

Getah 바이러스를 이용한 신규 유용 단백질
발현 체계 개발

Development of a New Expression System
Derived from Getah virus

충남대학교 수의과대학

농림부

제 출 문

농림부 장관 귀하

본 보고서를 “ Getah 바이러스를 이용한 신규 유용 단백질 발현 체계 개발에 관한 연구”
과제의 최종 보고서로 제출합니다.

2004년 8월 6일

주관연구기관명 : 충남대학교
총괄연구책임자 : 김철중
연구원 : 김현수
연구원 : 신광순
연구원 : 김민정
연구원 : 오윤이
연구원 : 서승희
연구원 : 임 송
연구원 : 최용수
연구원 : 이준한
연구원 : 유 정
연구원 : 신선아
연구원 : 오미정
연구원 : 이경호

요 약 문

I. 제목

Gatah 바이러스를 이용한 신규 유용 단백질 발현체계 개발

II. 연구 개발의 목적 및 필요성

외래유전자의 세포내 발현은 현재의 생명공학분야에 있어서 핵심이 되는 사항이며, 아직 고등진핵생물에서의 발현시스템은 효율적이지 못하다. 기존의 외래 단백질 발현 시스템은 낮은 단백질 발현량, 좁은 숙주범위, 실험조작시의 어려움, 강한 cytopathic effect, 실험실에서 RNA transcript 생성에 따른 고비용과 효율문제, 백신 후보자로서의 한계등 다양하다.

Getah virus는 Togaviridae의 Alphavirus에 속하며 이에는 Sindbis virus, Semliki Forest virus 등 이미 상용화된 바이러스 벡터 시스템과 같은 군에 속한다. 이러한 Alphavirus를 이용한 단백질 발현 시스템은 숙주세포의 범위가 넓고, 외래단백질을 대량생산할 수 있다는 잇점이 있으며, 최근에는 유전자 치료와 백신연구 등 많은 실험분야에 응용되고 있는 실정이다. 따라서 게타 바이러스를 이용한 외래단백질 발현 시스템을 구축함으로써 많은 연구기관에서는 외래유전자의 발현시스템으로 편리하게 이용될 수 있을 뿐만 아니라 이를 이용한 신약개발, 항암치료등 여러분야에서 보다 넓은 총체적인 생명공학 분야의 발전을 가져올 수 있으며, 또한 이 시스템 자체를 외래유전자를 전달하는 캐리어 백신으로 이용하여 돼지나 가금등에서 다른 질병을 예방하는 생백신으로서의 기능을 더불어 할 수 있을 것으로 기대된다.

III. 연구 개발 내용 및 범위

1. 게타바이러스를 진핵세포에 배양 후 배양 상층액으로부터 의 RNA 분리
2. 분리된 게타바이러스 RNA로부터 cDNA 제조
3. 게타바이러스 genome의 전체 염기서열 확인 및 다른 알파 바이러스와의 상동성 조사
4. 게타바이러스 genome을 이용하여 replicon 제조 및 helper plasmid 제조

5. 게타바이러스 replicon을 이용하여 reporter 유전자의 발현

6. 게타바이러스 replicon을 이용하여 동물용 항원의 발현 유도 및 발현 항원의 정제 그리고 정제 단백질의 characterization

IV. 연구 개발 결과 및 활용에 대한 건의

게타바이러스를 이용한 외래단백질발현 벡터시스템을 구축하기 위하여 본 연구진은 국내 돼지로부터 게타바이러스를 plaque purification 방법으로 순수 분리 분리된 게타바이러스를 진핵세포를 이용하여 배양한 후 배양 상층액으로부터 바이러스의 RNA를 추출하였고, 추출된 RNA로부터 RT-PCR을 수행하여 게타 바이러스 모든 genome의 염기 서열 분석을 실시 한 결과 3' 말단의 염기서열을 기존에 보고된 게타바이러스의 염기서열과 비교 분석하여 게타바이러스임을 확인하였고, 기존에 발표된 게타바이러스의 염기서열과 90%의 상동성을 보였다. 또한 기존에 발표된 게타바이러스의 염기서열을 토대로 하여 nsP1 부분과 E1 부분을 각각 증폭하여 염기서열을 확인하였다. 이 역시 nsP1의 경우 96%, E1의 경우 93%의 상동성을 보여주었다. 또한 다른 alphavirus에 속하는 Semliki Forest virus나 Sindbis virus, Sagiyama virus 등에서 conserve 한 것으로 보여지는 부분에 대한 search를 수행하여 다양한 부분에 대한 Getah virus와의 상동성을 조사한 결과 Sagiyama virus와는 약 97%의 상동성을 보였다.

그후 게타 RNA에서 cDNA를 합성하고 이것을 주형으로 Taq polymerase와 pfu polymerase를 이용하여 PCR을 실시하여 각각 5개, 6개의 PCR products를 얻어 pGEM-T Easy vector와 pBluescript II KS + 에 클로닝하였다. 그리고 nonstructural protein 부위를 coding하는 염기 서열 부위를 갖는 클론을 연결하여 게타바이러스 replicon을 제작하였고 structural protein 부위를 coding하는 염기 서열 부위를 갖는 클론을 연결하여 게타바이러스 helper를 제작하였다.

게타바이러스 replicon이 제대로 작동하는지 알아보기 위해 replicon의 multi cloning site에 reporter 유전자인 GFP gene을 삽입하였다. 그리고 in vitro RNA transcription을 한후 electroporation을 통해 mammalian cell인 BHK (Baby hamster kidney)-21 cell에 게타바이러스 replicon을 transfection 하여 GFP의 발현이 확인하였다. 앞으로는 in vitro에 mammalian 세포에서의 외래 유전자의 대량 발현을 확인 한 후 대상동물 및 실험동물에서의 캐리어 벡터로서의 항원발현 특성을 확인하여 백신 공학기술에 의한 생백신의 산업체로의 기술 이전 및 생산을 계획하고 있다.

S U M M A R Y

Getah virus is a member of the genus alphavirus of the Family Togaviridae, a group of enveloped viruses containing a single-strand positive RNA genome. The virus was first isolated by Elisberg and Buescher from *Culex gelidus* near Kuala Lumpur, Malaysia in 1955. A related strain, Sagiya virus, was isolated near Tokyo, Japan in 1956 by Scherer near Tokyo, Japan. And named after the town where it was discovered. Since then, several strains have been isolated, mostly in Japan and in Korea. Currently, Getah virus is classified under the Semliki Forest Complex group of alphaviruses.

The ecology of Getah virus is not yet fully understood. However, the groups of Marchette, Wada, Sentsui and Kono suggested that swines and horses serve as vertebrate hosts, with swines having the more important role as an amplifier of the virus. It is well established that the virus is transmitted between vertebrate hosts by a mosquito vector (*Culex* sp.). Evidence of fever, swelling of lymph nodes and some respiratory symptoms has been observed in infected horses. In humans, there is no disease associated with Getah virus infection. However, there have been reports of individuals possessing antibodies against Getah virus in the absence of antibodies against other alphaviruses, suggesting the possibility of asymptomatic infection in nature. This observation has led to studies on the potential applications of Getah virus as an expression vector and drug delivery system.

In recent years, the number of researches on the use of recombinant viruses for vaccination or therapeutics has increased tremendously. Several studies on baculovirus-, adenovirus- and rhabdovirus-based expression systems to display foreign antigen like HIV, HCV and Dengue are being done. Many of them demonstrated the ability of these recombinant viruses to allow recognition of the foreign antigen by host immune cells, thereby resulting in heightened, and probably protective, immune response against the pathogenic types. Other studies involve targeting and elimination of tumor cells using recombinant virus that has high tropism on the immortalized host cells. An example is the recombinant Sindbis virus, an alphavirus, in specifically infecting cancer cells in tumor-induced mice. Infection of these cells by the virus leads to cell death due to the cytotoxic gene that the virus carries. Another alphavirus, the Semliki-Forest Virus (SFV) that belongs to the same group as Getah virus, is being studied extensively for its potential use as a vaccine vectors for HIV and other important viruses and also as carrier of suicidal genes. Given the similarity of Getah virus and SFV, an expression system based on this virus was developed.

First, Getah virus was first isolated from a swine farm in Korea. This virus was propagated and purified by plaque assay using BHK-21 cells. Second, Viral genomic RNA was extracted from the purified virions and amplified by reverse-transcription, polymerase chain reaction (RT-PCR) using *Taq* and *Pfu* polymerase. To obtain the 5' and 3' non-translating regions (NTR) of the viral genome, the 5' /3' RACE (Rapid Amplification of cDNA Ends) methods were employed. Third, fragments of the double-stranded DNA version of the viral genome, ranging from 520 bases to 4 kilobase in length, were cloned into pGEM-T or pBluescript II. These cloned segments of the Getah virus was then sequenced using the dideoxy-dye termination method. Using a DNA analysis program, the full genome sequence of Getah virus was assembled then compared with that of Sagiyama strain (GenBank Accession Number AB032553), the only other strain whose full genome has been determined by the group of Shirako in Tokyo, Japan.

Sequence analysis of the nonstructural ORF between Getah and Sagiyama revealed a 97.9% similarity at the nucleotide level and a 99.3% similarity at the protein level. In the same manner, the structural ORF of Getah and Sagiyama were highly similar at the nucleotide level with 97.4%, while at the protein level, these two viruses have a high homology at 98.6%. The 5' NTR between the two viruses were identical. While at the 3' NTR, Getah and Sagiyama have a similarity index of 99.8% in the last 258 nucleotides. However, one notable difference between the two is the 102 base-deletion spanning the region nt 11331 to 11340, and nt 11350 to 11441 of Sagiyama. This region does not exist in Getah virus. Interestingly, this region in the Sagiyama strain contains the 3' terminal repeats, a characteristic that is typical of all alphaviruses. These 3' terminal repeats form secondary structures that are thought to play an important role in the synthesis of the complementary vRNA by nsP3 during viral RNA replication. The reason for the absence of this structural motif in Getah virus as well as its effect on viral RNA replication is still unknown. In general however, the sequence comparison supports the previous report that the Sagiyama strain and Getah virus are closely related.

The next goal was to construct an expression vector and a helper plasmid based on the nonstructural ORF and the structural ORF of Getah, respectively. For the expression vector, the nsP1-4 genes, including the 5' NTR, 3' NTR and the 26 subgenomic promoter region was assembled onto pSFV1 plasmid backbone, mediated by an SP6 promoter. A multi-cloning site was inserted between the nonstructural ORF and the 3' NTR, about 30 nucleotides downstream of the subgenomic promoter. As for the helper plasmid, the structural ORF, including a portion of nsP4 containing the 26 subgenomic promoter upstream and the 3' NTR downstream, was assembled into pGEM-T or pBluescript II backbone, mediated by T7 promoter. DNA sequencing was done at each step of the genome assembly to closely monitor any

mutation that may occur during the assembly process.

To test whether the Getah expression vector was functioning, two reporter genes, LacZ and IRES-GFP gene were inserted separately into the multi-cloning site of the expression vector. These constructs were transfected in BHK-21 cells by electroporation and tested for expression of the reporter genes after 24 hours. For the Getah helper plasmid, the construct was co-transfected with either pSFV-containing LacZ or IRES-GFP reporter gene. After 48 hours, the culture medium was harvested and centrifuged at 100,000 x g to recover mature virus-like particles (VLP) containing the SFV-Lac Z or SFV-IRES-GFP recombinant genome, if any. These recombinant VLPs are then inoculated into BHK-21 cells and tested for expression of the reporter genes after 24 hours.

Expression of the reporter genes LacZ and IRES-GFP were not observed in BHK-21 transfected with the Getah expression vector containing the LacZ or IRES-GFP reporter gene, respectively. This is in stark contrast to the SFV expression vector containing the same reporter genes, which clearly showed expression of the same reporter genes (20% of LacZ transfected cells and 90% of IRES-GFP transfected cells) in BHK-21, 24 hours after transfection. Similarly, the expression of the reporter genes were not observed in the cells inoculated with VLPs containing the recombinant SFV-LacZ or SFV-IRES-GFP genome in contrast to the positive expression of the same recombinant genomes using a pSFV helper plasmid.

To rule out the effect of the 102 base deletion on the failure of the Getah constructs to express the reporter genes, the 3' NTR was replaced with the SFV 3' NTR. Previous studies mentioned that the 3' UTR of togaviruses may be interchanged without having any detrimental effect on virus replication. However, in our study, the reporter genes were still not expressed BHK-21 transfected with Getah expression construct with modified 3' NTR. Also, the reporter genes were not expressed in BHK-21 inoculated with the Getah VLPs (source mRNA template contained the modified 3' NTR) carrying the recombinant SFV-LacZ or SFV-IRES-GFP genome in contrast with the positive expression of the reporter genes when pSFV helper was used.

These experiments led us to the conclusion that efficiency of Getah viral RNA replication is very low when the non-structural and structural ORFs are located on separate mRNA. This agrees with the group of Shirako et al. who demonstrated that the Sagiya viral genome was expressed efficiently when the full genome is located in *cis*. Therefore, a full genome clone of Getah virus may be required in order to utilize the virus' mechanism of gene expression.

C O N T E N T S

Chapter 1 Introduction

Chapter 2 State of Art

Chapter 3 Research Contents and Results

section 1 Getah virus RNA isolation and cDNA synthesis

section 2 Sequencing analysis of Getah virus genome and homology analysis with other alphaviruses

section 3 Construction of Getah replicon and helper plasmid

section 4 Confirmation of reporter gene expression in Getah replicon

section 5 Establishment of Getah carrier vector for a new animal vaccine

Chapter 4 Degree of research purpose achievement and its contribution to other further researches

Chapter 5 References

1. 연구개발과제의 개요

1. 연구 개발의 목적

다양한 유용성 단백질을 진핵세포에서 발현하는 것은 농업 특히 축산분야의 생산성을 향상시킬뿐만 아니라 고부가가치의 단백질 생산 및 정제를 통해 산업적으로 유용한 물질을 제조할 수 있는 기본 핵심기술 (platform technology)로 매우 중요하다. 현재 여러 가지 벡터를 이용하여 클론된 외래 유전자의 단백질 발현은 분자생물학 분야에서 중점적인 기술분야가 되고 있으며, 이러한 목적을 위해 많은 단백질 발현 시스템이 이용되고 있으나 이는 대부분 단핵세포(*E. coli* 종류 등)를 이용한 것이며 고등 진핵세포에서의 외래 단백질 발현시스템은 여전히 어려운 상황이다. 또한 바이러스 벡터를 이용한 발현시스템은 용도 및 효율면에서 매우 제한적이다. 이에 본 연구개발의 목적 및 대상기술은 병원성이 매우 낮은 돼지에서 분리된 Getah virus를 이용하여 진핵세포의 세포질내에서 외래 유용성 단백을 다양하고 고효율로 발현할 수 있는 신규(novel)체계를 구축하고자 한다.

Getah virus는 Togaviridae의 Alphavirus에 속하며 이에는 Sindbis virus, Semliki Forest virus 등 이미 상용화된 바이러스 벡터 시스템과 같은 군에 속한다. 이러한 Alphavirus를 이용한 단백질 발현 시스템은 숙주세포의 범위가 넓고, 외래단백질을 대량 생산할 수 있다는 잇점이 있으며, 최근에는 유전자 치료와 백신연구 등 많은 실험분야에 응용되고 있는 실정이다.

본사는 국내 돼지에서 분리한 게타바이러스를 이용하여 바이러스 유전자의 전체 염기서열을 관독하고 구조단백질을 제거한 replicon과 구조단백질만을 포함하고 있는 helper DNA를 제조하여 외래 단백질 발현 벡터 시스템을 구축하고, 또한 이를 이용하여 돼지가금류에서 외래유전자를 도입하여 발현시켜 단백질 발현능력을 검정하고 그 활용성을 단백질 발현은 물론 유용 유전자의 전달매체로써 확대할 수 있는 기술을 개발하고 제품화하고자 한다.

