Nucleic Acids Research, 2005, Vol. 33, No. 4 ***e42***
doi:10.1093/nar/gni042

A multipurpose vector system for the screening of libraries in bacteria, insect and mammalian cells and expression *in vivo*

Olli H. Laitinen[1], Kari J. Airenne[1], Vesa P. Hytönen[3], Erik Peltomaa[1], Anssi J. Mähönen[1], Thomas Wirth[1], Miia M. Lind[1], Kari A. Mäkelä[1], Pyry I. Toivanen[1], Diana Schenkwein[1], Tommi Heikura[1], Henri R. Nordlund[3], Markku S. Kulomaa[3] and Seppo Ylä-Herttuala[1,2,*]

[1]Department of Molecular Medicine and Biotechnology, A.I. Virtanen Institute and [2]Department of Medicine and Gene Therapy Unit, University of Kuopio, PO Box 1627 FIN-70211, Kuopio, Finland and [3]Department of Biological and Environmental Science, NanoScience Center University of Jyväskylä, Jyväskylä, Finland

Received December 22, 2004; Revised and Accepted February 8, 2005

ABSTRACT

We have constructed a novel tetra-promoter vector (pBVboostFG) system that enables screening of gene/cDNA libraries for functional genomic studies. The vector enables an all-in-one strategy for gene expression in mammalian, bacterial and insect cells and is also suitable for direct use *in vivo*. Virus preparation is based on an improved mini Tn7 transpositional system allowing easy and fast production of recombinant baculoviruses with high diversity and negligible background. Cloning of the desired DNA fragments or libraries is based on the recombination system of bacteriophage lambda. As an example of the utility of the vector, genes or cDNAs of 18 different proteins were cloned into pBVboostFG and expressed in different hosts. As a proof-of-principle of using the vector for library screening, a chromophoric Thr^{65}-Tyr-Gly^{67}-stretch of enhanced green fluorescent protein was destroyed and subsequently restored by novel PCR strategy and library screening. The pBVboostFG enables screening of genome-wide libraries, thus making it an efficient new platform technology for functional genomics.

INTRODUCTION

Currently, genomes of over 200 organisms have been sequenced and the work for a number of other species is in progress (http://www.tigr.org, http://www.genomesonline.org/). The Human Genome Project, in particular, is revolutionizing medicine and related biological sciences. Until recently, the focus in genome projects has been on the identification and sequencing of genes. The bare sequence information, however, is not enough to completely define the functions of the tens of thousands of proteins encoded by the revealed open reading frames. This problem has created a new area in molecular biology and genetics called functional genomics, which addresses the functions of genes and gene products at the level of whole organisms (1).

Conventionally, the most widely used approaches to link a gene with a function are mutation analysis and production of knock-out or transgenic animals. These are, however, slow and costly approaches at the genome scale. DNA (2) and protein (3) microarrays and proteomics (4) are evolving techniques used in functional genomics, although these techniques suffer from their inefficiency to detect weakly expressed gene products. Thus, a technology is needed that directly associates the whole genome libraries, such as arrays of isolated cloned cDNAs (5), with the possibility of high-throughput screening for phenotypic changes.

Viral vector systems, in which genes are cloned under strong promoters, have been explored as one solution to these problems. Arrayed libraries of thousands of genes can be constructed within suitable viral vectors, from which gene functions can be identified via cell-based assays using high-throughput screening technology (6,7). Considerable time and several cloning steps, however, are usually required for the construction of adenoviral and retroviral expression cassettes and libraries. In addition, cytotoxicity associated with adenoviruses and alphaviruses may affect the obtained results (7).

To address these issues we have developed a new vector system (pBVboostFG) for library construction which enables expression and screening of libraries in bacterial, insect and mammalian cells. The cloning of the libraries into the pBVboostFG vector is based on the efficient site-specific

*To whom correspondence should be addressed. Tel: +358 17 162075; Fax: +358 17 163751; Email: Seppo.Ylaherttuala@uku.fi

recombination system of bacteriophage lambda (8). The delivery of the cloned genes can also be done directly *in vivo* without any further subcloning steps via baculovirus-mediated transduction (9). In contrast to adenovirus and retrovirus systems, the benefits obtained by using the baculovirus are that at least 50 kb fragments of the DNA can be incorporated into its genome (10) and no cytotoxity is associated with baculoviral vectors in vertebrate cells (11). Moreover, the production of baculoviral libraries by pBVboostFG system is fast, compatible with high-throughput technology, and produces a high diversity of clones ($\sim 10^7$ c.f.u./μg) in a single step without background (12). Here, we describe the construction of pBVboostFG system and show examples of its usefulness in areas relevant to functional genomics.

MATERIALS AND METHODS

Vector construction

In order to allow recombinational cloning we first inserted the recombination cassette of bacteriophage lambda (8) (Gateway, Invitrogen) into a vector backbone that contained promoters functional in bacteria, insect and mammalian cells. The expression cassette was constructed so that as a rule, the first available ATG codon starts translation. Therefore, all that is required is to clone an open reading frame or a suitable library according to bacteriophage lambda's cloning scheme (8). The constructed cassette was cloned into the pBVboost vector (12) and the resultant vector was named as pBVboostFG (Figure 1A). In practice, a linker sequence was first cloned into an NcoI/BamHI digested pTriEx-1.1 vector (EMD Biosciences). The linker was constructed by annealing the following oligonucleotides: F526 (5′-GATCGATATCGGATC-3′) and F527 (5′-CACGGATCCGATATC). The underlined oligonucleotides correspond to the EcoRV site. This linker was also designed to omit the inherent start codon (ATG) of NcoI in order to allow the precise initiation of the protein translation of desired inserts after the recombinational cloning. The resultant plasmid was named as pBVboostFGpre1 and the blunt-end gateway cassette A was cloned into the introduced EcoRV site. This vector was named as pBVboostFGpre2. A second linker prepared by annealing FgL35′ (5′-CCTAGGCATGCCCGG-3′) and FgL33′ (5′-GCATGCCTAGGCCGG-3′) oligonucleotides was then cloned into the FseI cut pBVboostFGpre2. The underlining indicates SphI sites in the oligonucleotides. The resultant vector was named as pBVboostFGpreOB. The pBVboostFGpreOB was then digested with SphI and the recombinational cloning cassette under the triple promoter was ligated into a similarly treated pBVboost vector (12). The resultant vector was named as pBVboostFG (Figure 1A).

Red fluorescent protein was cloned under the polyhedron promoter (pPolh) by subcloning a SacI and XbaI insert from pDsRed2-N1 vector (Clonetech) into a SpeI and SacI cut pBVboostFG.

The chimeric construct containing the bacterial ompA secretion signal fused to avidin cDNA (13) flanked with *att*L1 (5′) and *att*L2 (3′) sites (8) was obtained using SES-PCR (14) (Figure 1). Briefly: in SES-PCR (stepwise elongation of sequence) the amplified insert was elongated (60–70 nt) from its 5′ end using primers that had desired extra sequences in their 5′ end. The procedure was repeated stepwise in order to achieve longer constructs (see Figure 1). The avidin SES-PCR product after three cycles was cloned into pBVboostFG. The enhanced green fluorescent protein (EGFP)-construct was prepared similarly and cloned into pBVboostFG and pBVboostFGR. Other constructs were created in a similar manner. However, those constructs that contained N-terminal His-tag were cloned by adding the tag into the oligonucleotides that contained gene-specific sequences and *att*B sites and the fragments were cloned by subsequent BP and LR reactions (8). The used constructs are summarized in Table 1.

