東海大學環境科學與工程研究所

碩士論文

台灣地區水體環境中腺病毒之親源分析

Phylogenetic Analyses of Human Adenoviruses (HAdVs) In Sewages of Taiwan Area

研究生 ： 施孟欣 Meng-Hsin Shih

指導教授 ： 郭獻文 Hsion-Wen Kuo

中華民國一 0 二年八月

August 2012

致 謝

「天將降大任於斯人也，必先苦其心志，勞其筋骨，餓其體膚，空乏其身。」短短的幾句話，道盡了碩士生涯這兩年來的心路歷程，兩年前的我至今日，經歷了大大小小的事情，不論是人、事、物都徹底了改變我的看法和心態，過程中的挫折與眼淚已成為了最寶貴的經驗，這本論文的完成，算是人生當中一個重要的里程碑。本論文能夠順利的完成，最感謝的還是我的指導老師—郭獻文老師，老師的積極態度、細心指導，無一不讓我深感敬佩，尤其老師更是盡其所能，適時的提供我們研究上的專業資源，讓我受益無窮。還要感謝的就是成功大學醫學院醫技所的王貞仁老師，王老師在病毒與專業檢測上的諸多幫忙，才能使學生的論文順利完成。另外要感謝的是這次也參與論文口試的黃啟裕老師，黃老師提供了專業的實驗室，讓我可以順利完成研究。此外亦感謝成功大學環工所的鄭幸雄老師，撥冗參與學生論文口試，並提供諸多寶貴意見，使論文可以更加完整。

研究的過程中，可以如此順利，其背後當然有著諸多的支持、鼓勵與協助，在這邊要先感謝志鴻學長，在研究剛開始的階段協助我有關生物方面的技術與資訊，讓我做實驗能夠得心應手;還有實驗室裡的百淳、昱智、學弟佳昆、良誌，祿翔及學妹宥鈞，大家陪我共度了實驗室悶煩的苦日子，感恩。除此之外還有管理實驗室的好朋友小葳，時常陪我去採樣，吃吃喝喝，助理美華，時常幫我解決技術上的困惑;當然還有已經畢業的玥禎和斯婷，因為有涂幫的陪伴，讓我覺得研究所的日子不孤單。另外，也謝謝實驗室已畢業的專題生雅雯與佩玲，幫了許多忙，是協助我完成研究不可或缺的一環，還有系辦的阿美學姊、美燕學姊、美君學姊及雅琪學姊，感謝你們協助我解決了實驗室的種種問題，大家的溫暖及互相扶持，融洽的氣氛成為了我進步的動力。最後，感謝愛我的家人與神明，讓我一路走來克服了許多的阻礙，闖過了種種的難關，擁有了美好的回憶。

孟欣 于台中 2010,8,13

ABSTRACT

The Human adenoviruses (HAdVs) have been identified as an etiologic agent causing gastroenteritis, conjunctivitis, and upper respiratory infection diseases in susceptible individuals or localized/ immunocompromised patients. The HAdVs currently have 52 serotypes classified into seven species, from A to G. Some studies in Taiwan have detected enteroviruses in spas, drinking water, hot tubs, and swimming pools of Taiwan area. However, the distribution of HAdVs in Taiwan has not been fully understood. The foremost objective of this study was to survey geological dispersal of HAdVs in sewages of Taiwan area based on phylogenetic analysis of partial hexon genes and quantification of real-time PCR assay. The results showed that nested PCR successfully amplified HAdV hexon genes (~143 bps) from all the 16 sewage samples. Sequencing results demonstrated HAdV-F41 was the dominant serotype (~87.8% and ~99.8% for cloning and high-throughput approaches, respectively) while rests of revealed sequences (~12.2 and ~0.2% for two sequencing methods) were belonged to C2, C6, and D8 serotypes. Serotype F41 was detected in 14 sewage samples. Based on Unifrac analyses, the seasonal difference had more significant effects on HAdVs distribution in Taiwan (80.8% and 68.3% variation for cloning and high-throughput approaches, respectively) than geological variation (7.2% and 10.1%). Furthermore, Quantitative results of real-time PCR assay for all six adenovirus species (A-F) were ranged from 8.8×10^2 to 1.7×10^8 viral particles/L; and, concentrations of serotypes F41 were from 7.9×10^2 to 1.8×10^8 viral particles/L. This study demonstrated sewages in Taiwan contained significant amount of HAdVs (especially species F and C).