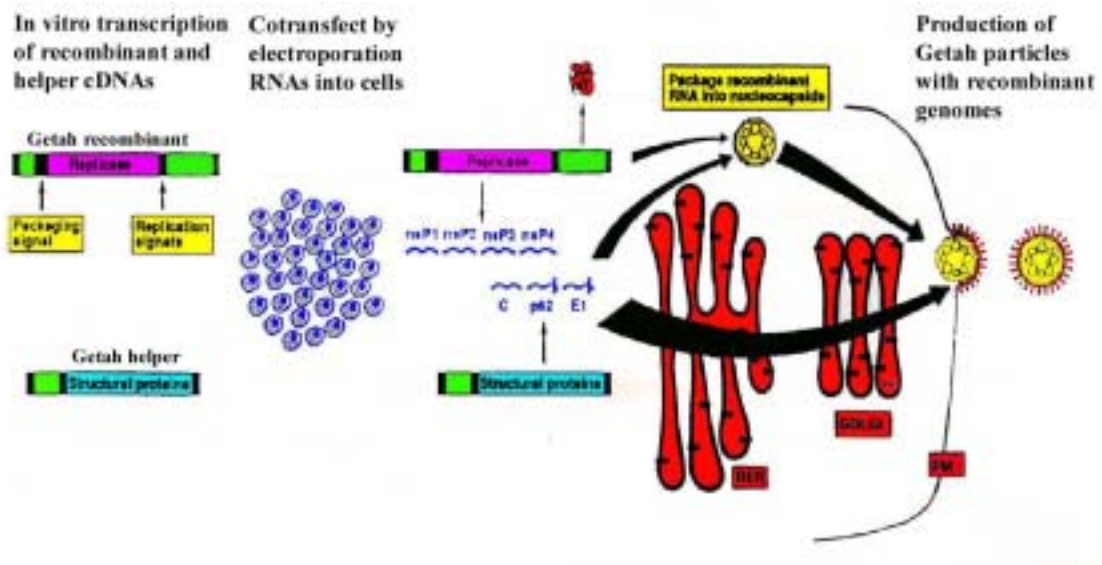
2. 필요성 및 범위

최근 생명공학 분야에 있어서 관심이 되고 있는 유전자 치료, 유전자 면역뿐만 아니라 여러 생백신에 이르기까지 원하는 외래유전자를 효율적으로 발현하는 시스템은 이들을 위한 전제조건이라 할 것이며, 결국 총체적인 생명공학 분야의 발전을 위해서 이러한 시스템은 필수요건이라 할 것이다. 따라서 효율적인 외래 단백질 발현 벡터 시스템은 많은 연구기관에서의 실험실내에서 원하는 결과를 얻기 위한 상용도구로써뿐만 아니라 새로운 최신 기술분야에 있어서의 보조자로서 그 이용가치는 현재의 가시적인 결과보다 무궁무진하다 할 것이다.

- 진핵세포 발현 바이러스 벡터 시스템 상용화
- 외래 유전자 발현을 통한 각각의 발현 특허 상용화

- 외래 유전자 전달 시스템 상용화
- 병원체 항원 발현을 통한 subunit 백신의 후보물질 생산
- 특히 단백질의 세포내 작용기전 규명화
- 바이러스 유사입자 개발을 통한 새로운 백신 후보물질 생산

○ Getah virus를 이용한 외래단백질 발현시스템 제조공정



2. 국내외 기술 개발 현황

현재 생명공학 분야에 있어서 기능성 단백질(기능성 효소, 호르몬, 백신사용 항원, 항체 및 치료용 단백질 등)을 대량으로 발현하는 것은 가장 중요한 핵심사항이라 할 수 있다. 이러한 단백질 발현 시스템으로는 현재 원핵세포나 효모에서의 단백질 발현 시스템이 개발되어 사용되고 있으나 진핵세포 유래의 단백질 발현을 위해서는 낮은 효율과 실험시 조작의 난이성 등으로 인해 여러 가지 어려움이 있으며, 또한 생산할 수 있다 해도 진핵세포에만 존재하는 post-translational modification이 다름으로 인해 생산된 단백질의 기능과 효능을 유지하는데 어려움이 있다. 이에 진핵세포에서는 여러 가지 바이러스를 이용한 발현체계가 개발되어 Vaccinia virus, Baculovirus 및 Adenovirus 등이 대표적으로 사용될 수 있다. 그러나 Baculovirus 벡터는 곤충세포에만 감염되어 외래단백질 생산에 유용하게 사용할 수 있으나 생성된 단백질의 glycosylation pattern의 차이로 인해 실질적으로 사용의 제약이 따르며, Vaccinia virus나 Adenovirus는 단백질 발현량이 적을 뿐 더러 homologous recombination을 이용하여 screening해야 하는 문제점 등 발현벡터 체계로는 문제가 있다. 이러한 점에서 Alphavirus는 광범위한 숙주세포, 높은 수준의 외래단백질 생산량 등 많은 유용한 특징을 가지고 있으며, 이중 Sindbis virus와 Semliki Forest virus는 이미 상용화된 벡터로서 근래에는 백신, 유전자 치료 분야에서도 연구되고 있다. 그러나 이들 바이러스를 이용한 벡터는 cytopathic effect가 강하고, Sindbis virus는 사용시 기술적인 어려움으로 인해 광범위하게 사용되고 있지는 않다. Getah virus는 Sindbis virus, Semliki Forest virus와 같은 Alphavirus에 속하며, 사람과 관련된 질병과는 관련이 없고, 말 등 약간의 동물에서 질병을 일으키는 것으로 알려지고 있으며, 돼지는 병변없이 바이러스 증폭자로서 중요한 역할을 하고 있다. 따라서 게타 바이러스를 이용하여 새롭고, 보다 간편한 외래단백질 발현 벡터 시스템을 구축함으로써 고등진핵생물에서 항암물질등 실험실에서 원하는 단백질 발현 실험을 상용화하고, 상대적으로 병변을 일으키지 않는 돼지나 가금 등에서 외래단백질을 유도하는 캐리어벡터로서 생백신으로 이용될 수 있을 것으로 보인다.

Alphavirus system은 Alphavirus의 비구조단백 부분과 multicloning site를 포함하는 플라스미드 벡터와 Alphavirus의 구조단백 부분을 포함하고 패키징 시그널이 없는 helper를 이용하여 defective infectious virus particle을 생성하여 원하는 단백질을 발현하는 등 다양한 패키징 시스템을 이용한 transient expression system과 이들의 염기서열의 일부에 변이를 가하여 CPE를 줄임으로써 진핵세포 내에서 지속적으로 외래단백질을 발현할 수 있는 cell line 구축, 백신분야, gene therapy 영역에서도 활발히 연구되고 있다.

Sindbis virus와 Semliki Forest virus는 최근 외래 단백질 발현 벡터 시스템으로 개발이 되어 여러 곳에서 상용화 되고 있으며, 현재에도 이를 이용한 새로운 stable packaging cell line 형성이나 백신, 유전자 치료 등에 관한 실험이 계속중이다. 본사도 Semliki Forest virus를 이용하여 많은 바이러스의 구조단백, 항암물질 등의 단백질 발현 실험을 하였으며, *in vitro*에서 RNA transcript를 생성할 필요 없이 DNA를 이용한 단백질 발현 시스템, 여러

외래 유전자를 동시에 발현하는 시스템 등을 개발하여 이미 특허출원 상태에 있다. 또한 Semliki Forest virus를 이용한 HIV의 gag, env 유전자를 함유한 벡터 제조 및 바이러스 유사입자 생산에 대한 특허를 출원(10- 2001-894)중이며, 여러 가지 유용성 단백질 생산 및 유전자 치료를 위한 특허를 출원할 예정이다. 그러나 계타바이러스에 대한 연구는 아직까지 국내외에서 거의 전무한 상태이며, 따라서 본사는 같은 알파바이러스에 속하는 Semliki Forest virus를 이용한 단백질 발현시스템을 연구한 경험과 자료 등을 바탕으로 보다 편리하고 유용한 시스템을 만들고자 한다.

제 3 장. 연구수행 내용 및 결과

1. 바이러스의 분리

게타바이러스를 이용한 외래단백질발현 벡터시스템을 구축하기 위하여 본 연구진은 국내 돼지로부터 게타바이러스를 plaque purification 방법으로 순수 분리하였다(Fig. 1).



Fig. 1. Plaque purification of Getah virus

2. 전체 염기서열의 분석

국내돼지에서 게타 바이러스를 plaque purification의 방법으로 순수분리하여 분리된 바이러스를 진핵세포에 배양 후 배양 상층액으로부터 Qiagen viral RNA extraction kit를 이용하여 viral RNA를 추출하였다. 추출된 RNA를 주형으로 RT-PCR을 수행하여 각각의 fragment의 염기서열을 분석하였다. RT-PCR은 두가지 방법으로 수행하였다. 우선 첫 번째로, Invitrogen사의 SuperscriptII를 이용하여 42℃에서 50분간 reverse transcription을 수행하여 cDNA를 만든후, 이를 주형으로 하여 PCR을 수행하였고, 둘째로는 Intron사의 One-step RT-PCR kit를 사용하여 RT-PCR을 수행하였다. 각각에 사용된 primer는 표1(1~31)에 기술하였다. 증폭된 PCR band는 pGEM-T vector에 클로닝 하였고, T7 및 SP6 promoter sequencing primer를 이용하여 게타 바이러스 전체의 염기서열을 분석하였다(Fig. 8). 염기서열은 automatic sequencer를 이용하여 분석하였으며, NCBI의 blast search를 통하여 기존에 발표된 게타바이러스 및 다른 Alphavirus의 염기서열과의 상동성을 비교하였다. 우리가

분리한 게타바이러스의 3' UTR 염기서열을 기존에 보고된 게타바이러스의 염기서열과 비교 분석하여 게타바이러스임을 확인하였고(Fig. 2), 기존에 발표된 게타바이러스의 염기서열과 90%의 상동성을 보였다. 또한 이미 발표된 게타바이러스의 염기서열을 토대로 하여 nsP1 부분(Fig. 3)과 E1 부분(Fig. 4)을 각각 증폭하여 염기서열을 확인하였다. 이 역시 nsP1의 경우 96%, E1의 경우 93%의 상동성을 보여주었다. 또한 특히 alphavirus에 속하는 Semliki Forest virus, Sindbis virus, Sagiyama virus, Aura, VEE 그리고 EEE 등에서 nonstructural proteins 인 nsp 1,2,3,4 와 structural proteins 인 capsid ,E2 의 염기서열의 상동성을 비교한 결과 앞서 보고되어있는 Sagiyama virus와 97% 이상의 상동성이 있음을 phylogenetic trees 로 나타내었다. (Fig. 5, 6). Fig. 10 에 검은색 화살표로 표현된 부분이 염기서열을 분석하기 위한 각각의 분절을 나타내주고 있다. 게타바이러스의 전체 염기서열 분석 결과, 1~78까지와 11289~11528까지는 nontranslated region, 79~7482까지는 nonstructural protein 으로 nsP1234를 coding하는 것으로 보이며, 기존의 alphavirus의 특징인 nsP3부분에 stop codon이 있어 79~5638까지 nsP123로 끊어지기도 한다. 또한 7587~11288까지는 structural protein인 C-E3-E2-6K-E1을 coding하는 것으로 보인다(Fig. 8, 9).

```

Query: 55  cccctgaatagtaacaagtacataaaaaaccaaacagtagttcaaagggctatacaacc 114
          |||
Sbjct: 155  cccctgaatagtaacaaaaacataaaaaactaaaacagtagttcaaagggctatacaacc 214

Query: 115  cctgaatagtaacaaaatacagaannccataactaaattataaaaattaactaatcagatc 174
          |||
Sbjct: 215  cctgaatagtaacaaaacacagaaaaaccata--aaaatcttaaaataaataaatctgatc 272

Query: 175  atct-aaatgactaattggaaatagccgaactctacggagatgtaggcgtccgaactc 233
          |||
Sbjct: 273  atctaaaatgactaattggaaatagccgaactctacggagacgtaggcgtccgaactc 332

Query: 234  cacggagacgta 245
          |||
Sbjct: 333  cacggagacgta 344

```

Fig. 2. Nucleotide sequence analysis of 3' UTR. Identities : 174/192 (90%)

```

Query: 2    aaactgattagcaagaggttccaacaggcgtcaccatcctggacgtgggtagtgcacc 61
          |||
Sbjct: 1    aaactgattgagcaagaggttccaacaggcgtcaccatcctggatgtgggtagtgcacc 60

Query: 62  gcaaggaggttgatgtctgaccacacctaccactgcatctgcccatgaaaagtgcggaa 121
          |||
Sbjct: 61  gcaaggaggttgatgtctgaccacacctaccactgcatctgccctatgaaaagtgcggaa 120

```

```

Query: 122 gacccagagaggtggcgaattacgctcгааagctggcгааagcatcggggactgtgcta 181
          |||
Sbjct: 121 gacccagagaggtggcgaattacgctcгааagctggcгааagcatcggggactgtgcta 180

Query: 182 gacaagaatgtgtccgгааagataacggacctacaagacgtcatggccactccagacttg 241
          |||
Sbjct: 181 gacaagaatgtgtccgгааagataacggaccttcaagacgtcatggccactccagacttg 240

Query: 242 gaatccccgactttttgcctgcatactgacgagacgtgccgcactagggctgaggtcgcc 301
          |||
Sbjct: 241 gaatccccgactttttgcctgcacaccgacgagacgtgccgcactagggctgaggtcgct 300

Query: 302 gtgtaccaggacgtatagcgtgtgcacgcaccgacgtcactgtaccaccaggccatcaaa 361
          |||
Sbjct: 301 gtgtaccaggacgtatagcgtgtgcacgcaccgacgtctctgtaccaccaggcaattaa 360

Query: 362 ggtgtaaggacggcgactgg 382
          |||
Sbjct: 361 ggtgtaaggacggcgactgg 381

```

Fig. 3. Nucleotide sequence analysis of nsP1 portion. Identities : 369/381 (96%)

```

Query: 2 actgcagcagtgctcagccccgagtcggccggctaccgctgcacccatgtcattgcc 61
          |||
Sbjct: 1005 actgcagcagtgctcagccccggaatccgccagctaccgctgcacccatgtcattgcc 946

Query: 62 gtgccagacatgtcaggaaaaacttggttgttatggctcgccccataagggacgatgtgg 121
          |||
Sbjct: 945 gtaccagacatgtctggaaaaacttggttgttatggctcgctccgtaagggacgatatga 886

Query: 122 tccttcggcggctcacacgctgccatgcatgtcgttttggcactgcacacagataaccttg 181
          |||
Sbjct: 885 tccttcggcggctcacacgctgccgtgcatgtcgttttggcactgcacacagataaccttg 826

Query: 182 aatgccggggatgctgatgctgtagagaaatgcagggttatcttgccatctgttttgatg 241
          |||
Sbjct: 825 aatgccggggaagctgatgctgtagagaaatgcagggttatcttgccatctgwtgtgatg 766

Query: 242 tccacagctgcctcctgtatggtagctacgttcgagtgagagwggacagcacatcttccg 301
          |||
Sbjct: 765 tccacagctgcctcctgtatggtagctacgttcgagtgagagwggacagcacatcttccg 706

Query: 302 ggtttgtcagttttgaaagtacagatcgcgatcccgccgaagtccgatgagtgcgatgcag 361
          |||
Sbjct: 705 ggtttgtcagttttgaaagtacagatcgcaatcccgccgaagtccgatgagtgcgatgcaa 646

Query: 362 accgccacttggcactccaggtttgtgacggcaggtgcatcaatcacgcgctaacgcg 421
          |||
Sbjct: 645 accgccacttggcactccaggtttgtgacggcaggtgctcaatcacgcgctaacgcg 586

Query: 422 gtgtccgggatgtccatggagactgggatgttgccaacggcgcaatcttctgctctgact 481
          |||
Sbjct: 585 gagtccggaatgtccatggagactgggatgttgccaacggcacaatcttctgctctgact 526

Query: 482 gggttggtcttgattacgcatacгааaggagccttgtcattcagcgacgtgcctctctct 541
          |||
Sbjct: 525 gggttggtcttgattatgcaccaaaggagccttgtcattcagcgatgtgcctctctct 466