Expression of test genes in bacterial, insect and mammalian cells

Bacterial expressions were carried out in *Escherichia coli* BL21. For the expression of ompA–avidin protein, the production was switched on by adding IPTG (1 mM). After overnight incubation, the cells were fractionated into total, periplasmic and insoluble fractions, subjected to SDS–PAGE and transferred onto nylon beat filters. The proteins were detected using polyclonal rabbit anti-avidin (15), while goat anti-rabbit IgG-AP was used as a secondary antibody. EGFP expression was carried out by growing the bacteria on Luria–Bertani plates containing IPTG (1 mM) and gentamycin (7 ug/ml), and EGFP-expressing colonies were detected directly under UV light.

Recombinant baculoviruses were prepared using the vectors pBVboostFG + EGFP, pBVboostFGR + EGFP and pBVboostFG + ProtX (Table 1) as described previously (12). Sf9 cells (1×10^6 in 1 ml) were infected with corresponding baculoviruses for 3 days. For the purification of His-tagged proteins, large-scale infections (200–500 ml) were performed and the His-tagged proteins were purified from the culture medium with Talon™ resin (BD Biosciences). The glycosylation pattern of the purified VEGF-A was studied with Endo H_f glycosidase and western blotting as described previously (15). The SEAP activity was measured by chemiluminescent assay according to the manufacturer's instructions (BD Great EscAPe™, NJ) and by dot plotting.

To test the constructed expression vector in mammalian cells, human hepatocarcinoma (HepG2) cells and chinese hamster ovary (CHO) cells were used for the expression of EGFP driven by the CAG promoter. The functionality of the vector was tested both by baculoviral transduction and by plasmid transfection (FuGENE™ 6, Roche) using pBVboostFG + EGFP. About 150 000 cells were plated on 6-well plates and after 24 h the cells were either transfected with 1–2 μg of plasmid DNA or transduced by the virus with MOI (multiplicity of infection) 300. The cells were incubated for 24 h and viewed under a fluorescence microscope. For BVboostFG + VEGF-A and BVboostFG + VEGF-$D^{\Delta N \Delta C}$ virus treatments BT4C cells were plated on 6-well plates (Nunclon™, Science kit & Boreal Laboratories, Tonawanda, NY) at a density of 150 000 cells per well. After 24 h of cultivation, viruses were added to the cells (MOI 1000) and incubated for 2 h at 37°C. Samples for VEGF ELISA analyses (R&D Systems, Minneapolis, MN) were taken after 24 h of culturing.

In vivo injection of baculovirus into rat brain

Eight female Wistar rats (200–250 g) were injected as described by Lehtolainen *et al.* (16) using 2×10^8 p.f.u. of

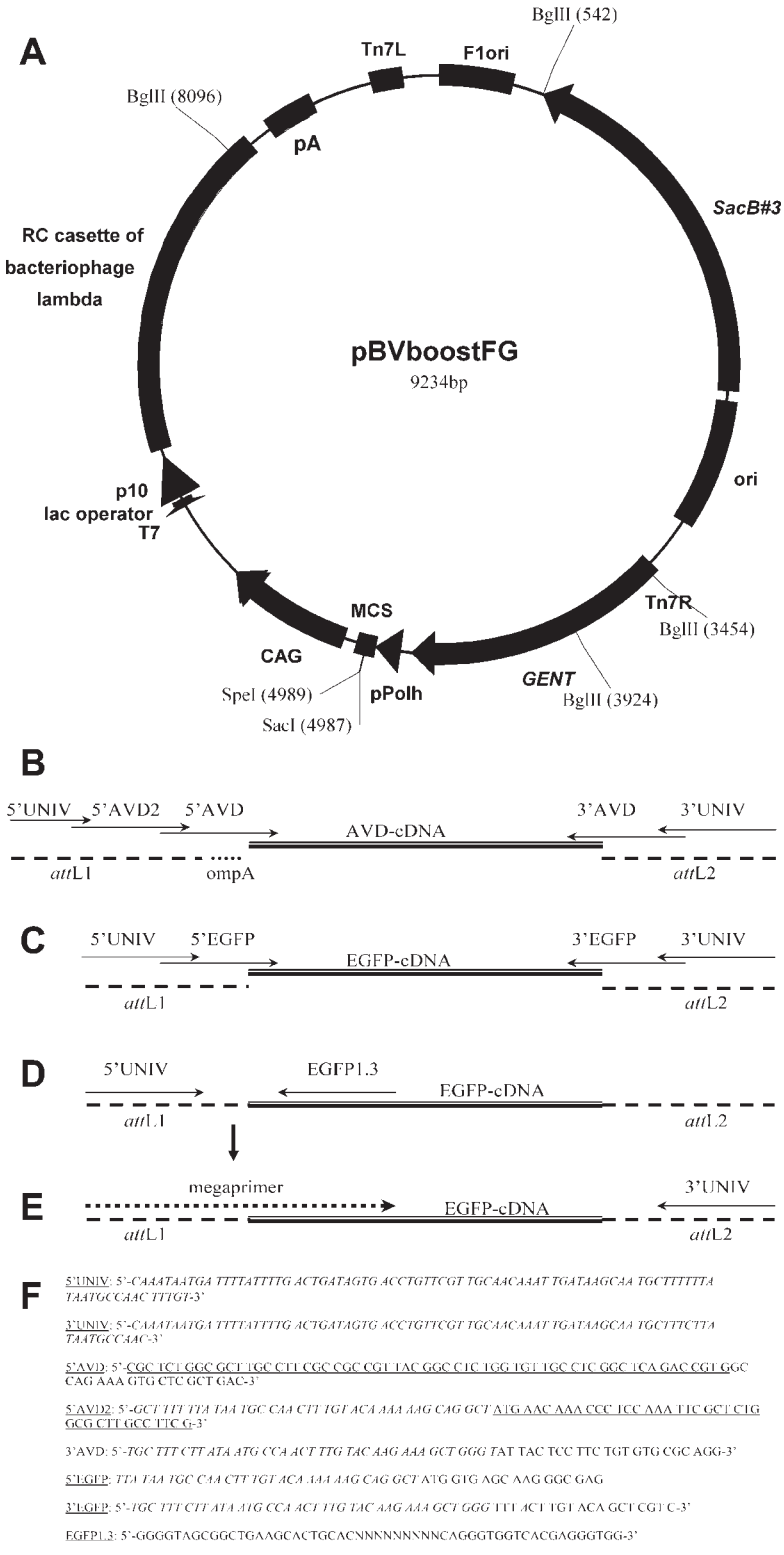

Figure 1. Map of the pBVBoostFG vector (**A**), SES-PCR strategy to construct avidin and EGFP cassettes for cloning into pBVboostFG (**B** and **C**) and the principle of the construction of the chromophore library (D and E). (A) SacB#3, a mutant form of the levansucrase gene (12); GENT, gentamycin resistance gene; Tn7R/L, left and right ends of bacterial transposon TnT7; pPohl, polyhedrin promoter; CAG, chicken β-actin promoter; T7, bacteriophage T7 promoter;. p10, p10 promoter; pA, transcriptional terminator area. (B) and (C) The dashed lines show the *att*L sites compatible with LR reaction of the recombinational cloning system and bacterial ompA signal (in avidin). (**D**) The first PCR reaction to construct the EGFP chromophore library: the product amplified in (C) was used as a template and amplification was performed with primers 5′UNIV and EGFP1.3. (**E**) The second PCR reaction for library construction, using the same template as in the first reaction. The product of the first reaction was used as a megaprimer and 3′UNIV was used as a second primer. The product of this reaction was LR-cloned into pBVboostFG. (**F**) Oligonucleotides to synthesize avidin and EGFP constructs and EGFP chromophore library compatible with LR reaction. The *att*L sequences are shown in italics and a sequence encoding the ompA signal peptide is underlined.

