that the titer was 7.28×10^6 pfu/ml and the library consisted of 3.64×10^6 recombinants; the recombinant rate was 94%. For better preserving, the cDNA library was amplified, resulting at least 3.8×10^9 pfu/ml. Ten monoclonal recombinant phages were picked up randomly and PCR was performed to verify the size of cDNA inserts in the phages. The PCR products were electrophoresed on 1% agarose gel, and

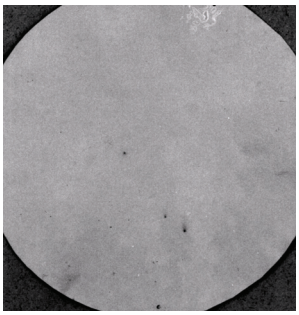

Figure 3. Positive dots of phage clone from first cycle of screening by sera. Immunoscreening was performed by taking "one versus one" means with autologous serum as previously. In first immunoscreening, of a total of 1×10^6 clones tested, 132 were found to be reactive with IgG antibodies in the serum sample from the autologous patient at a dilution of 1:100.

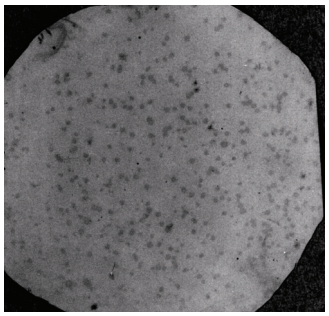

Figure 4. **Positive dots of phage clone from second cycle of screening by sera.** The second and third cycle immunoscreening were to exclude false positive clones and subclone to monoclonality. The method performed was the same as the first cycle immunoscreening. After second and third screening, 23 of 132 clones above were obtained, which were real positive to the autologous serum.

ranged from 0.5 kb to 3.0 kb (Figure 2).

Immunoscreening of cDNA library

The cDNA library was expressed in *XL1-Blue* using a λ TriplEx2 system. In first immunoscreening, of a total of 1×10^6 clones tested, 132 were found to be reactive with IgG antibodies in the serum sample of the autologous patient at a dilution of 1:100 (Figure 3). After second and third screening, 23 of 132 clones above were obtained, which were real positive to the autologous serum (Figure 4).

Sequencing and identification of positive clones

The inserts of positive clones were subcloned to monoclonality, and the nucleotide sequences of the cDNA inserts were determined. Results revealed that the 23 positive clones contained 16 different inserts; thus, some inserts might be expressed by multiple clones in the cDNA library, and these cDNA inserts were named NPC-A-1~16 (Table 1), which might be potential tumor antigenic genes. The cDNAs were identified by homology search through the GenBank database. It showed that 10 of 16 cDNAs were high homologous to known genes in GenBank respectively, 3 additional cDNAs with low homology to known genes, and 3 other cDNAs were hitherto unknown molecules. The sequences of these latter three molecules (NPC-A-5, NPC-A-6, NPC-A-7) have been deposited in the GenBank database (Accession Nos. AY320401, AY320414, AY320402) (Table 2).

Discussion

Construction of cDNA expression library is an important molecular biological technique. Through sequencing clones of cDNA library, researchers can not only identify some known genes but also obtain some novel genes. The quality of cDNA library consists of two aspects: the capacity and the sequence integrality of recombinant cDNA. As concerning the capacity, the constructed cDNA library that is screened by serum or antibody must contain at least 2×10^6 recombinants according to the report previously. In our research, SMART technique was used to construct cDNA library of human nasopharyngeal carcinoma tissue, resulting in at least 3.64×10^6 recombinants in primary library, which completely meets the need of cDNA library capacity.

The integrality of recombinant cDNA segment is another important factor that reflects the quality of cDNA library. With SMART technique, cDNA was synthesized by LD-PCR. This procedure has some merits as follows: 1) The percentage of full-length clones in the library can be increased. 2) As the superiority of LD-PCR in cDNA synthesis, the amount of available RNA starting material needed is very small (i.e., 50 ng of total RNA or 25 ng of

Table 1. Size of the inserted cDNA in the positive clones

Clones No.	Size of inserted DNA (bp)
NPC-A-1	384
NPC-A-2	523
NPC-A-3	553
NPC-A-4	462
NPC-A-5	383
NPC-A-6	620
NPC-A-7	372
NPC-A-8	344
NPC-A-9	150
NPC-A-10	600
NPC-A-11	406
NPC-A-12	1465
NPC-A-13	589
NPC-A-14	524
NPC-A-15	415
NPC-A-16	573