```

```

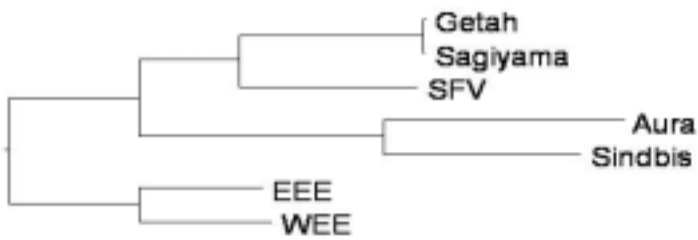
Query: 542 tttatccagaacttaagccagaaggggtctgtgtgtaaggcacgtgaacagtaccggac 601
          ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 465 ttaatccagtacttaagccggaaggggtctgtgtataaggcacatgaacagtaccggac 406

Query: 602 gaaggtcttgacaacttgagggcgggtgtggcatacaggctcttgctctctaccgtcctg 661
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct: 405 gaaggtcttgacaacttgagggcgggtatggcatacaggctcttgctctctaccgtcctg 346

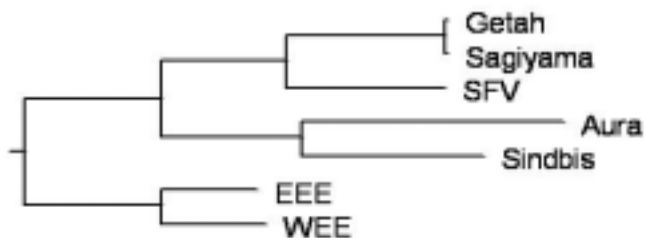
Query: 662 ctctggatgtctcc 675
          ||||| ||||| |||||
Sbjct: 345 ctctggatgtctcc 332

```

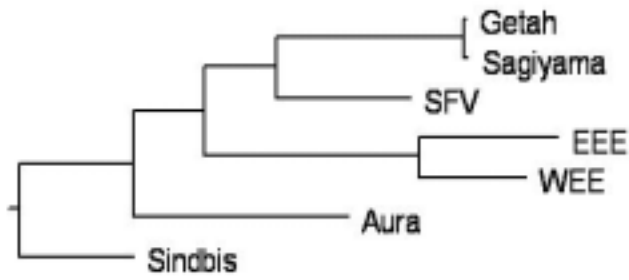
Fig. 4. Nucleotide sequence analysis of E1 partial sequence. Identities : 633/674 (93%)



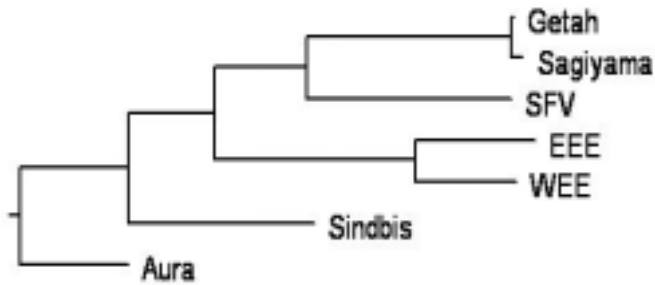
nsp1



nsp2

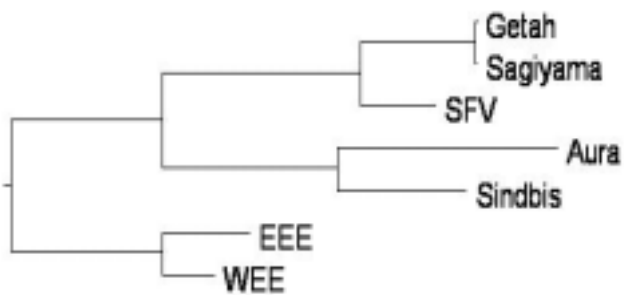


nsp3

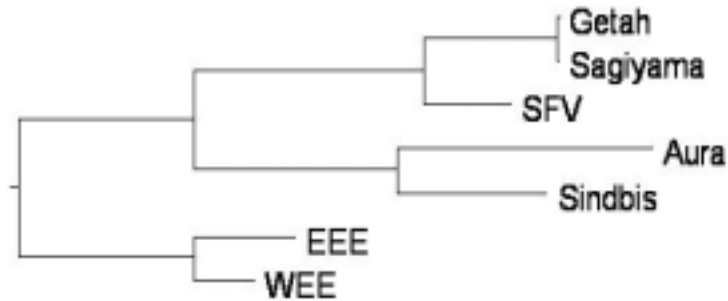


nsp4

Fig. 5. Alphavirus phylogenetic trees based on nonstructural proteins (nsp 1~4).



capsid



E2

Fig. 6. Alphavirus phylogenetic trees based on structural proteins (c, E2).

3. Full clone

게타바이러스를 이용한 벡터 시스템을 확립하기 위한 replicon 및 helper를 제조하기 위한 전과정으로 게타바이러스의 full clone 작업을 수행하였다. 우선 pGEM-T vector를 self-ligation한 뒤, *Mlu*I으로 cutting하여 linearize 한다. Multicloning site(*Mlu*I-*Eco*RI-*Bgl*II-*Xho*I-*Xba*I)를 만들기 위하여 primer 32와 33을 98°C에서 10분간 denaturation을 실시하여 4°C에서 급속하게 cooling한 뒤, 37°C에서 10분간 anealing을 수행하였다 (Fig. 7). 이렇게 만들어진 multicloning site를 미리 준비해둔 vector와 ligation하여 벡터를 준비하였다. 게타바이러스의 크기가 10kb가 넘는 크기이므로 한번에 클로닝을 하는데 무리가있어 , 분석된 전체 염기서열을 기준으로 하여 one-site로 cutting 되는

표 1. List of primers

1	1F	5'-ATGGCGGACGTGTGACAT-3'
2	3'endR	5'-GTAAAATATTAAAAAAC-3'
3	98F	5'-TTGAGGCTGACAGCCATT-3'
4	541F	5'-GTGCACGCACCGACGT-3'
5	1387F	5'-CAGACTATCGTCAAAGTGC-3'
6	1445R	5'-TGACCACAGGCTCGGTAT-3'
7	2424F	5'-CCTTTACGTGATGAGGCGTT-3'
8	2500F	5'-GTCGCCACACAGTACTACCT-3'
9	3279F	5'-ATTCTCAGCGCAGTCAGTT-3'
10	3359R	5'-TTGAACCCATACATGCGTCC-3'
11	3420R	5'-CGAGTTCATGTTGCCACGTAG-3'
12	3740F	5'-TGAGTACCGGCAACCACT-3'
13	4460R	5'-AATAAGTGGTTCAAGGACTG-3'
14	4650F	5'-CAGTGCAACTGACGGTA-3'
15	4886F	5'-AGCGAGTTGCCAGACTTAGG-3'
16	5036R	5'-GGACTTACTAGTGACG-3'
17	6040F	5'-AATTATCCGACCGTGGCG-3'
18	6143R	5'-GGGCAGAACGTAGCTCTGTC-3'
19	7030F	5'-GCCTTCATAGGTGATGACAA-3'
20	7283R	5'-CTGTCTTCATCCTGGTTGTC-3'
21	7856F	5'-ACGAATGTGCATGAAGATAGA-3'
22	8030R	5'-TCATACTTGCTCGATTTCTT-3'
23	8525F	5'-GTGACGGAACACTTCAATGT-3'
24	8689R	5'-ATGCCAATTTGCGCTGCGACC-3'
25	8821F	5'-TCATCGTGGCCTACTGCCCT-3'
26	9500F	5'-TFFFATCGAGTACAGATGGG-3'
27	9573R	5'-CCATGGGGTTTGCCTTCAGT-3'
28	10476F	5'-GCCTGGACGCCTTTCGACAAC-3'
29	10617R	5'-TGGCATAACAGGTCTTTGTC-3'
30	11080F	5'-TATCTGTGTGCAGTGCCAAA-3'
31	11204R	5'-CTGCACCCATGTCATTGCCGT-3'
32	Mlu-S	5'-CGCGTGAGAATTCGAAGATCTGACTCGAGGATCTAGAT-3'
33	Mlu-AS	5'-CGCGATCTAGATCCTCGAGTCAGATCTTCGAATTCTCA-3'
34	1S-MluI	5'-GAGACGCGTATGGCGGACGTGTGACATCACCC-3'
35	1981S	5'-CGTAAGCTACACCACATAGC-3'
36	8400S	5'-GTTGCTATGAAAAACAACCA-3'
37	End-XbaI	5'-GGCTCTAGAT ₍₄₃₎ GTAAAATATT-3'

restriction enzyme site를 찾아서 그를 응용하여 full clone을 잡고자 하였다. Multicloning site는 염기서열 분석결과, SP6 promoter의 downstream으로 삽입이 되었다. 만들어진 벡터에 게타바이러스를 네 개의 fragment로 연결하여 클로닝을 실시하였다. 클로닝 시에는 2167에 위치하는 *EcoRI*과 4760에 위치하는 *BglIII*, 그리고 8478에 있는 *XhoI* site를 각각 이용하였다. 첫 번째 fragment는 1~2500, 두 번째는 1981~5036, 세 번째는 4650~8689, 마지막

은 8400~끝까지로, 마지막에 43개의 poly-A를 삽입하여 주었다(Fig. 10). 각각의 네 개의 분절(Fig. 10. A, B, C, D)을 만들기 위하여 overlapping PCR을 통하여 PCR을 수행하였고, full clone 작업을 수행하였다. 그러나 결과적으로 이 stratege는 잦은 PCR 반응으로 인한 바이러스 gene의 mutation이 발생할 가능성이 충분히 제기되어 다른 방향에서 이 project에 접근을 시도 하였다.

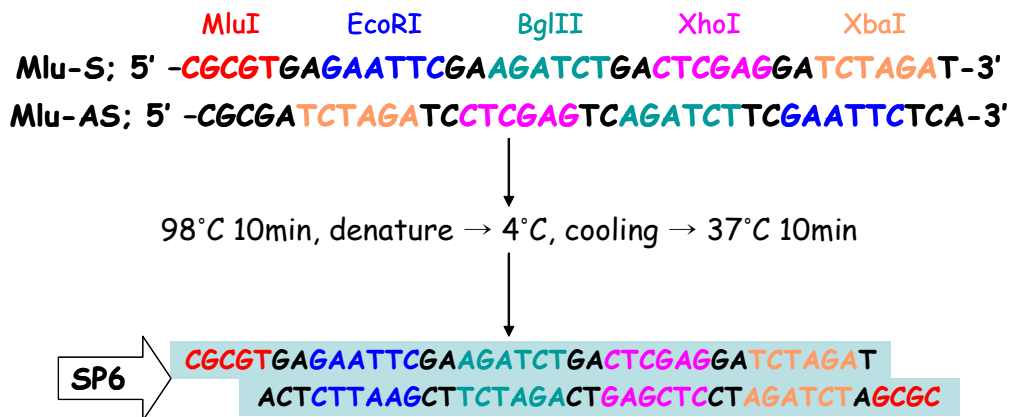


Fig. 7. Formation of multicloning site.

atggcggacgtgtgacatcaccttcgctctttctaggatcctttgctactccacatagt 60
gagagacaaacaaccccaaatgaaggtaacctggacgttgaggctgatagccattcctt 120
M K V T V D V E A D S P F L
aaggcccttcagaaggcgtttcccgctttgaggttgaatcacagcaggtcacaccgaat 180
K A L Q K A F P A F E V E S Q Q V T P N
gaccatgctaacgctagagcattttcgcatctggctactaaactgattgagcaagaggtt 240
D H A N A R A F S H L A T K L I E Q E V
ccaacaggcgtcaccatcctggacgtgggtagtgcaccgcgaaggaggttgatgtctgac 300
P T G V T I L D V G S A P A R R L M S D
cacacctaccctgcatctgccccatgaaaagtgcggaagaccagagaggtggcgaat 360
H T Y H C I C P M K S A E D P E R L A N
tacgtcgaagctggcgaagcatcggggactgtgctagacaagaatgtgtccggaag 420
Y A R K L A K A S G T V L D K N V S G K
ataacggacctacaagcgtcatggccactccagacttggaaatccccgactttttgcctg 480
I T D L Q D V M A T P D L E S P T F C L
catactgacgagacgtgccgactagggctgaggtcgctgtgtaccaggacgtatagct 540
H T D E T C R T R A E V A V Y Q D V Y A
gtgcacgcaccgacgtcactgtaccaccaggccatcaaaggtgtcaggacggcgtattgg 600
V H A P T S L Y H Q A I K G V R T A Y W
attggattcgacaccactccattcatgttcgagggcactagcgggcgctaccctgcttac 660
I G F D T T P F M F E A L A G A Y P A Y
tctaccaactgggcagatgatcaagtgtgcaggctcgtaacatcgccctgtgcgcgaca 720
S T N W A D D Q V L Q A R N I G L C A T
ggcctctccgaggggctcgccgcaaacctctctatcatgagaagaagtgtttgcgaccg 780
G L S E G R R G K L S I M R K K C L R P
agcgatagagtaatgttttcggctcgggtccaccttgtacaccgagagccgaaagctgtg 840
S D R V M F S V G S T L Y T E S R K L L
cgcagctggcatttaccttccgtgtttcacctgaaggtaagaacagttttacctgcagg 900
R S W H L P S V F H L K G K N S F T C R
tgcgacacgggtggtgtcatgcaaggttacgtggtaagaagatcaccaataagcccgggc 960
C D T V V S C E G Y V V K K I T I S P G
atataaggaaaaacagtcgattacgcagttacccatcacgcagaggggtttcctgatgtgt 1020
I Y G K T V D Y A V T H H A E G F L M C
aagatcactgatacagtcagaggagaaagagtcctttcccggtctgtacctatgtgcct 1080
K I T D T V R G E R V S F P V C T Y V P
gcaaccatatagcgaccagatgacgggtatacttgccactgacgtgacaccagaggatgcc 1140
A T I C D Q M T G I L A T D V T P E D A
cagaagctcctggttggattgaaccaacgcatagtggtgaacggtaggacgcaaagaaac 1200
Q K L L V G L N Q R I V V N G R T Q R N
acaaacacaatgaaaaactacctactgccagtggtagcgcaagcattcagtaaatgggca 1260
T N T M K N Y L L P V V A Q A F S K W A
cgagaggcacgcgacagacatggaggacgaaaaaccctaggcaccagagaacgcacgttg 1320
R E A R A D M E D E K P L G T R E R T L
acgtgtgttgccctgtggcggtttaaagccacaaaatccacaccatgtataagcggcct 1380
T C C C L W A F K S H K I H T M Y K R P
gaaacgcagactatcgtcaaagtccttccacttttgactcctttgtgataccgagcctg 1440
E T Q T I V K V P S T F D S F V I P S L
tggtcatccagcttttccatgggtatcagacagaggatcaaattgctactcagcgcaaga 1500
W S S S L S M G I R Q R I K L L L S A R
atggcccaaggcctaccatactcaggagaccgcactgaagctcgcgcgccagaagaagaa 1560
M A Q G L P Y S G D R T E A R A A E E E
gagaaggaggtgcaggaggtgaacttacgagggctgcgctgccaccgctagtgagcggc 1620
E K E V Q E A E L T R A A L P P L V S G