Table 1. Description of the vectors used in this study

Vector	Description
pBVboost	Parent vector for other constructs, allows high-throughput production of recombinant baculoviruses (12)
pBVboostFG	A derivative of the pBVboost, compatible with recombinational cloning and universal expression
pBVboostFGR	A derivative of the pBVboostFG, contains additional marker gene DsRed that is functional in insect cells
pBVboostFG + ProtX[a]	A derivative of the pBVboostFG for the expression of desired proteins
pBVboostFGR + EGFP	A derivative of the pBVboostFGR for the expression of EGFP

[a]ProtX represents different proteins cloned into pBVboostFG in the present study.

pBVboostFG + EGFP virus. The rats were sacrificed on day 5. The brains were removed and snap-frozen in liquid nitrogen for cryo-sectioning. These sections were visualized by fluorescence microscopy. Frozen sections were fixed in acetone for immunostaining. Samples were detected with anti-GFP (Biogenesis) as a primary antibody and with an AP conjugated anti-sheep (Bethyl Laboratories Inc.) as a secondary antibody followed by staining using AP substrate (Vector Alkaline Phosphatase Substrate Kit III, Vector Labs). Tissue samples were counterstained in Mayer's Carmalum and observed by microscope.

Screening of EGFP chromophore library

For the construction of the chromophore library the EGFP reading frame from pBVboostFG + EGFP was amplified using SES-PCR. This product was used for the construction of the chromophore library using a megaprimer method (17). The megaprimer was prepared by using primers 5′UNIV and EGFP1.3, of which the latter contains a random nucleotide mixture in the region that encodes the chromophore [amino acids 65–67 (18)]. In the second PCR reaction, this product was used as the megaprimer and the second primer was 3′UNIV (Figure 1D and E). The PCR product was then LR-cloned into pBVboostFG. The reaction end product contained the constructed chromophore library, which was transformed into BL21-AI™ (Invitrogen) *E.coli* that were plated onto LB plates containing gentamycin (7 ug/ml) and arabinose (0.2% w/v). After overnight incubation the fluorescent colonies were screened under UV illumination, transferred to LB medium and induced with L-arabinose (0.2% w/v). Following overnight cultivation, the cells (25°C) were collected by centrifugation (5000 *g*, 10 min) and suspended into 0.15 mg/ml lysozyme, 0.5 M glucose, 1 mM EDTA and 200 mM Tris (pH 7.4), and incubated for 30 min at room temperature and freeze-thawed. The lysates were supplemented with 50 mM $NaPO_4$ and 150 mM NaCl, pH 7.0, and the insoluble fraction was pelleted. Fluorescent excitation and emission spectra were measured from the soluble fraction with a spectrofluorometer (PerkinElmer LS 55 Luminescence Spectrometer) and the obtained spectra were normalized according to the fluorescence maximum. A sample from one non-fluorescent clone was prepared identically to obtain a negative control. The identified fluorescent clones and several non-fluorescent clones were then sequenced using Li-cor 4200S-2 DNA Sequencer. Fluorescent viruses of the identified clones were produced and tested in SF9 cells.

RESULTS

Construction of a multipurpose vector compatible with library screening and gene expression in mammalian, insect and bacterial cells

In order to construct a single vector that allows expression of cloned genes or cDNA libraries in different host systems without any further subcloning, we combined four different promoters into the same vector. This tetra-promoter cassette is composed of CAG (CMVie enhancer + chicken β-actin promoter), T7*lac*, pPolh and p10, all of which direct efficient expression of genes in vertebrate cells, *E.coli*, and baculovirus-infected insect cells, respectively (Figure 1A). The multiple cloning site following the pPolh promoter allows an option to modify the properties of the baculoviruses or to express a marker gene to detect the synthesis of recombinant baculoviruses, as exemplified here with red fluorescent protein. To allow the efficient cloning of the desired libraries (or genes/cDNAs) into the vector, a site-specific recombination cassette of the bacteriophage lambda containing *attR1/2* sites (8) was included into the vector. To further enable fast and high-throughput production of the recombinant baculoviruses, we cloned the tetra-promoter cassette as part of the pBVboost vector that enables zero background generation of recombinant baculovirus genomes with high diversity ($\sim 10^7$ c.f.u./μg) (12), making the vector optimal for library screening. The resulting vector was designated as pBVboostFG. The cloning strategy of test genes into pBVboostFG is shown in Figure 1. For easy detection of the produced recombinant baculoviruses, red fluorescent protein was cloned into the multiple cloning site under the pPolh promoter. The vectors used in this study are summarized in Table 1.

Expression of genes in multiple hosts

To test the functionality of the tetra-promoter in different expression hosts, we cloned EGFP into pBVboostFG. *E.coli* was transformed with this plasmid and the efficient production of EGFP was easily detected from bacterial cultures growing on LB plates (Figure 2A). Baculoviruses derived from the same plasmid were used to infect Sf9 cells. After 3 days infection the cells were studied using a fluorescence microscope and almost all cells showed bright green fluorescence (Figure 2B and C). Simultaneous function of both baculoviral promoters in the Sf9 cells was shown by infecting Sf9 cells with the virus construct containing red fluorescent protein under pPohl and EGFP under P10 promoters (Figure 2E–G). The BVboostFG + EGFP baculovirus was used to transduce HepG2 (Figure 2H and I) and CHO cells in order to show that the tetra-promoter construct is also active in mammalian cells. Green fluorescence was also detected in the plasmid transfected CHO cells (data not shown).

The broad feasibility of the pBVboostFG vector was also demonstrated by producing several different proteins in different host cells (Figures 2 and 3) (Table 2). The bacterial