Table 2. Genes identified by SEREX analysis of human NPC tissue cDNA library

Name	Identity	GenBank No.	Chr.location	Product, function, feature, expression pattern
NPC-A-1	RPL31	NM_000993	2q11.2	Ribosomal proteinL31, a component of the 60S subunit. Higher levels expressed in familial adenomatous polyps.
NPC-A-2	S100 A2	NM_005978	1q21	S100 calcium-binding protein A2, a member of the S100 family, ubiquitously expressed, involved in cell cycle progression and differentiation, have a tumor suppressor function.
NPC-A-3	H2AFZ	NM_002106	4q24	H2A histone family, member Z.A basic nuclear proteins that are responsible for the nucleosome structure of the chromosomal fiber in eukaryotes,is required for embryonic development.
NPC-A-4	MT2A	BC007034	16q13	Metallothionein 2A,ubiquitously expressed in multiple cancers.
NPC-A-5		AY320401		Unknown
NPC-A-6		AY320414		Unknown
NPC-A-7		AY320402		Unknown
NPC-A-8	KIAA1630	NM_018706	10p14	KIAA1630 protein, lies within an 82 kb intron of the regulator of nonsense transcripts 2 (RENT2) gene and is transcribed in the opposite direction from it. The protein has some similarity to dehydrogenase and transketolase protein domains.
NPC-A-9	RPS5	NM_001009	19q13.4	Ribosomal protein S5, belongs to the S7P family of ribosomal proteins. It is located in the cytoplasm. This gene expresses variably in colorectal cancers
NPC-A-10	KIAA0672	NM_014859	17p12	KIAA0672 gene product, function unknown.
NPC-A-11	CTB-95A14	AC005081	chr.7	Unknown
NPC-A-12	HS333H23	AL022326	22q12.1-12.3	Unknown
NPC-A-13	HSBP1	NM_001537	16q23.3	Heat shock factor binding protein 1, overexpression of HSBP1 in mammalian cells, represses the transactivation activity of HSF1, as a negative regulator of heat shock response.
NPC-A-14	RPL6	NM_000970	12q24.1	Ribosomal protein L6, belongs to the L6E family of ribosomal proteins. The protein can bind specifically to domain C of the tax-responsive enhancer element of human T-cell leukemia virus type1, and it has been suggested that the protein may participate in tax-mediated transactivation of transcription.
NPC-A-15	NCL	NM_005381	2q12-qter	Nucleolin, a eukaryotic nucleolar phosphoprotein, is involved in the synthesis and maturation of ribosomes.
NPC-A-16	LOC254948	NG_002468	15q25.1	Ribosomal protein L9 pseudogene, function unpublished.

mRNA), which is significative to those limited by difficult gaining samples, such as NPC tissue. 3) Double-strand cDNA digested by *Sfi I* are ligated to the arms of λ TriplEx2 vector, consequently achieve the directional cloning of cDNA simply and conveniently.

Along with the development of molecular biological technology, especially the cDNA library technique having been used, Pfreundschuh's group developed the SEREX technique firstly in 1995 (1), which inaugurates a new era of identifying antigen with serum. Since then, it has been applied to a variety of human tumor types resulting in the detection of more than 2,000 potential serologically active tumor antigens (2). However, up to now there is not any report about SEREX technique being used to NPC.

By implementing molecular cloning techniques into the original strategy of autologous typing, SEREX allows an unbiased search for an antibody response and the direct molecular definition of immunogenic tumor proteins based on their reactivity with autologous patient sera, which is technically characterized by several features: 1) The use of fresh tumor specimens to produce the cDNA libraries obviates the need to culture tumor cells *in vitro* and therefore circumvents artifacts, such as loss or neoexpression of antigens, generated by extended culture. The analysis is therefore restricted to antigen-encoding genes that are expressed by the tumor *in vivo*. 2) By using patient sera at dilution of 1:100 to 1:1,000, only high-titer IgG antibodies are detected in the initial screening procedure; this limits the analysis to those antigens that elicit a strong immune response in the autologous host and therefore imply cognate T-cell help. 3) Using cDNA expression cloning, the serological analysis is not restricted to cell-surface antigens, but covers a more complete repertoire of proteins coded by the genes of the tumor. 4) In contrast to techniques using monoclonal antibodies, SEREX uses polyspecific sera to scrutinize monoclonal antigens that are highly enriched in

lytic bacterial plaques. This allows for a direct molecular definition of antigens, because the cDNA of the tumor antigen can be sequenced directly.