tcttgtgctgacgatatcgcccaggtggatgtagaggaactaaccttcagagccggagcc 1680
S C A D D I A Q V D V E E L T F R A G A
gggggtgtggaacaccagggaatgccctgaaggttacaccgcaagcacacgaccatctc 1740
G V V E T P R N A L K V T P Q A H D H L
atagctcctacttgatcctttcccccaaacgggtgtgaaaagcgagaagctggcacc 1800
I G S Y L I L S P Q T V L K S E K L A P
atccaccctcttgctgagcaagtacgggtcatgaccactctggaagatccggcagatac 1860
I H P L A E Q V T V M T H S G R S G R Y
ccagtcgacaagtacgacggacgggtattgatcccaacaggagcagccatcccagtgagt 1920
P V D K Y D G R V L I P T G A A I P V S
gagttccaggcactcagcgagagcgcaaccatgggtgacaatgagaggaatttataaat 1980
E F Q A L S E S A T M V Y N E R E F I N
cgcaagctacaccacatagcgctatccggccagccttgaataccgacgaggaaagctac 2040
R K L H H I A L Y G P A L N T D E E S Y
gaaaaagtgagagctgaaagggcagagacagagtagtggttcgacgtggacaagaaggca 2100
E K V R A E R A E T E Y V F D V D K K A
tgtatcaagaaggaggagcatcaggccttgggttaacaggggacctaatcaatccacct 2160
C I K K E E A S G L V L T G D L I N P P
ttccacgaattcgcatagcaaggactcaagatccgcccagcagccccgtaccacacgacg 2220
F H E F A Y E G L K I R P A A P Y H T T
atcatgggtgtgtttggcgttccgggttcgggcaagtgggtatcattagaacatgggtg 2280
I I G V F G V P G S G K S A I I K N M V
acgactcgcgatctgggtggccagtgaaagaaggagaactgccaagagatcatgaatgat 2340
T T R D L V A S G K K E N C Q E I M N D
gtaaagaggcaacgggggttgacgtaaccgctaggaccgtcgactcaatcttactgaat 2400
V K R Q R G L D V T A R T V D S I L L N
gggtgcaagaaggcgtagaaaacctttacgtcgatgaggcgttcgctgtcactcgggt 2460
G C K K G V E N L Y V D E A F A C H S G
actttgctagcgctcatcgcgctggtagaccgtcaggttaaggtagtactgtcggcgac 2520
T L L A L I A L V R P S G K V V L C G D
cctaagcagtggtttcttcaatttgatgcaactgaaggtgcactataaccacaacatt 2580
P K Q C G F F N L M Q L K V H Y N H N I
tgtacaagggtgctccataagagcatctccagaagatgcactctacctgttacggcgatc 2640
C T R V L H K S I S R R C T L P V T A I
gtgtccaccttgactaccaaggaagatgagaacgacgaaccgatgcaacaccctatt 2700
V S T L H Y Q G K M R T T N R C N T P I
cagattgacaccaccggttctccaaccagcctcaggagatcgtgttaacgtgcttc 2760
Q I D T T G S S K P A S G D I V L T C F
cgcggtgggtgaagcaactgcaaatcgactatcgtggacacgaggtgatgaccgcagct 2820
R G W V K Q L Q I D Y R G H E V M T A A
gcttcccagggtctgacaaggaaaggcgttacgccgtgagacagaaagtgaacgaaaac 2880
A S Q G L T R K G V Y A V R Q K V N E N
ccactgtactcacctctgtcgggacacgtcaacgtgctgttgaccgaactgaaaaccga 2940
P L Y S P L S E H V N V L L T R T E N R
ctgggtggaagacactgtcgggtgaccctggataaagggtgtaaccaatgttccacgt 3000
L V W K T L S G D P W I K V L T N V P R
ggggatttcagtgcaactctggaggaatggcatgaagaacatgacgggtatcatgagagtg 3060
G D F S A T L E E W H E E H D G I M R V
ttgaacgagcgaccggcgagggtgatccattccaaaacaaggctaagggtgtgctgggca 3120
L N E R P A E V D P F Q N K A K V C W A
aatgtctgggtgcaagtcttgagacggccggaatcgtatgacggcagatgaatggaac 3180
K C L V Q V L E T A G I R M T A D E W N
accatcttggctttcagagaggacagagcgtactaccagaagtcgctctcaacgagatt 3240
T I L A F R E D R A Y S P E V A L N E I

tgcactcgttactacggcgttgacctagacagcggcctattctcagcgcagtcggtttcc 3300
 C T R Y Y G V D L D S G L F S A Q S V S
 ctcttttatgagaacaaccactgggacaacaggcctggaggacgcatgtacgggttcaac 3360
 L F Y E N N H W D N R P G G R M Y G F N
 catgaagttagccaggaatacgcagccaggtttccatttctacgtggcaacatgaactcg 3420
 H E V A R K Y A A R F P F L R G N M N S
 gggctacaactaaacgtccctgagaggaagctccagcccttagcgctgaatgcaacata 3480
 G L Q L N V P E R K L Q P F S A E C N I
 gtcccatccaatcgtcggttaccgcaagctctgggtcacaagttatcagcagtgctcgtggg 3540
 V P S N R R L P Q A L V T S Y Q Q C R G
 gagagggtagagtgggtgctgaaaagattccaggtcaccaaatgttacttgtaagtgag 3600
 E R V E W L L K K I P G H Q M L L V S E
 tacaacctgggtgatacctcacaaaagagcttctggattgcacctccgcggtgtcaggg 3660
 Y N L V I P H K R V F W I A P P R V S G
 ggggaccgcagctacgacttggacctagggttacctatggatgcaggccgttacgacctg 3720
 A D R T Y D L D L G L P M D A G R Y D L
 gtattcgtcaacatccatactgagtaccggcaacaccactaccaacaatgcgtcgaccat 3780
 V F V N I H T E Y R Q H H Y Q Q C V D H
 tcaatgcgcctgcagatgctgggaggggattcactacacctgcttagaccaggaggctcg 3840
 S M R L Q M L G G D S L H L L R P G G S
 ctgctgatgagagcataatggttacgcagacagagtacggagatgggtggtagacgccctg 3900
 L L M R A Y G Y A D R V S E M V V T A L
 gctagaaaattctcggcgttccgtgtcctgagaccggcgcgtgtgacgagcaacacagaa 3960
 A R K F S A F R V L R P A R V T S N T E
 gtgttctcgtgttttctaactttgataacggcagaagagcggtaaccttgcaccaagct 4020
 V F L L F S N F D N G R R A V T L H Q A
 aaccgaaaacttagctcaatgtatgcctgcaacggattgcacactgctggttgtgcaccg 4080
 N Q K L S S M Y A C N G L H T A G C A P
 tcatacaggggtccgcccgcagatatacaggacacagtgaggaagcggctgtaaatgct 4140
 S Y R V R R A D I S G H S E E A V V N A
 gccaatgccaaaggtagcgtgagcgcaggagtgtagggcggctcgctaagaagtggcca 4200
 A N A K G T V S D G V C R A V A K K W P
 tcatctttcaaggggctgcaactccagtcggcacagccaaaatgatccgcgagatggc 4260
 S S F K G A A T P V G T A K M I R A D G
 atgaccgtaatccacgcagtgaggacaaaacttctccaccgtaacagaagccgaaggggac 4320
 M T V I H A V G P N F S T V T E A E G D
 agagagctagcggccgctatcgagctgtggctagcataatagtagcaacaacataaag 4380
 R E L A A A Y R A V A S I I S T N N I K
 agcgtcgcagtagcgtgctgtccacaggcaccttttccggcggtaaggacaaagtgatg 4440
 S V A V P L L S T G T F S G G K D K V M
 cagtccttgaaccacttattcacggcactggacgcaaccgatgcagacgtggttatctac 4500
 Q S L N H L F T A L D A T D A D V V I Y
 tgcagagataaaaactgggaaaagaagattcaggaagccatcgacaggcggacggcaatc 4560
 C R D K N W E K K I Q E A I D R R T A I
 gagctcgtatctgaagacgtgaccttggaaaccgatctggttagggtacaccggacagt 4620
 E L V S E D V T L E T D L V R V H P D S
 tgcttagtcggcagaaaatggttacagtgcaactgacggtaactgtactcctacctgag 4680
 C L V G R N G Y S A T D G K L Y S Y L E
 ggcacagggttccaccagcggcggctgcacatggctgaaatatcaactttatggccaaga 4740
 G T R F H Q T A V D M A E I S T L W P R
 ctccaagatgctaacgagcagatctgcctgtacgccctaggggagacgatggacagcata 4800
 L Q D A N E Q I C L Y A L G E T M D S I
 cgcactaaatgccagtagaagacgccgattcgtctacggcgggaaaacggtagcgtgt 4860
 R T K C P V E D A D S S T P P K T V P C

ctatgtcggtagcggatgaccgaggagggttgccagacttaggatgaataaacaccaa 4920
L C R Y A M T A E R V A R L R M N N T K
aacatcatcgtgtgctcctcctttccattaccgaagtacaggatagaaggcgtgcagaag 4980
N I I V C S S F P L P K Y R I E G V Q K
gtgaagtgtgaccgagtgctaatTTTTgaccagaccgtcccgtcactagtaagtcccaga 5040
V K C D R V L I F D Q T V P S L V S P R
aagtacatacagcagccgccggaacagctggataatgtgagcctgacttctacgacgtcg 5100
K Y I Q Q P P E Q L D N V S L T S T T S
acgggatccgcatggtcatTTccgtcggaacagacctacgaaacctggaagtcgtagcc 5160
T G S A W S F P S E T T Y E T M E V V A
gaggtacacaccgaacctccaatccctccgcctcgccgacgtagagcagccgtcgccaa 5220
E V H T E P P I P P P R R R R A A V A Q
cttagacaggatctggaagtaccgaggagatcgagccgtacgtgacacagcaagcagag 5280
L R Q D L E V T E E I E P Y V T Q Q A E
atcatggtcatggagagggcgcgacgacagacatacgcgctatcccagtcccggcacgg 5340
I M V M E R V A T T D I R A I P V P A R
cgggccattacaatgccagtcaccgccccagggttcgtaaggctgctactgaacctcca 5400
R A I T M P V P A P R V R K V A T E P P
ttagaacgggaagctcctatcccggcaccaagagaagaaccactagcacctcacct 5460
L E P E A P I P A P R K R R T T S T S P
ccgcataaccccgaggatttcgctcccagggtacctgtgagttaccgtgggagccggaa 5520
P H N P E D F V P R V P V E L P W E P E
gacctagacatccaatTCGGTgactTGGAGccacgcccgggaacaccagggaccgagat 5580
D L D I Q F G D L E P R R R N T R D R D
gtcagcacaggaatacagttcggtagacatcgactTTAaccagtcctgactaggcagggt 5640
V S T G I Q F G D I D F N Q S * L G R A
ggcgcgtatactTTTTcgtctgacactggccccgggtcacctacaacagaagtccgtaagg 5700
G A Y I F S S D T G P G H L Q Q K S V R
caacatgaattgccatgcgagactctgtacgccatgaagacgaacgcataataccgccc 5760
Q H E L P C E T L Y A H E D E R I Y P P
gcattTgacggagagaaaagaaaggTactccaggcaaatgacagatggccccgacagaa 5820
A F D G E K E K V L Q A K M Q M A P T E
gccaataagagcaggtaccagtcgaggaaagttagagaacatgaaggcatTaatTtagaa 5880
A N K S R Y Q S R K V E N M K A L I V E
agactaccggaaggagcaagTgtacctccatgagcaaacgacaaagtaccacgtac 5940
R L R E G A K L Y L H E Q T D K V P T Y
accagcaagtaccctaggcctgtgtactcaccatcggTggatgacagcctgagcgatccg 6000
T S K Y P R P V Y S P S V D D S L S D P
gaagtggctgtggccgctgtaactcttcttagaggagaatTatccgaccgtggcgaac 6060
E V A V A A C N S F L E E N Y P T V A N
taccagataaccgatgagTatgacgctTatctggactTggtcgacggctctgaaagctgc 6120
Y Q I T D E Y D A Y L D L V D G S E S C
ctcgacagagctacgtTctgccccgcaactaagatgtTaccctaagcaccacgcatac 6180
L D R A T F C P A K L R C Y P K H H A Y
caccaaccacaaatcaggagcgcagTacctTcccctTTTcaaacacgtTacaaaacgtg 6240
H Q P Q I R S A V P S P F Q N T L Q N V
ctagccgcccactaaaagaaatTgcaacgtcacccaaatgagagaatTaccaacctg 6300
L A A A T K R N C N V T Q M R E L P T M
gactTcgggtgtTcaacgtagaaagctTcaaaaaatcgcagTgTaccggcgaatattgg 6360
D S A V F N V E S F K K Y A C T G E Y W
caagaattTaaagacaatcctatacggatcaccaccgaaaacataacgacgtacgtggct 6420
Q E F K D N P I R I T T E N I T T Y V A
aaactcaagggtccaaaggctgctgccctTTTgccaagacgcataaacTggtgccgctt 6480
K L K G P K A A A L F A K T H N L V P L

caggaggtgccaatggaccgcttcgtgatggatatgaagagagatgtgaaagttacacca 6540
Q E V P M D R F V M D M K R D V K V T P
ggcaccaagcataccgaagaaaggccaaaagtgaagtaattcaagcggcgaaccactg 6600
G T K H T E E R P K V Q V I Q A A E P L
gccacggcatatattatcggaatccacagagagttagttagcggcgtaaaagccgttctg 6660
A T A Y L C G I H R E L V R R L K A V L
accccgaaacattcatactctgtttgacatgtcggcggaggactttgatgccatcatagcg 6720
T P N I H T L F D M S A E D F D A I I A
gcacatttccaaccgggagatgctgtactggaggcagatatcgcatccttcgacaagagc 6780
A H F Q P G D A V L E A D I A S F D K S
caagacgactcctttgcgctaaccggcgtgatgcttctggaagacctcggggtcgaccaa 6840
Q D D S F A L T A L M L L E D L G V D Q
gaactgctggaccttatcgaagctgcgtttgggtgagatcacgagtgatcctacctgcc 6900
E L L D L I E A A F G E I T S V H L P A
gggtacaagatttaaatcgggtgctatgatgaagtcaggaatgtttcttacactcttcatc 6960
G T R F K F G A M M K S G M F L T L F I
aacacgctgctgaacattgtcatagcgtgccgctcttacgcgacaaattatcgtcctcg 7020
N T L L N I V I A C R V L R D K L S S S
gcgctgcgcccttcataggtgatgacaacatagtcacggcgtgaggtcagaccgcta 7080
A C A A F I G D D N I V H G V R S D P L
atggcagaaaggtgcgcgagttgggtcaacatggaagtaagatcatcgatgccacaatg 7140
M A E R C A S W V N M E V K I I D A T M
tgtgagaaccaccgctacttttgtggaggattcatcctgtacgacagtgctgcgggtaca 7200
C E K P P Y F C G G F I L Y D S V A G T
gcggttaggggtgcagaccggttaagagagctgttcaaactcgggaaaccgctcccggcg 7260
A C R V A D P L K R L F K L G K P L P A
gacgacaaccaggatgaagacagaagaaggcactaaaagatgaaacagtttaagtggtcc 7320
D D N Q D E D R R R A L K D E T V K W S
cgcataggttagagagaagaattagacgtggcattgagctcaagataccaagtcagcggc 7380
R I G L R E E L D V A L S S R Y Q V S G
gtcggaaacatcactagagcgtatccacgctgtctaaagatttgaagcttttaggaaa 7440
V G N I T R A M S T L S K S L K S F R K
ataagaggtcccatacatctgtacggcggctcctaaatagatgcaggattacactaca 7500
I R G P I I H L Y G G P K *
tctaaagaccacgtattacagacaccatgaattacattccaactcaaaccttttacggac 7560
M N Y I P T Q T F Y G
gccgttggcgaccacgcccggcgtaccgtccatggcgggtgccgatgcagccggcccccac 7620
R R W R P R P A Y R P W R V P M Q P A P
ccatggatgctcctgagctgcaaactccgatcgtccaggcccaacagatgcagcagctaa 7680
P M V I P E L Q T P I V Q A Q Q M Q Q L
tcagtgcagtttctgcctgacgaccaagcaaaatggcaaagcaccgaagaagccgaaga 7740
I S A V S A L T T K Q N G K A P K K P K
aaaagccgcaaaaagcgaaggctaaagaaaacgaacagcaaaaagaagaacgagaacaaga 7800
K K P Q K A K A K K N E Q Q K K N E N K
aaccaccgctaaagcagaagaatccggctaaagaagaagaaccaggaaaaagggaaacgca 7860
K P P P K Q K N P A K K K K P G K R E R
tgtgcatgaagatagagaatgatgcatcttcgaggccaagcttgacggtaaggtcacgg 7920
M C M K I E N D C I F E V K L D G K V T
gatacgctgcctagtcggggataaagtgatgaagccggcacacgtcaaaggtgtgatcg 7980
G Y A C L V G D K V M K P A H V K G V I
acaaccccgacctagcgaagcttacctacaagaaatcgagcaagtatgacctggagtgcg 8040
D N P D L A K L T Y K K S S K Y D L E C
cccagataaccagtcacatgaagtcagatgcttcaaagtacacccatgaaaaaccagaag 8100
A Q I P V H M K S D A S K Y T H E K P E