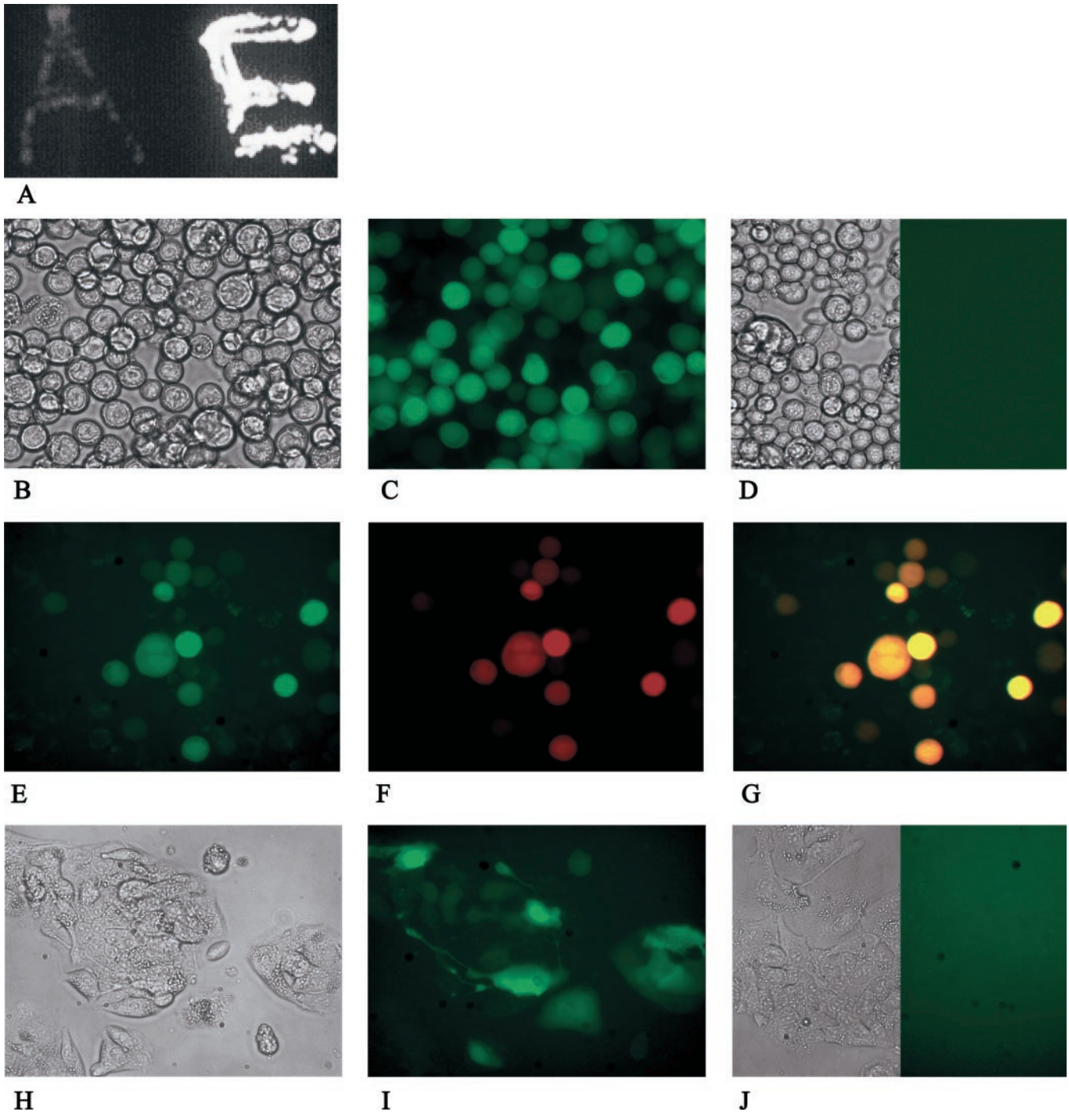

Figure 2. Expression of EGFP in different host cells. (**A**) EGFP expression in bacteria: the culture forming an alphabet A contains an avidin cassette (negative control in this experiment) whereas E denotes the culture that expresses EGFP. The plate was illuminated under UV light. (**B**) Sf9 cells infected with pBVboostFG + EGFP-derived baculovirus, three days after infection, in normal light. (**C**) Green fluorescence from the same cells detected using fluorescence microscopy. (**D**) A negative control (non-infected cells; left, phase contrast; right, fluorescent image). In (E–G) Sf9 cells infected with pBVboostFGR + EGFP derived baculovirus. (**E**) Cells were studied by detecting green fluorescence whereas in (**F**) the red fluorescence was detected. (**G**) The merged image of (E and F). (H and I) HepG2 cells 24 h after transduction by pBVboostFG+EGFP derived baculovirus. (**H**) Cells viewed using phase contrast microscopy and (**I**) the same cells detected by the emitted green fluorescence. No virus was used for the negative control (**J**). The magnification of the cells is 20-fold.

ompA secretion signal was fused to the avidin gene in order to transport the synthesized avidin into the periplasmic space of *E.coli* (Figure 1B). The periplasmic expression of avidin was efficient in *E.coli* and a remarkable proportion of the total cellular protein was composed of avidin after overnight induction. Some of the avidin was produced as insoluble inclusion bodies (Figure 3A) and the inclusion bodies, as well as the total cell sample, also contained a non-processed form of the

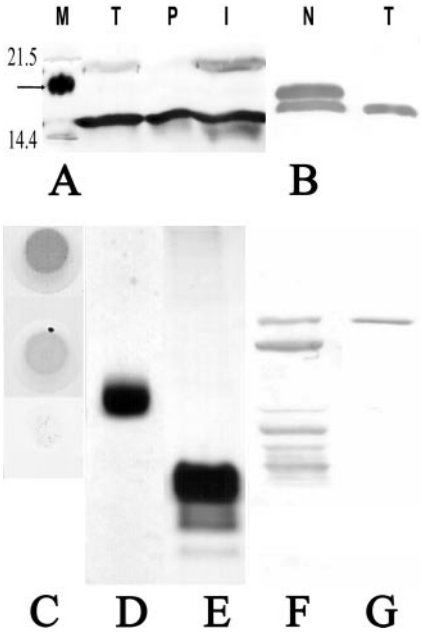

Figure 3. Analyses of the proteins expressed by the BVboostFG system. (**A**) Western blot of the expressed avidin detected with an avidin antibody. M, molecular weight markers; control avidin marked with an arrow. T, sample of total *E.coli* proteins; P, periplasmic fraction; I, insoluble fraction. The upper faint band represents ompA secretion signal-containing avidin whereas the lower band represents the fully processed avidin. Control avidin has a higher molecular weight because it contains a carbohydrate side chain (38). (**B**) Deglycosylation analysis of purified VEGF-A produced in insect cells. N indicates a sample not treated with endo H_f and T indicates a sample deglycosylated with the enzyme. (**C**) A dot-blot analysis of SEAP secreted into insect cell medium. The uppermost dot is 200 µl of medium diluted 1:10, whereas the middle dot is a positive control containing 300 ng of SEAP. The lowest dot is a negative control. (**D**) and (**E**) show purified sKDR(1–7) (silver staining) and VEGF-$D^{\Delta N\Delta C}$ (Coomassie staining) proteins in SDS–PAGE gel, respectively. The proteins were purified to milligram amounts with no apparent contaminants which indicate high purity. The lower bands seen in VEGF-$D^{\Delta N\Delta C}$ preparation are different glycosylation forms of the protein. (**F**) and (**G**) show immunoblots of HIV-integrase produced in *E.coli* and insect cells, respectively. Most of the protein produced in bacteria was degraded whereas protein produced in insect cells was intact. Proteins were detected with an integrase-specific antibody.

protein (i.e. protein that still contained the signal peptide). In contrast, the ompA signal was cleaved off from virtually all periplasmic avidins. The biological activity of the periplasmic avidin was shown by its binding to biotinylated agarose (data not shown). The purified VEGF-A produced in the insect cells contained N-linked glycomoieties typical for wild-type VEGF-A (Figure 3B).