It is very important how to select serum to screen cDNA library with SEREX. Many researchers use mixed sera from several patients with the same tumor, which can increase the possibility of obtaining positive clones, but raise the false positive percentage at the same time. Therefore, the serum from the same patient whose diagnostic biopsy was used to construct cDNA library was selected in our research. With this serum, immunoscreening was performed in "one versus one" means, which obviated the cross-reaction in homo-allogenous and might increase the specificity of antibody identifying in a certain extent.

In our research, 10 of 16 cDNA inserts identified were high homologous to known genes in GenBank such as RPL31 (NPC-A-1), S100A2 (NPC-A-2), MT2A (NPC-A-4), etc., and they represented known molecules for which immuno- genicity in humans has not been described. Ribosomal protein L31 (RPL31) overexpresses in colorectal tumor, is associated with a malfunction of normal growth regulatory mechanisms in the tumor, and plays a role in proliferation and neoplasia (7). S100 calcium-binding protein A2 (S100A2), a member of the S100 family, ubiquitously expresses and involves in cell cycle progression and differentiation with a tumor suppressive function (8). Metallothionein 2A (MT2A), ubiquitously overexpresses in multiple cancers and plays a role in proliferation and apoptosis of tumor cells (9, 10). According to these reports at present, these genes identified in the cDNA library of NPC tissue are closed to the neoplasia, development and prognosis of tumor. In addition, three unknown cDNAs have been submitted to the GenBank database (Accession Nos. AY320401, AY320414, AY320402). Because all these genes are firstly found and identified in NPC tissue, their association to NPC and their function in neoplasia are still unknown, so they are worthy to be investigated further.

References

1. Sahin U, Tureci O, Stenner F, et al. Human neoplasms elicit multiple specific immune responses in the autologous host. Proc Natl Acad Sci U S A. 1995;92:11810-11813.
2. Preuss KD, Zwick C, Pfreundschuh M, et al. Analysis of the B-cell repertoire against antigens expressed by human neoplasms. Immunol Rev. 2002;188:43-50.
3. Sun QB, Guo M, Li XL, et al. Monoclonal antibody against nasopharyngeal carcinoma: establishment of hybrid cell lines and preliminary characterization. Hu'nan Yi Ke Da Xue Xue Bao. 1987;12:1.
4. Li GC, Zhu JG, Sun QB, et al. Preparation of monoclonal antibodies against nasopharyngeal cancer by *in vivo* and *vitro* combining immunization. Hu'nan Yi Xue. 1999;1:326.
5. Gure AO, Altorki NK, Chen YT, et al. Human lung cancer antigens recognized by autologous antibodies: Definition of a novel cDNA derived from the tumor suppressor gene locus on chromosome 3p21.3. Cancer Res. 1998;58:1034-1041.
6. Chen YT, Gure AO, Old LJ, et al. Identification of multiple cancer/testis antigens by allogenic antibody screening of a melanoma cell line library. Proc Natl Acad Sci USA. 1998;95: 6919-6923.
7. Chester KA, Robson L, Begent RH, et al. Identification of a human ribosomal protein mRNA with increased expression in colorectal tumours. Biochim Biophys Acta. 1989;1009:297-300.
8. Kyriazanos ID, Tachibana M, Dhar DK, et al. Expression and prognostic significance of S100A2 protein in squamous cell carcinoma of the esophagus. Oncol Rep. 2002;9:503-510.
9. Jin R, Chow VT, Tan PH, et al. Metallothionein 2A expression is associated with cell proliferation in breast cancer. Carcinogenesis. 2002;23:81-86.
10. Cui Y, Wang J, Zhang X, et al. ECRG2, a novel candidate of tumor suppressor gene in the esophageal carcinoma, interacts directly with metallothionein 2A and links to apoptosis. Biochem Biophys Res Commun. 2003;302:904-915.