ggcactacaattggcatcacgggtgcagtgacagcgggtggcaggttcacaatcccga 8160
G H Y N W H H G A V Q Y S G G R F T I P
caggcgcaggt aagccaggagacagcggccggcgcgatcttcgacaacaaggacgcgtgg 8220
T G A G K P G D S G R P I F D N K G R V
tggccatcgtcctgggaggggccaacgaaggagccaggactgccctatctgtcgtaacct 8280
V A I V L G G A N E G A R T A L S V V T
ggaccaaagacatggtcacacgggtacaccccagaaggaacagaagaatggtccgcccgt 8340
W T K D M V T R Y T P E G T E E W S A A
tgatgatgtgcgtcttagccaacgttacattcccatgctcagagcccgcgtgtgcaccct 8400
L M M C V L A N V T F P C S E P A C A P
gttgcgtgtgaagaacaaccagaacagacactgaggatgttagaggacaacgtggaccgcc 8460
C C C E E Q P E Q T L R M L E D N V D R
cgggctactacgacctgctcgaggccacgatgacgtgtaacaatagtgcacgccacggtc 8520
P G Y Y D L L E A T M T C N N S A R H G
gcagtgtagcgaaacacttcaacgtctacaaggccacgaaaccgtatctagcgtattgcg 8580
R S V T K H F N V Y K A T K P Y L A Y C
cggactcgggagacgggagcttctgttacagcccgggtggctatagaaaaaattagggatg 8640
A D C G D G Q F C Y S P V A I E K I R D
aggcttccgatggcatgataaaaaatccaggctcgagcgcgaaatggcatcaacaaaggag 8700
E A S D G M I K I Q V A A Q I G I N K G
gaacacacgaacacaacaaaatcagggtacatcgccgggcatgacatgaaagaggcaaacc 8760
G T H E H N K I R Y I A G H D M K E A N
gggactctttacaagtgataacttccgggtgtgtgcgctattcgaggcacgatgggccact 8820
R D S L Q V H T S G V C A I R G T M G H
tcatcgtggcctactgccctccaggggacgaactaaagggtccagttccaagatgcagaat 8880
F I V A Y C P P G D E L K V Q F Q D A E
cgcacaccaggcctgcaaagtgagtaacaacacgcaccggccccagtaggcagagaaa 8940
S H T Q A C K V Q Y K H A P A P V G R E
aatcaccgtcaggccccacttccggtatcgaagtgccatgcacaacgtaccagctgacta 9000
K F T V R P H F G I E V P C T T Y Q L T
ccgcaccgacggaggaagagatcgacatgcataccccaccggatcccagacataacgt 9060
T A P T E E E I D M H T P P D I P D I T
tgctgtcgagcagtgaggtaatgtaagatcacagcaggagggaaaaaccatcagataca 9120
L L S Q Q S G N V K I T A G G K T I R Y
actgcacgtgtggtagtggcaacgtgggcaccaccagtagcgacaagactatcaattcgt 9180
N C T C G S G N V G T T S S D K T I N S
gcaaaatagcacagtgccacgctcgggtgactaaccacgataagtggcagtacacctcct 9240
C K I A Q C H A A V T N H D K W Q Y T S
cgtttgtccctagagccgaccagttgtctcgaaaggtaagtgacgtacctttccctc 9300
S F V P R A D Q L S R K G K V H V P F P
tgaccaactccacatgcaggggtgcctgtgcacgtgcaccaggtgtcacatcggaaaga 9360
L T N S T C R V P V A R A P G V T Y G K
gagaactgacagtgaaactgcaccagatcatcccacgctgttgacgtaccggagtctag 9420
R E L T V K L H P D H P T L L T Y R S L
gagcagatccgcgcccgtatgaggagtggatagaccgatcgtcgaacggaccataaccgg 9480
G A D P R P Y E E W I D R Y V E R T I P
tgaccgaagatgggatcgagtacagatggggaacaaaccaccggcgtgcgcttgtgggccc 9540
V T E D G I E Y R W G N N P P V R L W A
agctgacaactgaaggcaaacccatgggtggccgcagagatcatactctattactatg 9600
Q L T T E G K P H G W P H E I I L Y Y Y
ggctataccagcagccaccatcgccgcccgtctcagccgcccgtctcgagctcgtactat 9660
G L Y P A A T I A A V S A A G L A V V L
cgctgctggcgtcatgttacatgttcgccactgcacgccgcaagtgccctgaccccatacg 9720
S L L A S C Y M F A T A R R K C L T P Y

ccctgacccccggagctgtcgtcccggtaacactaggagtactatgctgcgaccacgag 9780
A L T P G A V V P V T L G V L C C A P R
cgcatgccgctcatttgcggaatctatggcgtatctatgggatgagaatcaaaccctgt 9840
A H A A S F A E S M A Y L W D E N Q T L
tttgctggagcttgcaacgccgctcgtgccataatcatacttgtatgctgcctgaaga 9900
F W L E L A T P L A A I I I L V C C L K
acctgctttgctgctgcaaaccgctttcttttttagtgggtgagcctgggaactcccg 9960
N L L C C C K P L S F L V L V S L G T P
tcgtaaaatcttacgaacacaccgcaacgatcccgaatgtgggtgggattcccgtataagg 10020
V V K S Y E H T A T I P N V V G F P Y K
ctcacattgagaggaacggcttctccccgatgaccctacagcttgaagtacttgaacca 10080
A H I E R N G F S P M T L Q L E V L G T
gcttgaacccacgctaacttagagtacataaacctgtgaatacaagacagtctgacct 10140
S L E P T L N L E Y I T C E Y K T V V P
caccttatatacaagtgtcggggacatcagaatgcagatccatggagcgcggcggactatc 10200
S P Y I K C C G T S E C R S M E R P D Y
aatgccaggctacacaggagtgtaccatttatgtggggcggcgatactgcttctgcg 10260
Q C Q V Y T G V Y P F M W G G A Y C F C
aactgagaacgccagctgagtgagcatacgttgatagatcggacgtatgcaagcagc 10320
D T E N A Q L S E A Y V D R S D V C K H
accatgccgcccctacaaggcgcatactgcccgaatgaaagccaccatccgaataagct 10380
D H A A A Y K A H T A A M K A T I R I S
acgggaacctcaatcagacaacaacggcgttcgtcaacggggagcacacagtgtaccgtcg 10440
Y G N L N Q T T T A F V N G E H T V T V
gaggcagcaggtttacttttgggtccaatctccactgcctggacgcctttcgacaacaaga 10500
G G S R F T F G P I S T A W T P F D N K
tcgtcgtctacaagaacgagcgtctacaaccaggacttcccaccctacgggtcaggacaac 10560
I V V Y K N D V Y N Q D F P P Y G S G Q
cagggaggcttggagacatccagagcaggacggttagagagcaaagacctgtatgccaaca 10620
P G R L G D I Q S R T V E S K D L Y A N
ccgccctcaagtgtcaagacctcgtccggtactgttcacgtgccttacacacagacc 10680
T A L K L S R P S S G T V H V P Y T Q T
cttctggctttaagtactggataaaagagagaggcacgtcgtgaatgacacggctcctt 10740
P S G F K Y W I K E R G T S L N D T A P
ttggatgcgtaatacaagaccaaccagtcagagcagaaaatgcgccgttggcaacatcc 10800
F G C V I K T N P V R A E N C A V G N I
cagctccatggacatcccggacaccgctttacgcgctgatgatgcacctgccgtca 10860
P V S M D I P D T A F T R V I D A P A V
caaacctggagtccaagtggcggctgtcacgcactcatcggacttcggcgggatcgga 10920
T N L E C Q V A V C T H S S D F G G I A
ctctgactttcaaaactgacaaaccgggaaaatgtgctgtccactctcatcgaacgtag 10980
T L T F K T D K P G K C A V H S H S N V
ccaccatacaggaggcagctgtggacatcaaacagatggcaagataaccctgcatttct 11040
A T I Q E A A V D I K T D G K I T L H F
ctgcagcatcagcatccccggcattcaaggatctgtgtgcagtgccaaaacgacatgca 11100
S A A S A S P A F K V S V C S A K T T C
tggcagcgtgtgagccgccgaaggaccacatcgtcccttatggggcgagccataacaacc 11160
M A A C E P P K D H I V P Y G A S H N N
aagtttttctgacatgtctggcacggcaatgacatgggtgcagcgggttagccggcgagc 11220
Q V F P D M S G T A M T W V Q R V A G G
tcggcgggctaactcgcggcagtgccagctacttatactgggtgacgtgtgtgactatgc 11280
L G G L T L A A V A V L I L V T C V T M
gccgctaaccgggaggcttgacataatgtatatataagcatcatagttttaataaagc 11340
R R *

ccctgaatagtaacaaaacataaaaaacaaaaacagtagttcaaagggctatacaacccc 11400
 tgaatagtaacaaaatacagaaaaaccataaaaaattataaaat taactaatcagatcatc 11460
 taaatttgactaattggaatatgccgaactctacggagatgtaggcgtccgaactccacg 11520
 gagacgta 11528

Fig. 8. Nucleotide sequence of Getah virus. *; stop codon.

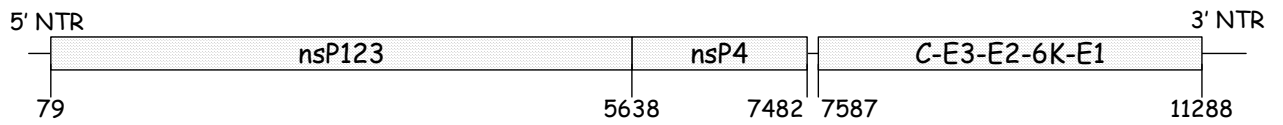


Fig. 9. Genomic organization of Getah virus.

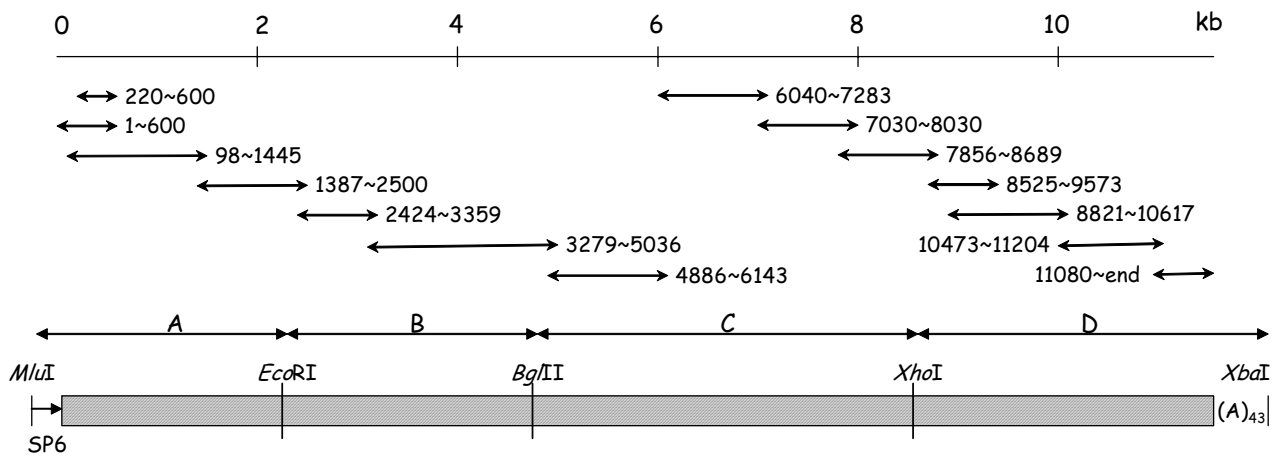
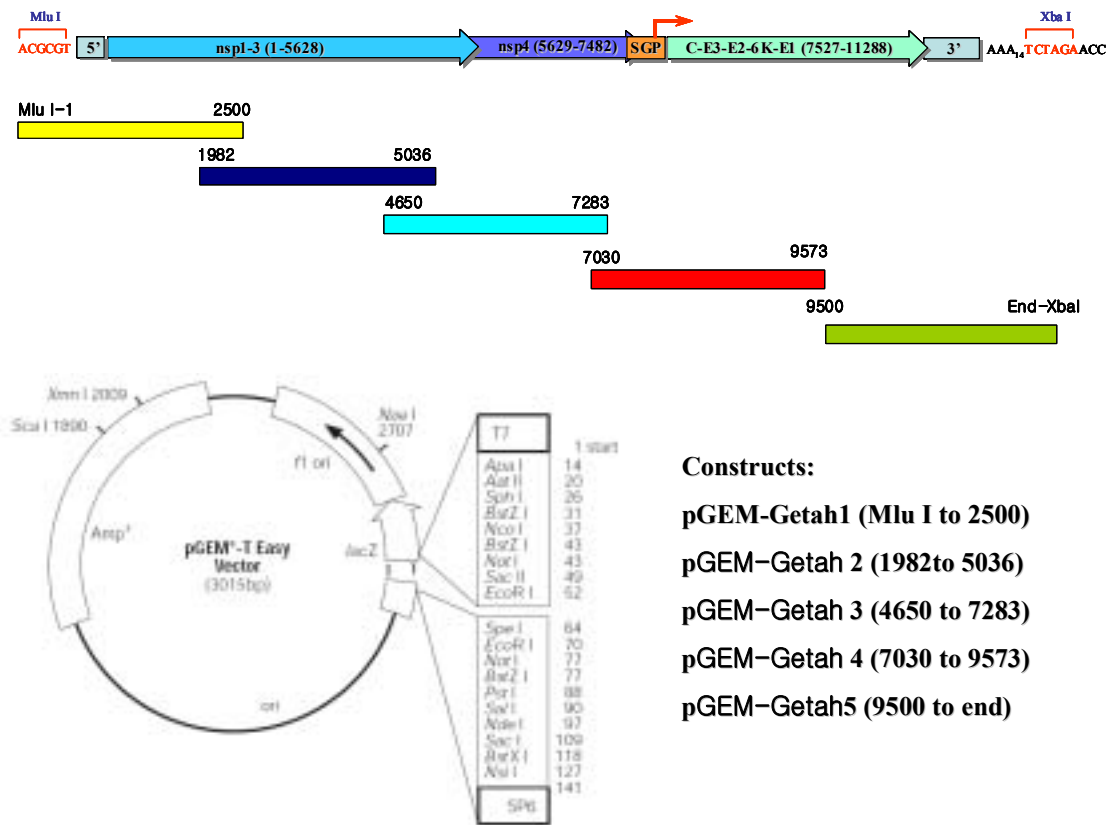


Fig. 10. Cloning strategy of Getah virus full clone.

4. PCR을 통한 게타 바이러스 gene의 증폭

전 실험을 통해 국내돼지에서 plaque purification 방법으로 순수 분리된 게타바이러스를 진핵세포를 이용하여 배양한 후 배양 상층액으로부터 Qiagen viral RNA extraction kit를 이용하여 viral RNA를 추출하였다. 추출된 RNA를 주형으로 oligo-dT₁₀를 이용하여 cDNA를 합성하였다. Reverse transcription은 GIBCO사의 Superscript II를 이용하여 42°C에서 50분간 수행되었다. 만들어진 cDNA를 주형으로 하여 NCBI의 blast search를 통하여 기존에 발표된 사기야마 바이러스 (sagiyama virus)의 염기서열을 토대로 primer를 제작하여 다섯 차례의 PCR을 각기 수행해서 게타 바이러스의 모든 염기서열을 클로닝하였다. 첫 번째 PCR에 사용된 primer는 sense; 1S-Mlu I 5'-GAGACGCGTATGGCGGACGTGTGACATCACCC-3', antisense; 2500R 5'-GTCGCCA CACAGTACTACCT-3'이며 두 번째 PCR에 사용된 primer는 sense; 1981S 5'-CGTA AGCTACACCACATAGC-3', antisense; 5036R 5'-GGACTTACTAGTGACG-3'이며, 세 번째 PCR에 사용된 primer는 sense; 4650F 5'-CAGTGCAACTGACGGTA-3', antisense; 7283R 5'-CTGTCTTCATCCTGGTGTC-3'이며, 네 번째 PCR에 사용된 primer는 sense; 7030F 5'-GCCTTCATAGGTGATGACAA-3', antisense; 9573R 5'-CCATGGGGTTTGCCT TCAGT-3'이며, 다섯 번째 PCR에 사용된 prim는 sense; 9500F5'-TGGGATCGAGTAAGAT GGG-3', antisense; End-Xba I 5'-GGCTCTAGAT₍₄₃₎GTAAAATATT-3'이다. PCR 반응은 TAKARA EX Taq polymerase를 사용하였고, 94°C 1분, 58°C에서 1분 30초, 72°C 2분의 과정을 40회 반복하여 증폭하였다. 그 결과 첫 번째 fragment는 앞쪽에 Mlu I 이 삽입된 1 ~2500 두 번째는 1981~5036, 세 번째는 4650~7283, 네 번째는 7030~9573 마지막은 9500~끝까지로, 뒷쪽에 43개의 poly-A와 Xba I이 삽입된 5개의 PCR product band를 얻을 수 있었다 (Fig. 11). 또한 Taq polymerase보다 좀더 fidelity가 좋고 blunt end PCR product를 만드는 pfu polymerase (solgene)을 이용하여 게타 RNA를 갖고 oligo-dT₁₀를 이용하여 합성한 cDNA를 주형으로 표 (2)에서 제시한 primer로 6개의 blunt end PCR product band를 얻을수 있었다(Fig. 12).