A collection of proteins expressed in different host cells are shown in Table 2. The yields of produced proteins varied between different proteins and hosts. Protein yields were typically in the range of 3–20 mg/l from *E.coli* periplasm and several mg/l from insect cell cultures. Some proteins, such as SEAP, were expressed in very high concentrations in insect cells (Figure 3C). High purity after a single affinity-chromatography step was typically achieved using immobilized metal affinity chromatography or biotin/2-iminobiotin chromatography. As an example no contaminating bands were detected in sKDR(1–7), VEGF-$D^{\Delta N\Delta C}$ (Figure 3D and E) and several avidin preparations (data not shown). HIV-integrase shows the strength and flexibility of our system: its production was strong in *E.coli* but the protein was mostly degraded (Figure 3F). Instead, infection of insect cells led to the production of intact protein (Figure 3G), which showed normal biological activity in *in vitro* tests (data not shown) after affinity chromatography. The VEGF-A and VEGF-$D^{\Delta N\Delta C}$ expression in mammalian cells was comparable with that achieved with adenoviruses containing similar expression cassettes under the CMV promoter (19): at the 24 h time point, expression yields for VEGF-A and VEGF-$D^{\Delta N\Delta C}$ were 441 and 1094 pg/ml, respectively with the pBVboostFG vector. Correspondingly, 861 pg/ml (VEGF-A) and 982 pg/ml (VEGF-$D^{\Delta N\Delta C}$) were produced by the adenovirus.

EGFP expression in rat brain

To verify the function of the vector system *in vivo*, EGFP expressing baculovirus was injected into the rat right lateral ventricle. Choroid plexus epithelial cells were shown to be transduced with high efficiency as reported previously by Lehtolainen *et al*. (16). Strong EGFP production was detected in choroid plexus cells as seen by anti-GFP immunostaining and fluorescence microscopy (Figure 4).

Screening of the chromophore library

The synthesized chromophore library was screened as described in Materials and Methods. Because three consecutive amino acid residues were randomized, the theoretical diversity of the library was 9261 (21^3, twenty different amino acids + termination) at the amino acid level and 262144 (4^9) at the nucleic acid level. Out of 20 000 screened colonies, three colonies showed fluorescence under UV light. These clones were amplified and sequenced (Table 3), and their chromophore sequences were found to be SYG, LYG and THG instead of the TYG in the wild-type EGFP. In order to verify the diversity of the constructed library, eight non-fluorescent colonies were also sequenced and they all had different sequences when compared to wild-type EGFP and to each other (Table 3). Some non-fluorescent clones were also induced to express the non-functional EGFP and the produced protein was recognized on western blot by the EGFP antibody, which confirmed that products with the correct size were synthesized (data not shown).

The excitation and emission spectra of the fluorescent clones were measured to study their spectroscopic properties (Table 3) (Figure 5). The EGFP-mutant SYG resembles the wild-type GFP in its spectral properties (λ^{ex}_{max} = 395–397 nm, λ^{em}_{max} = 504 nm). The slight change towards red in the emission maximum of this variant could be due to the F64L change in the EGFP-template as compared with the wild-type GFP. The LYG variant shows spectral properties close to EGFP. The THG variant shows spectral properties similar to those reported for blue fluorescent protein.

DISCUSSION

The post-genomic era requires tools to translate the enormous nucleotide sequence data into biologically relevant knowledge, thus, there is a clear need for versatile vector systems like the pBVboostFG, developed in this study (5,20,21). One of the main benefits of these kinds of vectors is their suitability for many alternative host systems: the library or the single gene/cDNA can be expressed in *E.coli*, insect cells and mammalian

Table 2. Proteins expressed with pBVboostFG system in different host cells

Protein	Description and reference	Added tag	Used host	Analyses
EGFP	Enhanced green fluorescent protein and its screened mutants	—	*E.coli*/ insect/ mammalian	Fluorescent images, spectral analyses, western blots
DsRed	Fast maturation variant of *Discosoma* sp. Red fluorescent protein	—	Insect	Fluorescent images
VEGF-A_{165}	Vascular endothelial growth factor A, 165 amino acid long form (39)	C-terminal 6-His	Insect/mammalian	Purified by affinity chromatography/ELISA
VEGF-B_{186}	Vascular endothelial growth factor B, 186 amino acid long form (39)	C-terminal 6-His	Insect	Purified by affinity chromatography
VEGF-$D^{\Delta N \Delta C}$	Vascular endothelial growth factor D, N and C terminally truncated form (39)	C-terminal 6-His, N-terminal Flag taq	Insect/mammalian	Purified by affinity chromatography/ELISA
sKDR(1–7)	Soluble form of vascular endothelial growth factor receptor 2, includes IgG like domains 1-7 (39)	C-terminal 6-His	Insect	Purified by affinity chromatography
sKDR(1–4)	Soluble form of vascular endothelial growth factor receptor 2, includes IgG like domains 1-4 (39)	C-terminal 6-His	Insect	Purified by affinity chromatography
Flt1(1–5)	Soluble form of vascular endothelial growth factor receptor 1, includes IgG like domains 1-5 (39)	C-terminal 6-His	Insect	Purified by affinity chromatography
Flt1(1–3)	Soluble form of vascular endothelial growth factor receptor 1, includes IgG like domains 1-3 (39)	C-terminal 6-His	Insect	Protein detected with His-tag antibody
Avidin	(13,40)	—	*E.coli*	Purified by affinity chromatography
AVR2	Avidin related protein 2 (15)	—	*E.coli*	Purified by affinity chromatography
AVR4/5	Avidin related protein 4/5 (15)	—	*E.coli*	Purified by affinity chromatography
DcAvd	Engineered dual chain form of chicken avidin (41)	—	*E.coli*	Purified by affinity chromatography
LDAvdN54A	Modified monoavidin (42)	—	*E.coli*	Purified by affinity chromatography
Streptavidin	Avidin analogue from *Streptomyces avidiini* bacteria (43)	—	*E.coli*	Protein detected by immunoblotting
HCL-1	Hypothetical calycin like protein from *Bradyrhizobium japonicum* (45)	—	*E.coli*	Purified by affinity chromatography
HIV integrase	(44)	N-terminal 6-His	*E.coli*/insect	Purified by affinity chromatography
SEAP	Secreted alkaline phosphatase, GenBank accession no. U89937	—	Insect	Dot-blot and chemiluminescence analyses

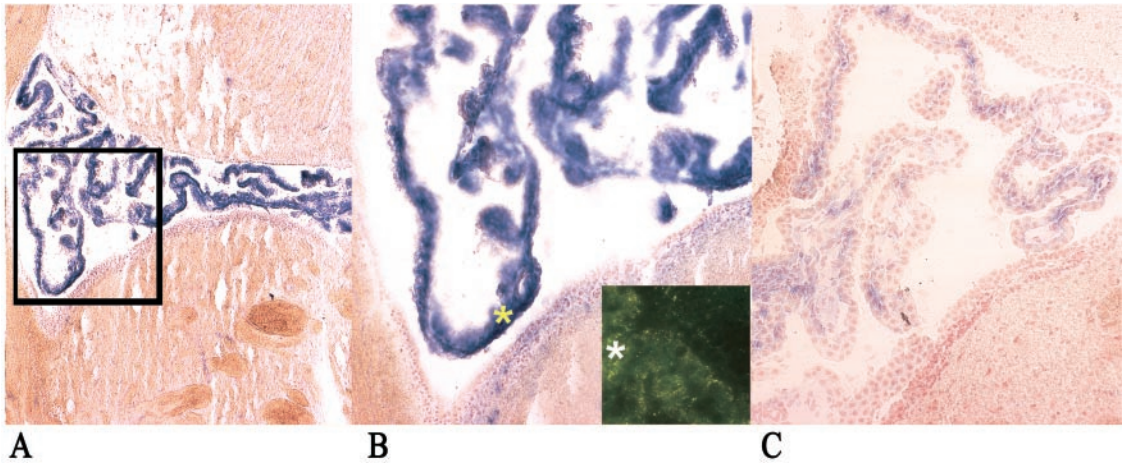

Figure 4. Baculovirus-mediated EGFP (pBVboostFG + EGFP) gene transfer in rat brain. (**A**) Rat brain choroid plexus cells show ~100 % transduction efficiency (immunostaining, original magnification 4×). In (**B**), the area boxed in (A) is shown at 20× magnification, insert shows fluorescent image of the choroid plexus cells transduced with the virus. Asterisks indicate choroid plexus cells. (**C**) A negative control from a non-transduced rat at the same magnification as in (B).