Constructs:

- pGEM-Getah1 (Mlu I to 2500)
- pGEM-Getah 2 (1982to 5036)
- pGEM-Getah 3 (4650 to 7283)
- pGEM-Getah 4 (7030 to 9573)
- pGEM-Getah5 (9500 to end)

Fig. 11. PCR amplification of Getah genimic DNA fragments with Taq polymerase and cloning of Getah genomic fragments into pGEM-T Easy vector

2. list of primers and PCR products

Fragment	Primer pair	Primer sequence	Amplicon Size
Getah 1	Mlu I-Sagi-1F Sagi-2500R	5' <i>ACG CGT</i> ATG GCG GAC GTG TGA CAT 3' 5' GTC GCC ACA CAG TAC TAC CT 3'	2,526 bp (-6 to 2520)
Getah 2	Sagi-1981F Sagi-5036R	5' CGT AAG CTA CAC CAC ATA GC 3' 5' GGA CTT ACT AGT GAC G 3'	3,056 bp (1981 to 5036)
Getah 3	Sagi-3273F Sagi-8466R	5' CGG CCT AT CTC AGC GCA GTC AGT TTC 3' 5' AGC CCG GGC GGT CTA CGT TGT CCT C 3'	5,194 bp (3273 to 8466)
Getah 4	Sagi-7255F Sagi-11204R	5' CCA GCG GAC GAC AAC CAG GAT GAA GAC 3' 5' CTG CAC CCA TGT CAT TGC CGT 3'	3,950 bp (7255 to 11204)
Getah 5	Sagi-10476F Sagi-11560R	5' GCC TGG ACG CCT TTC GAC AAC 3' 5' GAT GGT CTG GGG TTC GGC AG 3'	1,082 bp (10476 to 11558)
Getah 6	Sagi-11071F Xba I-EndSagiR	5' TAT CTG TGT GCA GTG CCA AA 3' 5' GGC <i>TCTAGA</i> TTT ₁₄ TGT AAA ATA TT 3'	578 bp (11071 to 11648)

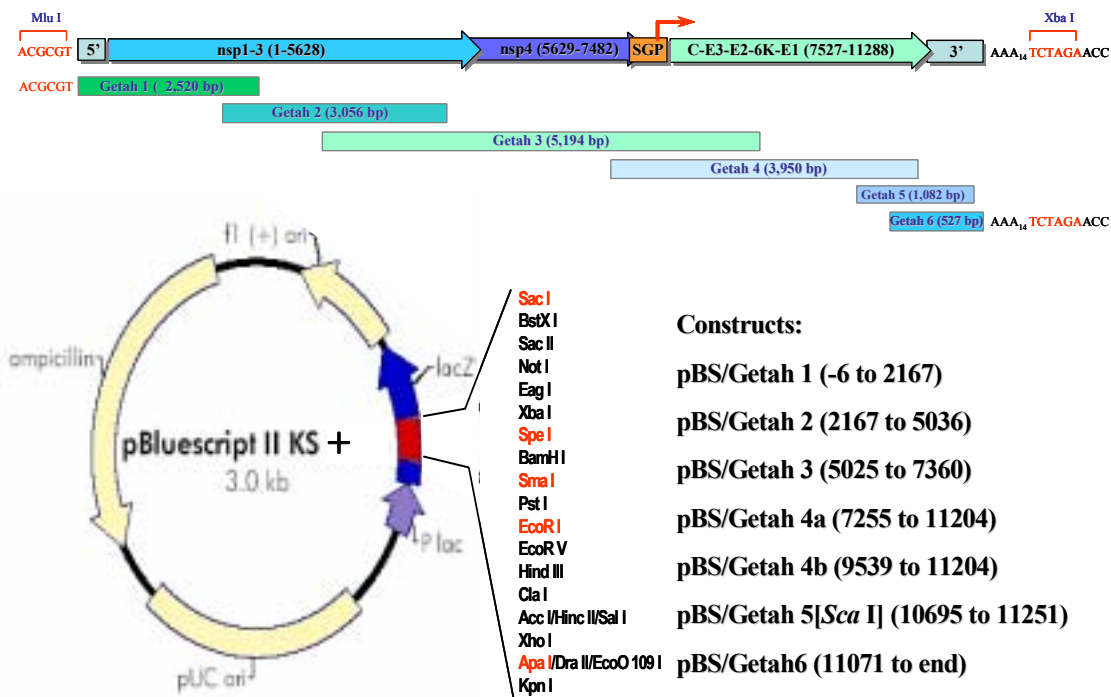


Fig.12. PCR amplification of Getah genomic DNA fragments with pfu polymerase and cloning of Getah genomic fragments into pBluescript II KS + vector

5. PCR product의 클로닝 및 염기 서열 분석

이렇게 하여 Taq polymerase로 증폭된 다섯 개의 cohesive end PCR product를 promega의 pGEM-T Easy vector에 클로닝 하였다(fig. 11). 이 다섯 개의 plasmid는 게타 바이러스 genome의 순서에 따라 Mlu I 이 삽입된 1 ~2500을 클로닝한 plasmid를 pGEM-Getah1, 1981~5036은 pGEM-Getah2, 4650~7283은 pGEM-Getah3, 7030~9573 은 pGEM-Getah4, 9500~끝은 pGEM-Getah5 라고 명명하였다. 또한 blunt end polymerase 로 증폭된 6개의 blunt end PCR product를 pBluescript II KS +(STRATAGENE)에 클로닝 하였다(Fig. 12) Mlu I 이 삽입된 1 ~2167을 클로닝한 plasmid를 pBS/Getah1, 2167~5036은 pBS/Getah2, 5025~7360은 pBS/Getah3, 7360~9539은 pBS/Getah4a, 9539~11204은 pBS/Getah4b, 10695~11251은 pBS/Geta5, 11071~end 은 pBS/Getah6 이라고 명명하였다. 이로써 게타 바이러스의 full genome 염기 서열 분석 및 각각의

㊦ 3. Getah genome comparison with Sagiya genome

nonstructural proteinORF : 1234

Gene	nucleotide difference				amino acid difference			
	position	Sagiya	Getah(ufu pol)	Getah(tao pol)	position	Sagiya	Getah(ufu pol)	Getah(tao pol)
nsP1 <i>nt 79-1680</i> <i>534 a.a.</i>	108	C	T	T				
	196	C	T	T				
	292	A	A	G	72	M	M	V
	333	C	T	T				
	417	G	A	A				
	432	T	A	A				
	483	C	T	T				
	486	C	T	T				
	519	T	C	C				
	651	T	C	C				
	657	A	G	G				
	753	T	C	C				
	786	C	T	T				
	819	T	C	C				
	966	C	T	T				
	978	T	C	C				
	983	A	A	G	302	Y	Y	C
	1016	G	A	A	313	V	M	M
	1072	T	T	A	332	Y	Y	N
	1107	C	T	T				
	1119	C	T	T				
	1182	T	C	C				
	1263	T	A	A				
	1269	G	A	A				
	1287	T	C	C				
	1386	G	A	A				
	1389	G	A	A				
	1483	C	T	T				
	1536	C	T	T				
	1566	A	G	G				
	1571	C	T	T	498	A	V	V
	1581	A	T	T				
	1596	A	T	T				
1623	C	T	T					
1656	A	G	G					
1660	T	C	C					
nsP2 <i>nt 1681-4074</i> <i>798 a.a.</i>	1708	T	C	C				
	1737	C	T	T				
	1776	T	G	G				
	1815	C	T	T				
	1863	G	A	A				
	1905	G	A	A				
	1983	T	C	C				
	2056	G	A	A				
	2082	T	C	C				
	2146	G	A	A				
	2244	A	G	G				
	2367	G	A	A				
	2395	T	C	T				
	2397	A	G	G				
	2411	G	A	A	778	R	K	K
2478	T	C	C					
2514	T	C	C					
nsP2 <i>nt 1681-4074</i> <i>798 a.a.</i> <i>(continuation)</i>	2580	C	T	T				
	2622	A	T	T				
	2631	C	T	T				
	2697	T	C	C				
	2913	T	C	C				
	2952	A	G	G				
	3019	T	C	C				

Gene	nucleotide difference				amino acid difference			
	position	Sasivama	Getah(ufu no)	Getah(tao no)	position	Sasivama	Getah(ufu no)	Getah(tao no)
	3033	A	T	T	985	Q	H	H
	3105	C	T	T				
	3108	A	G	G				
	3138	C	T	T				
	3219	T	A	A				
	3294	A	G	G				
	3306	C	T	T				
	3351	T	C	C				
	3381	T	C	C				
	3435	T	C	C				
	3456	A	G	G				
	3459	T	C	C				
	3477	T	C	C				
	3513	A	G	G				
	3519	T	A	A				
	3534	C	T	T	1176	A	V	V
	3611	C	T	T				
	3663	A	G	G				
	3717	T	C	C				
	4057	C	T	T				
	4071	C	T	T				
<i>nsP3 at 4075-5646 524 aa</i>	4101	T	A	A	1347	G	S	S
	4117	G	A	A				
	4167	T	C	C				
	4221	G	A	A				
	4242	G	A	A				
	4287	G	A	A				
	4290	T	C	C				
	4311	T	C	C				
	4341	C	T	T				
	4468	T	C	C				
	4476	C	A	A				
	4560	A	C	C				
	4596	C	T	T				
	4671	T	C	C				
	4749	C	T	T				
	4818	G	A	A				
	4821	A	G	G				
	4839	A	G	G				
	4890	A	G	G				
	4950	G	A	A				
	5003	T	T	C	1642	I	I	T
	5119	C	T	T	1681	L	F	F
	5121	C	T	T				
	5124	A	G	G				
	5181	G	A	A				
	5232	C	T	T				
	5267	T	C	C	1730	I	T	T
	5370	T	C	C				
<i>nsP3 at 4075-5646 524 aa (continuation)</i>	5391	C	T	T	1775	S	L	L
	5402	C	T	T				
	5406	G	A	A				
	5409	A	G	G				
	5433	G	A	A				
	5455	A	T	T	1793	T	S	S
	5474	G	A	A	1799	G	B	B
	5481	T	C	C				
	5496	G	T	T				
	5502	A	G	G				
	5575	T	C	C	1833	W	R	R
	5577	G	A	A				
<i>nsP4 at 5647-7484</i>	5685	G	A	A	1871	R	K	K
	5690	G	A	A				

Gene	nucleotide difference				amino acid difference			
	position	Sasivama	Getah(ofu pol)	Getah(tao pol)	position	Sasivama	Getah(ofu pol)	Getah(tao pol)
<i>611 s.s.</i>	5730	T	C	C				
	5781	G	A	A				
	5785	A	G	G	1903	I	V	V
	5901	A	G	G				
	5934	T	C	C				
	5995	A	G	G	1973	N	D	D
	6189	G	A	A				
	6216	C	T	T				
	6219	C	T	T				
	6229	C	T	T				
	6267	T	C	C				
	6342	C	A	A				
	6432	C	T	T				
	6435	T	A	A				
	6474	A	G	G				
	6480	C	T	T				
	6567	T	A	A				
	6579	G	A	A				
	6598	T	C	C				
	6658	T	C	C				
	6675	C	T	T				
	6774	T	C	C				
	6786	T	C	C				
	6810	A	G	G				
	6828	T	C	C				
	6864	C	T	T				
	6951	G	A	A				
	7009	C	T	T				
	7095	T	C	C				
	7221	A	G	G				
	7251	A	G	G				
7257	A	G	G					
7302	C	T	T					
7326	C	A	A					
7354	C	T	T					
7377	T	C	C					
7380	T	C	C					
7411	T	C	C					
7421	A	G	G	2448	N	S	S	

structural protein ORF C-E3-E2-6k-E1

Gene	nucleotide difference				amino acid difference			
	position	Sagiyana	Getah(pfu pol)	Getah(taq pol)	position	Sagiyana	Getah(pfu pol)	Getah(taq pol)
Capsid <i>nt 7527-8330</i> 534 a.a.	7538	C	T	T				
	7585	T	A	A	20	F	Y	Y
	7628	T	G	G				
	7698	C	C	A	58	L	L	M
	7787	A	G	G				
	7808	A	G	G				
	7811	C	T	T				
	7825	T	C	C	100	L	P	P
	7859	A	C	C				
	7877	A	G	G				
	7904	C	T	T				
	7928	A	C	C				
	7949	A	G	G				
	8033	A	G	G				
	8078	A	G	a				
	8102	A	G	G				
	8135	T	C	C				
	8174	G	A	A				
	8216	T	C	C				
	8258	A	G	G				
8270	T	C	C					
8276	A	G	G					
8321	T	A	A					
E3 <i>nt 8331-8522</i> 64 a.a.	8390	A	G	G				
	8441	G	A	A				
	8453	A	G	G				
	8471	T	C	C				
E2 <i>nt 8523-9788</i> 422 a.a.	8532	G	A	A	336	E	K	K
	8543	T	C	C				
	8558	T	G	G				
	8585	T	C	C				
	8645	C	T	T				
	8663	G	A	A				
	8697	A	G	G	391	R	G	G
	8734	T	C	C	403	V	A	A
	8742	T	G	G	406	Y	D	D
	8756	C	A	A				
	8765	T	C	C				
	8768	T	T	C				
	8798	C	T	T				
	8888	A	C	C	464	D	A	A
	8917	A	C	C				
	8918	T	A	A				
	8957	T	C	C				
	9032	C	T	T				
	9081	G	A	A	519	D	N	N
	9117	T	T	C	531	Y	Y	H
9147	G	G	A	541	G	G	S	
9221	C	T	T					
9267	C	T	T	582	P	S	S	
9293	A	T	T					

Gene	nucleotide difference				amino acid difference			
	position	Sasivama	Getah(ouf po)	Getah(tao po)	position	Sasivama	Getah(ouf po)	Getah(tao po)
B2 <i>at 8523-9788</i> 422 a.a (continuation)	9326	C	T	T	601	L	V	V
	9327	C	G	G				
	9379	T	T	C	618	L	L	P
	9383	T	C	C				
	9407	A	G	G				
	9431	A	G	A				
	9434	T	C	T				
	9458	G	A	G				
	9599	C	T	T				
	9623	T	C	C				
9716	T	A	A					
9755	G	A	A					
6K <i>at 9789-9971</i> 61 a.a	9794	G	A	A	798	Q	K	K
	9806	C	T	T				
	9869	C	T	T				
	9918	C	A	A				
	9939	T	C	C				
B1 <i>at 9972-11284</i> 438 a.a	10067	G	A	A	872	P	S	S
	10140	C	T	T				
	10148	C	T	T				
	10164	G	A	A				
	10232	C	T	T				
	10278	T	C	C				
	10298	C	T	T				
	10304	A	G	G				
	10325	C	T	T				
	10331	G	C	C				
	10346	C	T	T				
	10502	T	C	C				
	10553	T	A	A				
	10578	G	A	A				
	10604	A	G	G				
	10622	T	C	C				
	10688	T	C	C				
	10739	T	C	C				
	10769	T	C	C				
	10823	T	C	C				
	10824	T	A	A				
	10919	A	G	G				
	10987	C	T	T				
	10964	C	T	T				
	10967	C	T	T				
	11015	A	G	G				
	11068	T	A	A				
	11105	G	A	A				
	11126	T	C	C				
	11129	T	C	C				
11144	A	G	G					
11171	C	T	T					
11183	T	C	C					
11213	T	C	C					
11251	C	T	T					
					1242	A	V	V

fragment를 ligation 함으로써 게타 replicon 및 helper plasmid의 구축이 가능하게 되었다. 게타 full genome의 염기 서열 분석은 pGEM-T Easy vector와 pBluescript II KS + 에 클로닝된 상태에서 T7 sequencing primer 및 각 500bp마다 gene specific sequencing primer를 제작하여 각기 다섯 개의 plasmid 당 2~3개의 다른 클론을 주식회사 솔젠트에 의뢰하여 염기 서열 분석을 하였다. 그리고 NCBI의 blast search를 통하여 기존에 발표된 게타바이러스 및 다른 Alphavirus의 염기서열과의 상동성을 비교해 보았다. 비교결과 전 실험에서 이미 제시한대로 alphavirus 중 Sagiya virus와 97% 이상의 상동성을 갖는 것이 확인하였다. 또한 Taq polymerase와 blunt end polymerase의 fidelity의 차이에 의해 두개의 클론 사이에도 몇 개의 염기 서열에 차이를 보였다 (표. 3).