cells (22,23). This makes it possible not only to efficiently screen libraries in these hosts but also to utilize the best characteristics of each specific host for individual purposes as shown in the current study. *E.coli* is the most widely utilized prokaryotic expression system, which provides a cost-efficient alternative to produce proteins in large quantities. Protein production in *E.coli*, however, often leads to poor expression yields and insoluble end products. Insect cells are able to carry

Table 3. Sequences of the chromophore-encoding region of three positive clones found in the screening of the EGFP chromophore library and their fluorescence properties

Clone	DNA-sequence	Amino acid sequence	Fluorescence properties λ_{max}^{ex} (nm)	λ_{max}^{em} (nm)
EGFP[a]	ACCTACGGC	TYG	489	509
F1	AGCTACGGA	SYG	395(475)[b]	508
F2	CTATATGGC	LYG	478	506
F5	ACTCACGGC	THG	380	447
NF1[c]	CGCGCATTG	RAL		
NF2	GGACAACAC	GQH		
NF3	GGAACACGC	GTR		
NF4	AATATTCCG	NIP		
NF5	CTACCAACA	LPT		
NF6	CAGAAGAAC	QKN		
NF7	GCAAAGATA	AKI		
NF8	AAAAAAAGC	KKS		

The sequences of the eight non-fluorescent clones are also shown.
[a]wild-type EGFP sequence.
[b]secondary peak in excitation spectrum.
[c]non-fluorescent clone that was used for library construction to eliminate wild-type fluorescent background.

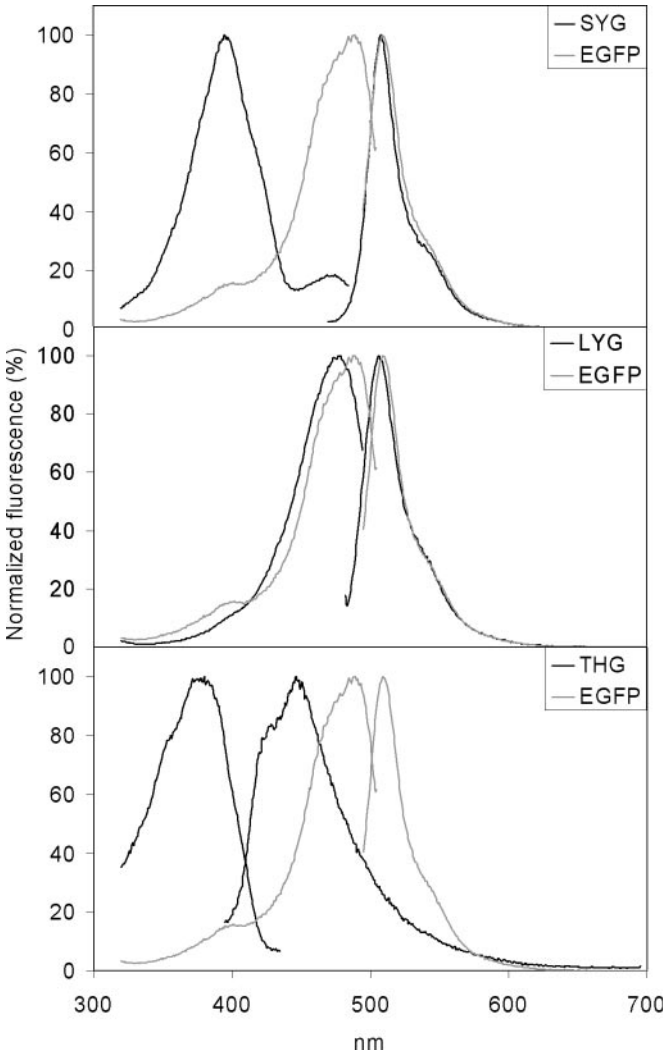

Figure 5. The excitation and emission spectra (black) of the three fluorescent protein mutants identified by library screening. The uppermost spectra are from the SYG variant, the middle spectra from the LYG variant and the lowest spectra from the THG variant. The measured spectra of the wild-type EGFP are shown for comparison (grey).

out most post-translational modifications typical for eukaryotic proteins that may be required for the full activity of the produced protein (24). Indeed, the baculovirus expression system has become a widely used eukaryotic protein-expression system. Some complex proteins may need to be produced in mammalian cells that have the unique ability to produce properly processed and biologically active molecules at reasonable level. Ideally, a protein should be expressed in a host as near to its origin as possible. The pBVboostFG system therefore provides a straightforward tool to address this issue.

In general, good expression rates were achieved using the pBVboostFG tetra-promoter vector in different hosts. This suggests that the tetra-promoter design allows protein production with an efficacy comparable with single promoter vectors. The universal expression compatibility does not, however, necessarily guarantee efficient production of all proteins in all hosts. For example VEGF-$D^{\Delta N \Delta C}$ and sKDR(1–7) were not produced successfully in *E.coli*. This is most likely due to the fact that the signal peptides of these eukaryotic proteins are not recognized by the bacterial secretion machinery (25). Correspondingly, to produce chicken avidin efficiently in *E.coli*, a bacterial signal sequence was utilized. Another well-recognized issue relevant to the universal expression concept is codon usage. In highly expressed genes, a narrow set of species-specific codons is used, which correspond to the more abundant tRNA species (26). However, this does not change the fact that the tetra-promoter concept allows rapid and straightforward screening through the different hosts and, depending on the most preferential host, known sequence constraints can be taken into account in cloning design.

The construction of baculovirus libraries with the pBVboostFG based system can be accomplished in as little as one week (12). Easy baculovirus generation and amplification provides a rapid transition from *in vitro* library screening to animal testing without any further subcloning steps and therefore, it markedly facilitates the screening of disease-related genes (11). In this context it is important to note that baculoviruses can deliver genes into diverse types of cells from different species (27,28). Viral replication is restricted to specific insect hosts and the baculovirus genome is silent in non-permissive cells (27). Further benefit of baculoviruses is their low cytotoxity (11). An additional advantage of using baculovirus is that it should be possible to screen also large DNA inserts (10,29). In contrast, adenoviral and retroviral gene transfer vectors can incorporate <8 kb of foreign DNA into their genomes.

The pBVboostFG system offers several advantages as compared with previously described multiple promoter vectors (22,23). First, the novel tetra-promoter construct allows dual-gene expression in insect cells. Two insect-cell-specific promoters were cloned into pBVboostFG as separate expression cassettes. This allows preparation of pseudotyped baculoviruses for enhanced gene delivery into vertebrate cells and animals (30–33) or visualization of virus infection (titering), which eases and speeds up virus preparation as shown in the current study. Second, because recombinant baculovirus

genomes are generated in *E.coli*, there is no need to carry out time-consuming plaque purifications to isolate separate clones. This will accelerate the screening and generation of recombinant viruses. Third, recombinant bacmids can be generated at high frequency ($\sim 10^7$ c.f.u./μg) virtually free of background (12), which yields high virus diversity more efficiently than traditional homologous recombination methods (34) or methods based on conventional restriction enzyme cloning (35). Fourth, because no restriction enzyme digestions are required for cloning, the yield of full-length clones is improved since the inserts will not be cut from internal restriction sites. The use of the bacteriophage lambda recombination system also provides better cloning efficiency than restriction–ligation based strategies (36). Furthermore, the site-specific recombination system of bacteriophage lambda is reversible, in contrast to many other corresponding site-specific recombinase systems. This feature ensures that, if desired, any fragment cloned into the pBVboostFG vector should be easily transferred to any other vector utilizing the same system, and *vice versa* (8).

To the best of our knowledge, the current study provides the first example where simple SES-PCR has been used to generate the long *att*L sites required for recombinational cloning utilizing the bacteriophage lambda system (8) on the scale of library construction. This strategy has the benefit that SES-PCR products can be cloned directly into recombinational cloning-compatible expression vectors like pBVboostFG. This makes the method fast and cost-effective. The SES-PCR strategy fits well for randomization and DNA shuffling of genes as demonstrated in the present study. For high-throughput setups, however, we would find it interesting to convert the ready-made entry format libraries such as those described by Pearlberg and LaBaer (5) into pBVboostFG format. This would be especially useful if such annotated expression clones already contained fusion tags suitable for the detection and purification of the produced proteins. The current form of the pBVboostFG does not include an option for fusion tags since His-tag, for example, can be readily incorporated into PCR-primers and this strategy avoids incorporation of extra amino acids into produced proteins.

In vivo transduction of rat brain choroid plexus with BVboostFG baculovirus provides an interesting target tissue in the central nervous system; it is responsible for the production and maintenance of cerebrospinal fluid (37). In addition, its involvement in the pathological processes has been documented (37). Therefore, it is intriguing to speculate the possibility of using BVboostFG viruses as a tool to produce secreted therapeutic compounds in the cerebrospinal fluid or cure disorders associated with the functions of the choroid plexus.

It is concluded that the pBVboostFG system is a novel tool for functional genomics that allows a simple, rapid and simultaneous analysis of libraries and cloned genes in *E.coli*, insect and vertebrate cells as well as directly in *in vivo* animal models without time-consuming subcloning steps.

ACKNOWLEDGEMENTS

The authors thank Ms Tarja Taskinen and Ms Irene Helkala for the first-class technical assistance. This study was supported by grants from the Finnish Academy (numbers 203982 and 205728), Ark Therapeutics Oy and the National Graduate School in Informational and Structural Biology. We would also like to thank Drs Mervi Ahlroth and Tomi Airenne for comments on the manuscript and Ms Nicola Wakelin for revising the language. This study is dedicated to the memory of Eeva Laitinen. Funding to pay the Open Access publication charges for this article was provided by Ark Therapeutics Oy.

REFERENCES

1. Marcotte,E. and Date,S. (2001) Exploiting big biology: integrating large-scale biological data for function inference. *Brief. Bioinformatics*, **2**, 363–374.
2. McGall,G.H. and Christians,F.C. (2002) High-density genechip oligonucleotide probe arrays. *Adv. Biochem. Eng. Biotechnol.*, **77**, 21–42.
3. Eickhoff,H., Konthur,Z., Lueking,A., Lehrach,H., Walter,G., Nordhoff,E., Nyarsik,L. and Bussow,K. (2002) Protein array technology: the tool to bridge genomics and proteomics. *Adv. Biochem. Eng. Biotechnol.*, **77**, 103–112.
4. Whitelegge,J.P. and le Coutre,J. (2001) Proteomics. Making sense of genomic information for drug discovery. *Am. J. Pharmacogenomics*, **1**, 29–35.
5. Pearlberg,J. and LaBaer,J. (2004) Protein expression clone repositories for functional proteomics. *Curr. Opin. Chem. Biol.*, **8**, 98–102.
6. Michiels,F., van Es,H., van Rompaey,L., Merchiers,P., Francken,B., Pittois,K., van der,S.J., Brys,R., Vandersmissen,J., Beirinckx,F. *et al.* (2002) Arrayed adenoviral expression libraries for functional screening. *Nat. Biotechnol.*, **20**, 1154–1157.
7. Koller,D., Ruedl,C., Loetscher,M., Vlach,J., Oehen,S., Oertle,K., Schirinzi,M., Deneuve,E., Moser,R., Kopf,M. *et al.* (2001) A high-throughput alphavirus-based expression cloning system for mammalian cells. *Nat. Biotechnol.*, **19**, 851–855.
8. Hartley,J.L., Temple,G.F. and Brasch,M.A. (2000) DNA cloning using *in vitro* site-specific recombination. *Genome Res.*, **10**, 1788–1795.
9. Airenne,K.J., Hiltunen,M.O., Turunen,M.P., Turunen,A.M., Laitinen,O.H., Kulomaa,M.S. and Yla-Herttuala,S. (2000) Baculovirus-mediated periadventitial gene transfer to rabbit carotid artery. *Gene Ther.*, **7**, 1499–504.
10. Cheshenko,N., Krougliak,N., Eisensmith,R.C. and Krougliak,V.A. (2001) A novel system for the production of fully deleted adenovirus vectors that does not require helper adenovirus. *Gene Ther.*, **8**, 846–854.
11. Kost,T.A. and Condreay,J.P. (2002) Recombinant baculoviruses as mammalian cell gene-delivery vectors. *Trends Biotechnol.*, **20**, 173–180.
12. Airenne,K.J., Peltomaa,E., Hytonen,V.P., Laitinen,O.H. and Yla-Herttuala,S. (2003) Improved generation of recombinant baculovirus genomes in *Escherichia coli*. *Nucleic Acids Res.*, **31**, e101.
13. Gope,M.L., Keinänen,R.A., Kristo,P.A., Conneely,O.M., Beattle,W.G., Zarucki-Schulz,T., O'Malley,B.W. and Kulomaa,M.S. (1987) Molecular cloning of the chicken avidin cDNA. *Nucleic Acids Res.*, **15**, 3595–3606.
14. Majumder,K. (1992) Ligation-free gene synthesis by PCR: synthesis and mutagenesis at multiple loci of a chimeric gene encoding ompA signal peptide and hirudin. *Gene*, **116**, 115–116.
15. Laitinen,O.H., Hytonen,V.P., Ahlroth,M.K., Pentikainen,O.T., Gallagher,C., Nordlund,H.R., Ovod,V., Marttila,A.T., Porkka,E., Heino,S. *et al.* (2002) Chicken avidin-related proteins show altered biotin-binding and physico-chemical properties as compared with avidin. *Biochem. J.*, **363**, 609–17.
16. Lehtolainen,P., Tyynela,K., Kannasto,J., Airenne,K.J. and Yla-Herttuala,S. (2002) Baculoviruses exhibit restricted cell type specificity in rat brain: a comparison of baculovirus- and adenovirus-mediated intracerebral gene transfer *in vivo*. *Gene Ther.*, **9**, 1693–1699.
17. Sarkar,G. and Sommer,S.S. (1990) The ''Megaprimer'' method of site-directed mutagenesis. *Biotechniques*, **8**, 404–407.
18. Cormack,B.P., Valdivia,R.H. and Falkow,S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, **173**, 33–38.
19. Rissanen,T.T., Markkanen,J.E., Gruchala,M., Heikura,T., Puranen,A., Kettunen,M.I., Kholova,I., Kauppinen,R.A., Achen,M.G.,