6. 게타 바이러스 5'UTR의 확인

pGEM-T Easy vector에 클로닝된 게타 바이러스 다섯 개의 fragments를 가지고 게타 바이러스의 genome중에서 5'UTR에 해당하는 1~78 염기 서열과 nonstructural protein에 해당하는 1~7526 염기 서열을 클로닝하기 위해 pGEM-Getah1,2,3 plasmid를 이용하기로 하였다. 그런데 게타 바이러스 replicon를 제작하기 앞서 게타 바이러스가 속하는 알파 바이러스는 5'UTR의 RNA 구조가 바이러스의 replication에 중요한 역할을 한다고 보고되어 있어 게타 바이러스의 5'UTR에 대해 좀더 관심을 갖게 되었다. 게타 바이러스의 5'UTR을 포함하고 있는 pGEM-Getah1은 NCBI의 blast search를 통하여 기존에 발표된 사기야마 바이러스 (sagiya virus)의 염기서열을 토대로 5'UTR primer를 제작하여 PCR을 통해 gene을 증폭 하였기 때문에 게타 바이러스의 5'UTR 과는 차이가 있을 수 있다는 가설을 갖고 실험을 실시 하였다. 게타바이러스를 진핵세포를 이용하여 배양한 후 배양 상층액으로부터 Qiagen viral RNA extraction kit를 이용하여 viral RNA를 추출한 후, Reverse transcription은 GIBCO사의 Superscript II를 이용하여 RNA를 주형으로 oligo-dT₁₀를 primer로 하여 42°C에서 50분간 수행하였다. 그리고 Roche에서 판매하는 5'/3' RACE KIT, 2nd Generation을 이용하여 protocol에 따라 게타 바이러스의 5'UTR과 nonstructural protein 1을 포함하는 약 600bp를 PCR로 증폭하였다. 그리고 pGEM-T easy vector에 클로닝된 상태에서 주식 회사 솔젠트에 의뢰하여 T7 sequencing primer로 염기 서열 분석을 한 결과 게타 바이러스의 5'UTR과 사기 야마 바이러스 5'UTR은 차이가 없음을 밝혔다. 또한 게타 바이러스 replicon을 제작하기 위해서 사용될 pGEM-Getah1는 게타 바이러스의 5'UTR (Fig. 13a)앞에 Mlu I 제한 효소 site를 포함 (Fig. 13b) 하고 있어 이 것으로 인해 5'UTR의 RNA구조가 달라질 수 있으므로 Mlu I

의 유무에 따른 RNA구조를 살펴 보았다. 그 결과 Mlu I 제한 효소 site를 포함하고 있어도 5'UTR의 RNA구조에는 별 차이가 없어 보였다(**Fig. 13**). 그래서 pGEM-Getah1,2,3 plasmid를 이용하여 게타 바이러스 replicon를 제작하기로 하였다.

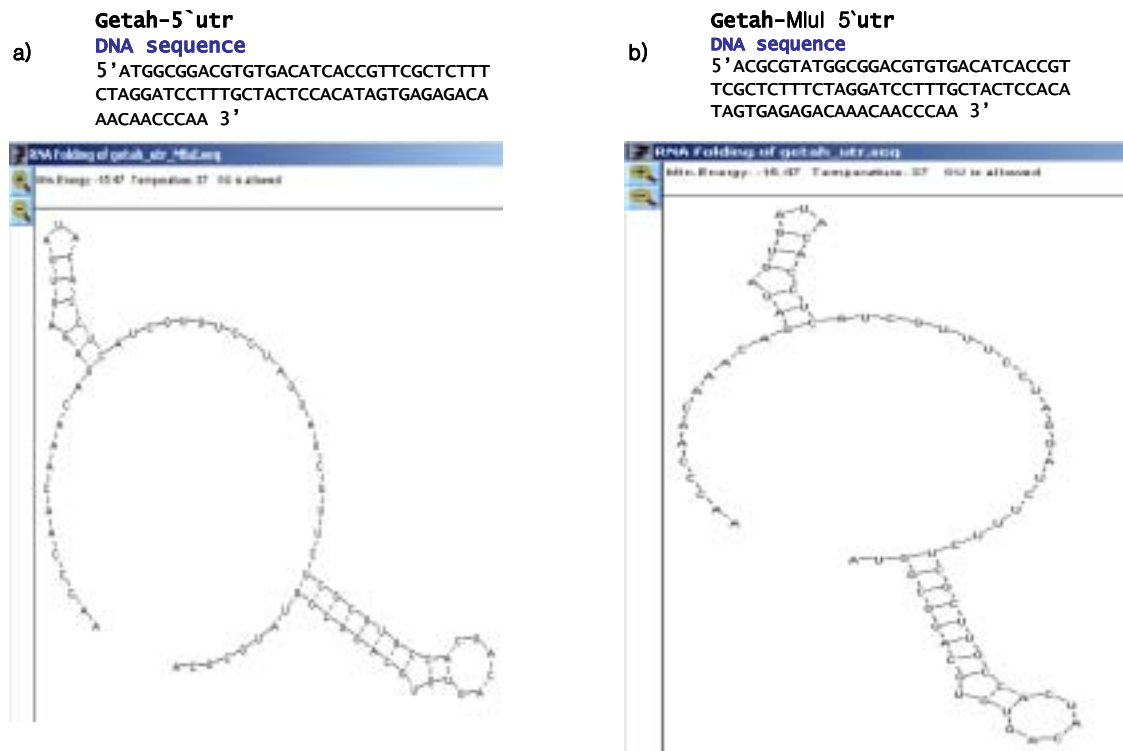
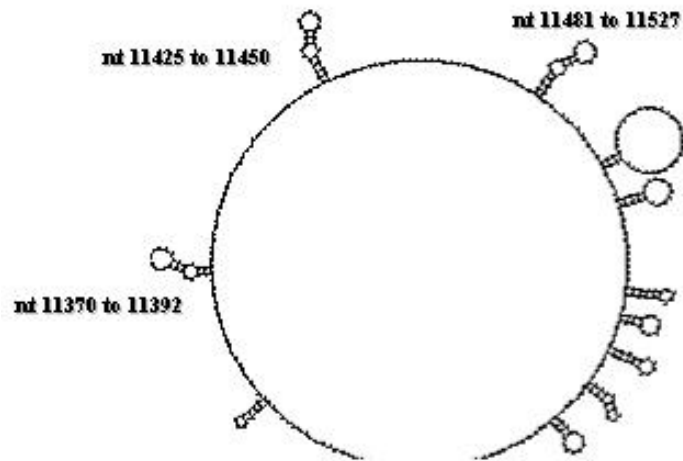


Fig. 13. Secondary structure of getah_5`utr RNA and getah_Mlu I 5`utr RNA

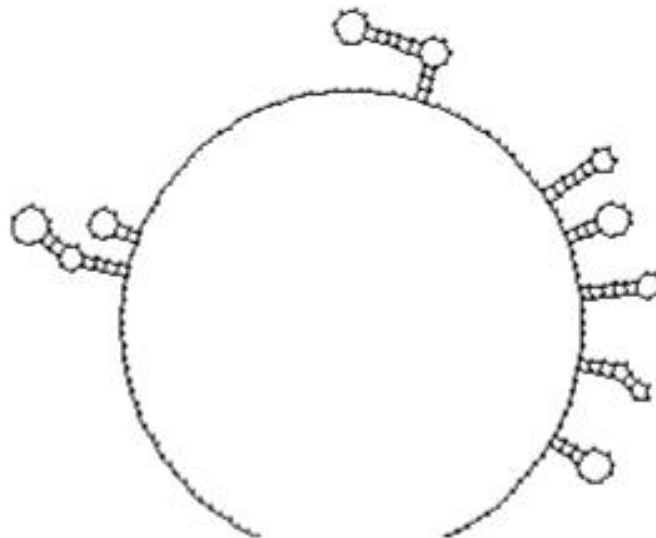
7. 게타 바이러스 3'UTR의 확인

알파 바이러스의 경우 3'UTR이 바이러스 replication에 중요한 역할을 한다고 보고되어진바 있다 우리가 조사한 게타 바이러스 gene coding region 및 5'UTR의 경우 사기야마 바이러스와 97%의 상동성을 보이는 것으로 결과가 나왔는데 3'UTR의 경우는 11330~11440사이의 약 100bp가 게타 바이러스에는 deletion이된 것으로 보인다 (**Fig. 14b**). 이결과 3'UTR의 RNA구조가 달라졌음을 (**Fig. 14a**)에서 보여주고 있다.

a) Sagiyama 3'UTR (410 nt)



Getah 3'UTR (309 nt)



b)

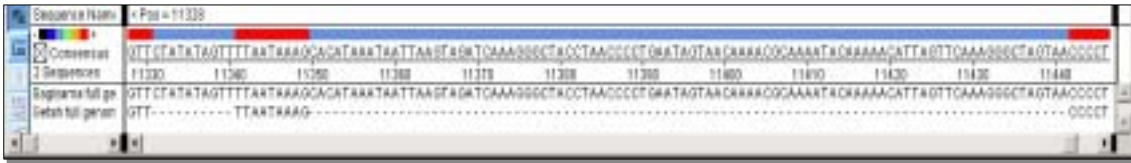


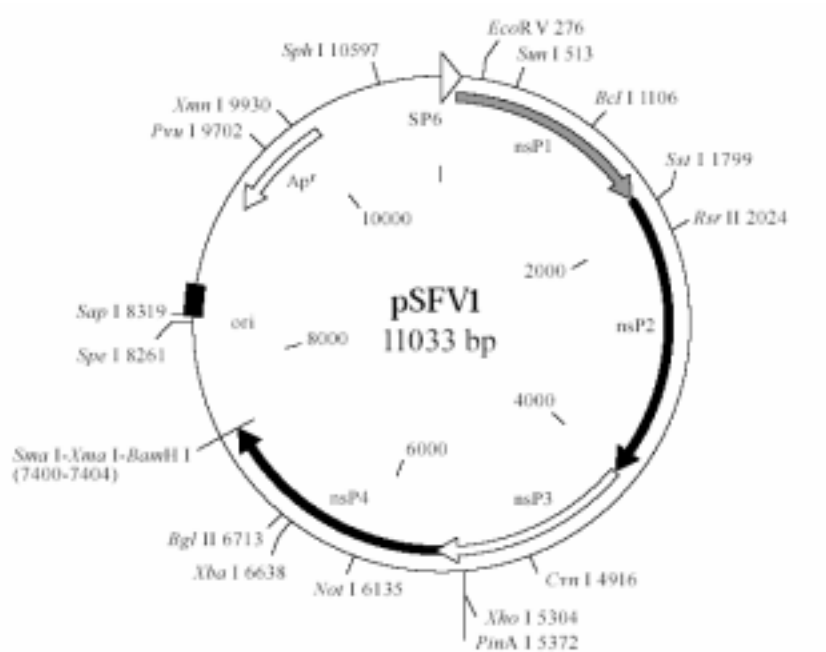
Fig. 14. Structural motifs at the 3'UTR

8. 게타 바이러스 replicon 및 helper plasmid의 구축

게타 바이러스 replicon을 제조함에 있어 이미 상용화 되고 있는 pSFV1 (**Fig. 15a**)을 이용하였다. 이 vector는 게타 바이러스와 같은 Alphavirus에 속하는 Semliki Forest virus를 상용화한 바이러스 벡터 시스템이다. pSFV1의 Semliki Forest virus의 nonstructural protein 1234부위를 제거하고 그곳을 게타 바이러스 nonstructural protein 1234 (**Fig. 15b**)로 교체했으며 외래 단백질 발현을 위한 외래 단백질 coding gene을 삽입할 multi cloning site를 subgenomic promoter뒤에 삽입 하였다. 우선 subgenomic promoter와 multi cloning site를 pSFV1에 삽입하기 위해서 sense; 7030F 5'-GCCTTCATAGGTGATGACAA-3', antisense; 7526-26sR 5'-AGATCTCTAGAGGCC CTGTAGGTGTCTGTAATACGTGGTCTTT-3' primer를 사용하여 게타 바이러스 cDNA를 주형으로 TAKARA EX Taq polymerase를 이용하여, 94°C 1분, 58°C에서 1분 30초, 72°C 2분의 과정을 40회 반복하여 nonstructural protein 4 뒷부분과 subgenomic promoter 그리고 Xho I, Apa I, Xba I, Bgl II의 multi cloning site를 포함하는 부위를 PCR로 증폭하였다. PCR product는 Qiagen pDrive cloning kit에 클로닝하여 염기 서열 분석후 Kpn I, Bgl II제한 효소 로 자르고 Kpn I, BamH I제한 효소로 자른 pSFV1에 ligation 하였다. 그 후 5'UTR 앞에 Mlu I 이 존재하는 pGEM-Getah1를 pSFV1에 ligation 하기 위한 전 과정으로 pSFV1의 SP6 promoter뒤에 Mlu I site를 삽입하였고 두 plasmid를 모두 EcoR I, Mlu I 제한 효소를 처리한 후 elution 하여 ligation 하였다. 그후 게타 바이러스 nonstructural protein 1234 나머지 부위의 클로닝 시에는 pGEM-Getah2,3에 one-site로 cutting 되는 restriction enzyme site인 2167에 위치하는 EcoR I과 4760에 위치하는 Bgl II, 그리고 7127 에 있는 Cla I site와 제한 효소를 이용하여 자르고 ligation 작업을 수행한 결과 게타 바이러스 replicon을 완성하였다. 게타 바이러스 helper 제조는 게

타 바이러스 structural protein과 3'UTR인 7527~끝까지를 포함하는 pGEM-Getah4,5를 one-site로 cutting 되는 restriction enzyme site인 9539에 위치하는 Apa I 을 이용하여 자르고 ligation하여 게타 바이러스 helper를 제조하였다 (Fig. 16).

a)



b)

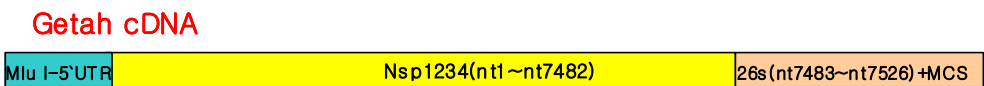


Fig. 15. Construction of pGetah replicon plasmid

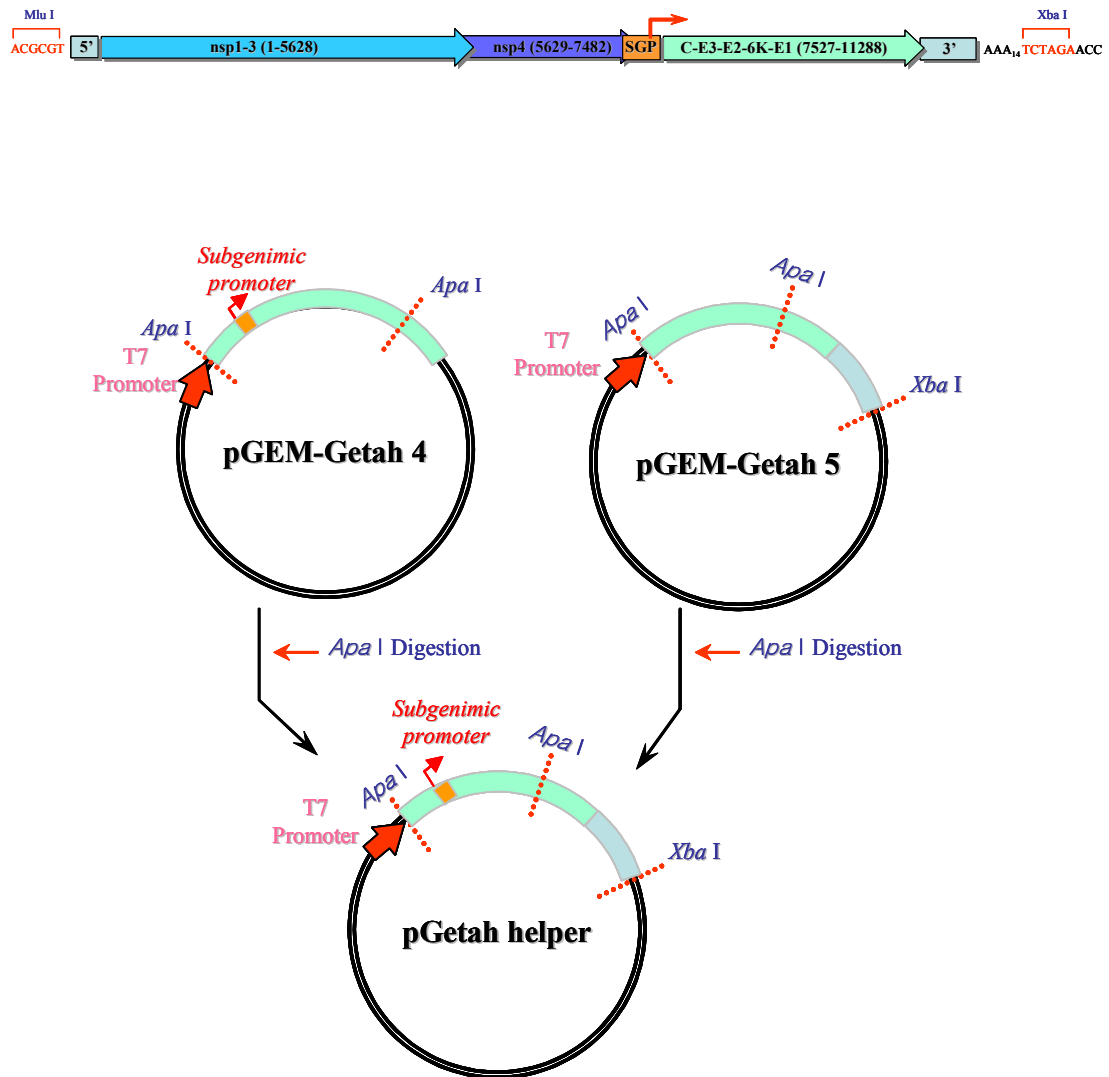


Fig. 16. Construction of pGetah helper plasmid

9. 게타 바이러스 replicon에 reporter 유전자인 GFP 삽입

클로닝이 완성된 게타 바이러스 replicon의 작동 여부를 확인하기 위해서 replicon에 reporter 유전자인 GFP와 게타 바이러스의 replication에 중요한 역할을 하는 3'UTR를 삽입하기로 하였다. GFP 720bp를 PCR로 증폭하여 pGEM-T Easy vector에 클로닝 하고 3'UTR을 포함하는 pGEM-getah5를 11100에 위치하는 Nsi I 제한 효소로 잘라 structural protein E1의 c-terminal 부위 약 100개의 아미노산과 3'UTR를 GFP뒤에 연결 하였다 그리고 Apa I, Xba I으로 잘라 역시 같은 제한효소로 잘린 게타 바이러스 replicon에

ligation 하였다.

10. Mamalian 세포에서 게타 바이러스 replicon의 작동 여부 확인

게타 바이러스 replicon에 삽입된 reporter 유전자인 GFP이 발현을 하는지 알아보기 위해 mamalian 세포 인 BHK (Baby hamster kidney)-21 cell에 in vitro RNA transcription 된 게타 바이러스 replicon을 electroporation 으로 transfection 하였다. GFP가 발현된 positive control은 pSFV1-GFP가 transfection된 BHK-21 cell로 하였고 negative control 은 plasmid를 transfection하지 않은 BHK-21 cell로 하였다.

가. cell culture

BHK-21 cell은 5% fetal bovine serum과 10% tryptose phosphate broth, 20mM HEPES, 2mM glutamine (Gibco BRL, USA) 그리고 0.1U penicillin and 0.1 μ l/ml streptomycin가 첨가된 minimum essential medium (α -MEM)을 사용하고 COS-1 cell은 10% fetal bovine serum이 함유된 Dulbecco's modified essential medium (DMEM, Gibco BRL)을 사용하여 5% CO₂, 37°C에서 배양한다.

나. in vitro RNA transcription과 mammalian 세포로의 transfection

pSFV1-GFP와 pGetah-GFP를 각각 Spe I, Xba I 제한 효소로 처리하여 선상 DNA로 만든 후 이 DNA 5 μ l, 40mM Tris-HCl (pH7.5), 6mM MgCl₂, 2mM spermidine, 5mM DTT(TaKaRa), 1mM each of ATP,CTP and UTP, 0.5mM GTP(BM), 1mM Cap analogue m7G(5')PPP(5')G(BM), 40U RNase inhibitor(TaKaRa), and 50U SP6 RNAPolymerase (TaKaRa), 로 mixture를 만든 후, 37°C에서 1시간 30분 동안 반응 시켰다. Transfected mRNA (50 μ l)를 PBS에 부유시킨 BHK-21 cell(10⁷ cell/ml) 800 μ l에 넣은후 0.4cm cuvette에서 830v/25 μ F로 2번 pulse를 주었다.

다. 형광 현미경 으로 GFP 발현의 확인

BHK-21 cell에 pSFV1-GFP와 pGetah-GFP를 각각 electroporation으로 transfection한 후5% CO₂, 37°C에서 12시간 배양한 후 형광 현미경으로 GFP 발현의 여부를 확인하였는데 positive control인 pSFV1-GFP이 transfection된 BHK-21 cell 약 70%의 발현률을 나타낸 반면 (**Fig. 17 b**)pGetah-GFP가 transfection된 BHK-21 cell은 그 발현 정도가 미약하였고 10%미만의 cell 만이 발현 되었다(**Fig. 17 c**). negative control인 plasmid를

제 4 장. 목표 달성도 및 관련분야에의 기여도

1. 연도별 연구 목표 및 달성도

가. 기술개발 목표

게타바이러스를 국내돼지에서 분리하여 전체 염기서열을 밝히고 비구조 단백 부분과 구조단백 부분을 구분하여 replicon과 helper system을 구축하여 이를 토대로 실험실내에서 외래단백질을 고등 진핵세포에서 발현시킬 수 있는 외래 단백질 발현 벡터시스템 구축하고, 또한 동물에서 외래유전자의 운반체로서의 생백신 시스템을 구축하는 것이다.

나. 평가방법 및 평가항목 및 달성도

평가방법은 각각의 연구개발 목표 대비 실적을 감안하여 평가하며 평가항목은 단계별, 연차별 개발목표 항목을 위주로 평가한다.

1) 연도별 주요개발 평가 항목

구 분	주요 개발내용 및 범위
1단계(기초연구)	<ul style="list-style-type: none"> . Plaque purification을 이용한 게타바이러스 분리 및 cDNA 클론 작업 . 전체 염기서열 결정 및 각각의 기능성 단백질 구성 확인 . 비구조단백유전자를 이용한 replicon제조 및 구조단백유전자를 이용한 helper plasmid 제조
2단계 (응용연구 및 시제품제작)	<ul style="list-style-type: none"> . Reporter 유전자를 이용하여 게타바이러스 발현벡터 시스템 구축 . 다양한 항원의 발현 실시 . Vitro에서 포유세포에서의 외래단백질 대량 발현 확인
3단계(사업화)	<ul style="list-style-type: none"> . 대상동물 및 실험동물에서의 캐리어 벡터로서의 항원발현 확인 . 축산분야의 유용 단백질 발현 및 생산 그리고 질병예방을 위한 생백신 시스템 구축

2) 달성도

국내돼지에서 게타바이러스를 plaque purification 방법으로 순수 분리하여 이에 대한 적정한 cDNA 클론 작업을 수행하였고 전체 유전자의 염기서열을 확인하였다. 이를 토대로 기존의 Alphavirus 염기서열과 비교 분석하여 유전자의 배열로 비구조 단백질 유전자와 구조 단백질 유전자의 구성을 확인하였다. 그리고 비구조 단백질 유전자를 이용한 replicon 제조 및 구조 단백질 유전자를 이용한 helper plasmid를 제조하였다.. 각각의 plasmid 제조를 위해 T7

promoter와 SP6 promoter를 이용하여 RNA transcript 및 DNA를 제조할 수 있는 plasmid를 제조함으로써 1단계 연구에 대한 당초 목표에 도달하였다. 그러나 2단계 목표로 제시한 Green fluorescent protein 유전자를 reporter 유전자로 단백질 발현 여부를 확인한 결과 발현량이 적어 발현량을 높이기 위해 여러 가지 시도를 하고 있는 중이다. 그래서 reporter 유전자로 단백질 발현 여부를 확인 후 수행될 수 있는 게타 replicon을 이용한 다양한 항원의 in Vitro 포유세포에서 대량 발현 확인 및 대상동물 및 실험동물에서의 캐리어 벡터로서의 항원발현 확인 그리고 최종적인 목표인 축산분야의 유용 단백질 발현 및 생산 그리고 질병예방을 위한 생백신 시스템 구축은 달성하지 못했다.

제 5 장. 연구 개발 결과의 활용 계획

1. 추가 연구의 필요성

현재 기준 진핵세포 발현 체계 자체만으로 국내·외 시장규모는 세계적으로 연간 약 3,000만 달러(약 360억원)에 이르며 이를 이용한 단백질 발현의 특허만도 100가지 이상(erythropoietin 등의 유용단백질 및 항원 발현 체계 등)이 되어 다양한 항원 발현 및 유전자 전달체로서의 기능 확대 등을 고려하면 매년 2배 이상의 성장을 나타낼 것이다. 또한 새로운 응용분야가 계속 개발되어 그 시장의 확대는 매우 급격하게 이루어 질 것이며 이에 따라 신규 발현 체계의 개발은 물질특허와 같은 novel system으로 특허를 획득할 수 있을 것이며 그 유용성이 매우 높다. 따라서 좀더 심도있는 연구를 통한 당초 연구 목표에 도달해야 할 것이다.

2. 활용방안 및 사업화 계획

게타바이러스를 이용한 외래 단백질 발현백터 시스템을 구축함으로써 고등 진핵생물에서 원하는 단백질을 효율적으로 대량생산 할 수 있으므로 많은 연구기관에서의 외래 단백질 발현 백터로서 이용될 것으로 보인다. 특히 알파바이러스는 다양한 숙주세포에 침입이 가능하므로 광범위한 분야에서 그들이 흥미있는 유전자를 삽입함으로써 실험을 할 수 있을 것이다.

또한 이러한 defective infectious virus particle을 이용한 시스템이 완성된 후 일부 염기서열에 변이를 가하여 포유세포 등에서 원하는 단백질을 계속 생산하는 시스템도 구축할 수 있을 것으로 보이며, 돼지, 가금류 등에서 다른 질병의 생백신으로서의 캐리어백터 역할을 수행할 수 있으므로 여러 백신 분야에 있어서 효율적인 운반체 역할을 할 것으로 보인다.

활용방안 및 사업화 계획표

핵심기술	활용유형	활용방안
- Getah 바이러스 발현 시스템	특허출원 및 상품화	유용단백질의 발현 및 분리정제를 통한 생산에 이용 (병원성 바이러스 항원 및 고가의 의약품 생산)
- 발현시스템이용 유전자 전달 시스템	특허출원 및 상품화	동물용 생백신 제제 개발 (안전한 동물용 백신 유전자를 도입하여 유전자 전달을 통한 새로운 백신개발)
- 발현 시스템 packaging 세포주 구축	특허출원 및 상품화	유전자 전달체 및 유용 단백질 발현을 위한 packaging cell line 구축 및 상용화

이러한 기술개발이 이루어질 경우 차후 유전자 면역이나 유전자 치료를 위한 캐리어벡터 분야에 있어서도 이용될 수 있을 것이다.

3년간의 기술개발 성공시 활용방안에서 언급한대로 특허권 활용을 위한 상품화 및 현 장애로 해결을 위한 기초를 확립하고 본사에서 적용되는 제품과 관련하여 발현체계로서의 상품화는 물론 백신으로의 제품을 개발할 것이며 이에 관련되는 업체와도 기술 전수를 통해 시장에 직접 판매할 수 있도록 할 것입니다.

제 6 장. 연구개발 과정에서 수집한 해외 과학기술정보

Getah virus는 Togaviridae의 Alphavirus에 속하며 이에는 Sindbis virus, Semliki Forest virus 등 이미 상용화된 바이러스 벡터 시스템과 같은 군에 속한다. 이러한 Alphavirus를 이용한 단백질 발현 시스템은 숙주세포의 범위가 넓고, 외래단백질을 대량 생산할 수 있다는 잇점이 있으며, 최근에는 유전자 치료와 백신연구 등 많은 실험분야에 응용되고 있는 실정이다. 그러나 이들 알파바이러스계열의 벡터 시스템인 Sindbis virus, Semliki Forest virus 시스템은 nonstructural protein gene이 coding 하는 protein에 의해 숙주 세포의 단백질 생합성을 억제하여 숙주세포에 cytopathic effect 가 나타난다. 이로 인해 실험상 in vitro에서 cell line에 transfection 하거나 infection 했을때 2~3일 이상이 지나면 모두 죽는다. 그래서 이런 벡터를 suicidal vector 라고도 부른다. 뿐만 아니라 in vivo로 실험동물에 직접 투여 했을 경우에도 이와 같은 cytopathic effect 가 나타난다. 그렇게 때문에 이런 알파 바이러스 계열의 벡터 시스템이 광범위하게 사용되는데 어려움이 겪고 있다. 그래서 이런 단점을 보완하기 위한 연구들이 시도 되고 있다. 이러한 Sindbis virus, Semliki Forest virus vector는 DNA vaccine으로 이용될 수 있는데 DNA vaccine의 경우 숙주 세포에 삽입되어 들어가 숙주에 암을 발생시킬 수 있는 위험성이 있다. Sindbis virus, Semliki Forest virus vector와 같은 suicidal vector는 DNA vaccine으로 이용되어 숙주 세포에 들어간 후 숙주세포가 암을 유발하기전에 죽게 하기 때문에 이런 DNA vaccine의 위험성을 피해갈수 있다. 반면 숙주 세포가 너무 일찍 죽게되어 지속적인 면역유도능이 떨어지는 단점이 있는데 이를 보완하기위해 suicidal vector에 antiapoptotic protein을 coding하는 gene을 항원 gene과 함께 삽입하는 전략이 많은 연구들에 이용되고 있다.

또한 Sindbis virus, Semliki Forest virus vector의 cytopathic effect을 적극적으로 이용하여 암세포를 제거하는 경우도 있다. 특정 유전자를 발현하는 암세포에 특징적인 수용체를 매개로 suicidal vector가 infection 되어 그 암세포만을 특이적으로 죽여 암을 치료하는 실험들이 수행되고 있다.

제 7 장 참고 문헌

1. Liljestrom P and Garoff H. A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. 1991 *Biotechnology* 9: 1356-1361
2. Powers AM, Brault AC, Shirako Y, Strauss EG, Kang W, Strauss JH, Weaver SC. Evolutionary relationships and systematics of the alphaviruses. 2001 *J Virol* 75: 10118-10131
3. Shirako Y and Yamaguchi Y. Genome structure of Sagiyama virus and its relatedness to other alphaviruses. 2000 *J Gen Virol* 81: 1353-1360
4. Yamaguchi Y and Shirako Y. Engineering of a Sagiyama Alphavirus RNA-based transient expression vector. 2002 *Microbiol Immunol* 46: 119-129

주 의

1. 이 보고서는 농림부에서 시행한 농림기술개발사업의 연구보고서입니다.
2. 이 보고서 내용을 발표할 때에는 반드시 농림부에서 시행한 농림기술개발사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니됩니다.