Stacker,S.A. *et al*. (2003) VEGF-D is the strongest angiogenic and lymphangiogenic effector among VEGFs delivered into skeletal muscle via adenoviruses. *Circ. Res.*, **92**, 1098–1106.
20. Braun,P. and LaBaer,J. (2003) High throughput protein production for functional proteomics. *Trends Biotechnol.*, **21**, 383–388.
21. Rual,J.F., Hill,D.E. and Vidal,M. (2004) ORFeome projects: gateway between genomics and omics. *Curr. Opin. Chem. Biol.*, **8**, 20–25.
22. Novy,R, Yaeger,K, Monsma,S and Scott,M (1999) pTriEx-1 Multisystem Vector for protein expression in *E.coli*, mammalian, and insect cells. *Innovations*, **10**, 2–5.
23. Buttner,C. and Skupin,A. (2002) Rapid analysis of the extracellular matrix protein decorin using the penta-his Alexa Fluor 488 Conjugate. *QIAGEN News*, **4**, 21–22.
24. O'reilly,D.R., Miller,L.K. and Luckov,V.A. (1994) *Baculovirus expression vectors. A laboratory Manual*. Oxford University Press, NY.
25. Humphreys,D.P., Sehdev,M., Chapman,A.P., Ganesh,R., Smith,B.J., King,L.M., Glover,D.J., Reeks,D.G. and Stephens,P.E. (2000) High-level periplasmic expression in *Escherichia coli* using a eukaryotic signal peptide: importance of codon usage at the 5′ end of the coding sequence. *Protein Expr. Purif.*, **20**, 252–264.
26. Gustafsson,C., Govindarajan,S. and Minshull,J. (2004) Codon bias and heterologous protein expression. *Trends Biotechnol.*, **22**, 346–353.
27. Volkman,LE. and Goldstein,R.A. (1983) *In vitro* Survey of *Autographa californica* nuclear polyhedrosis virus interaction with nontarget vertebrate host cells. *Appl. Biochem. Biotechnol.*, **45**, 1085–1093.
28. Leisy,D.J., Lewis,T.D., Leong,J.A. and Rohrmann,G.F. (2003) Transduction of cultured fish cells with recombinant baculoviruses. *J. Gen. Virol.*, **84**, 1173–1178.
29. Peters,J.E. and Craig,N.L. (2001) Tn7: smarter than we thought. *Nature Rev. Mol. Cell Biol.*, **2**, 806–814.
30. Huser,A., Rudolph,M. and Hofmann,C. (2001) Incorporation of decay-accelerating factor into the baculovirus envelope generates complement-resistant gene transfer vectors. *Nat. Biotechnol.*, **19**, 451–455.
31. Tani,H., Limn,C.K., Yap,C.C., Onishi,M., Nozaki,M., Nishimune,Y., Okahashi,N., Kitagawa,Y., Watanabe,R., Mochizuki,R. *et al*. (2003) *In vitro* and *in vivo* gene delivery by recombinant baculoviruses. *J. Virol.*, **77**, 9799–9808.
32. Raty,J.K., Airenne,K.J., Marttila,A.T., Marjomaki,V., Hytonen,V.P., Lehtolainen,P., Laitinen,O.H., Mahonen,A.J., Kulomaa,M.S. and Yla-Herttuala,S. (2004) Enhanced gene delivery by avidin-displaying baculovirus. *Mol. Ther.*, **9**, 282–291.
33. Kukkonen,S.P., Airenne,K.J., Marjomaki,V., Laitinen,O.H., Lehtolainen,P., Kankaanpaa,P., Mahonen,A.J., Raty,J.K., Nordlund,H.R., Oker-Blom,C. *et al*. (2003) Baculovirus capsid display: a novel tool for transduction imaging. *Mol. Ther.*, **8**, 853–862.
34. Granziero,L., Nelboeck,P., Bedoucha,M., Lanzavecchia,A. and Reid,H.H. (1997) Baculovirus cDNA libraries for expression cloning of genes encoding cell-surface antigens. *J. Immunol. Methods*, **203**, 131–139.
35. Ernst,W., Grabherr,R., Wegner,D., Borth,N., Grassauer,A. and Katinger,H. (1998) Baculovirus surface display: construction and screening of a eukaryotic epitope library. *Nucleic Acids Res.*, **26**, 1718–1723.
36. Ohara,O. and Temple,G. (2001) Directional cDNA library construction assisted by the in vitro recombination reaction. *Nucleic Acids Res.*, **29**, e22.
37. Strazielle,N. and Ghersi-Egea,J.F. (2000) Choroid plexus in the central nervous system: biology and physiopathology. *J. Neuropathol. Exp. Neurol.*, **59**, 561–574.
38. Nordlund,H.R., Hytonen,V.P., Laitinen,O.H., Uotila,S.T., Niskanen,E.A., Savolainen,J., Porkka,E. and Kulomaa,M.S. (2003) Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding. *FEBS Lett.*, **555**, 449–454.
39. Neufeld,G., Cohen,T., Gengrinovitch,S. and Poltorak,Z. (1999) Vascular endothelial growth factor (VEGF) and its receptors. *FASEB J.*, **13**, 9–22.
40. Hytonen,V.P., Laitinen,O.H., Airenne,T.T., Kidron,H., Meltola,N.J., Porkka,E.J., Horha,J., Paldanius,T., Maatta,J.A., Nordlund,H.R. *et al*. (2004) Efficient production of active chicken avidin using a bacterial signal peptide in *Escherichia coli*. *Biochem. J.*, **384**, 385–390.
41. Nordlund,H.R., Laitinen,O.H., Hytonen,V.P., Uotila,S.T., Porkka,E. and Kulomaa,M.S. (2004) Construction of a dual-chain pseudo-tetrameric chicken avidin by combining two circularly permuted avidins. *J. Biol. Chem.*, **279**, 36715–36719.
42. Laitinen,O.H., Marttila,A.T., Airenne,K.J., Kulik,T., Livnah,O., Bayer,E.A., Wilchek,M. and Kulomaa,M.S. (2001) Biotin induces tetramerization of a recombinant monomeric avidin. A model for protein–protein interactions. *J. Biol. Chem.*, **276**, 8219–24.
43. Green,N. (1990) Avidin and streptavidin. *Meth. Enzymol.*, **184**, 51–67.
44. Sioud,M. and Drlica,K. (1991) Prevention of human immunodeficiency virus type 1 integrase expression in *Escherichia coli* by a ribozyme. *Proc. Natl Acad. Sci. USA*, **88**, 7303–7307.
45. Nordlund,H.R., Hytonen,V.P., Laitinen,O.H. and Kulomaa,M.S. (2005) Novel avidin-like protein from a root nodule symbiotic bacterium *Bradyrhizobium japonicum*. *J. Biol. Chem.*, doi:10.1074/jbc.M414336200.