中文摘要

人類腺病毒(Human adenoviruses, HAdVs)為環境中常見的致病源之一，目前由 52 個血清型(serotypes)腺病毒分屬於 7 個屬種(species)，分別是 A-G。此病毒每年在臺灣地區因腺病毒所感染的疾病(如結膜炎、上呼吸道感染、及急性腸胃炎等)都有相繼被國內新聞報導。過去在台灣的研究結果顯示曾於溫泉、游泳池、及大眾浴池中檢測出腸病毒，然而目前並不能清楚地說明腺病毒在臺灣地區地理親源上的分布狀況。本研究主要針對台灣四個代表性區域(北、中、南、及東部)之污水廠中的民生污水於四季當中以親源分析進行腺病毒基因之鑑定，並探討水質與 HAdVs 之相關性。本研究以 hexAA1885/1913 及 nehexAA1893/1905 引子對進行巢式聚合酶連鎖反應(nested polymerase chain reaction)所放大之 hexon 基因片段(約 143 bps)以核酸定序與高通量定序法探討水體腺病毒的基因類別。此外，亦使用定量聚合酶連鎖反應(real-time polymerase chain reaction)，針對環境水體中腺病毒的所有血清種及單一血清種 F41 進行定量分析。研究結果顯示，血清種 F41 在台灣環境水體中的分佈趨勢是佔優勢的腺病毒屬種 (針對克隆與高通量定序法分別約佔 87.8% 和 99.8%)，於四季中的每個區域皆可鑑定之，並於北部區域的水樣中發現血清種 C6 出沒在秋季中，C2 在經過兩次鑑定後，才可從冬季與春季南部裡明顯的被鑑定出些許的分佈，而少數的 D8 則罕見地出現在東部的春季水樣中，另外，在中台灣的冬季中經過高通量的定序後才發現仍有少量 A 病毒。為了更能清楚地觀察其季節或區域兩者是否影響腺病毒基因型態的分佈，由統計工具 UniFrac 分析顯示，季節性因子的變異數(81%及 68%)大於區域性因子的變異數(7%及 10%)，可推斷季節性的影響較為顯著。定量的結果可算出每一升的水樣中所濃縮的病毒顆粒，在 16 個水樣中分析所有腺病毒血清種的數量為 8.8×10^{2} ~ 1.7×10^{8} (腺病毒顆粒/公升廢水)，而 F41 的數量為 7.9×10^{2} ~ 1.88×10^{8}。本研究結果證實相當數量的人類腺病毒存在民生廢水中，並有季節性的分佈趨勢。

Table of contents

LIST OF FIGURES

LIST OF TABLE

Chapter 1 Introduction

1.1 Water safety perspective and pathogenic viruses

Municipal water treatment techniques developed 95 years ago have used chlorine compounds to disinfect drinking water. Ironically, for the removal of water contaminants, although humans are continually exposed to a vast array of microorganisms in the environment, only a small portion of these microbes are pathogens (Grabow et al., 2001; Van Heerden et al., 2003). However, recent studies had identified a wide range of microbial problems for community drinking water supplies, delivered and store-bought bottled water, water coolers, water lines in dental offices, shower heads, air conditioning systems, ship holds, home plumbing systems, and even under-sink reverse osmosis water treatment systems purchased. Over the last 20 years, microorganisms have caused 75 percent of all waterborne illnesses associated with water.

The three primary groups of pathogens are virus, bacteria, and protozoa (Van Heerden et al., 2005). Pathogenic viruses transmitted by water usually are of human origin. Viruses are small enough to pass through the filters used in water treatment. Besides, they are believed to be more of a problem than enteric bacteria and parasites in groundwater because they can travel further distance through soil and survive longer. They are usually excreted in high numbers in the feces of infected individuals and may directly or indirectly contaminate drinking water. They are commonly found in domestic wastewater even after disinfection and may discharge into environments. In environments, they can survive for months under the right conditions.

The enteric viruses include the enteroviruses, rotaviruses, hepatitis A and E, noroviruses, adenoviruses, reoviruses, and others (Bosch, 1998). The type of treatment will determine the concentration of pathogens and the relative risk of disposal. Viruses are present in high numbers in raw wastewater and current water treatment practices fail to ensure the complete removal of viral pathogens; consequently, viruses become environmental pollutants. The fate of microbial enteric pathogens may take many potential routes in the water environment (Fig 1.1) (Bosch, 1998). Ground and surface waters may become normally contaminated from a variety of sources including sewage treatment plant effluents, on-site septic waste treatment discharges, agricultural, natural areas, land runoff from urban and leachates from sanitary landfills. More than 80% of the outbreaks were associated with deficiencies in treatment and distribution of water. Any risk assessment is the level of concentration of contaminant is important.

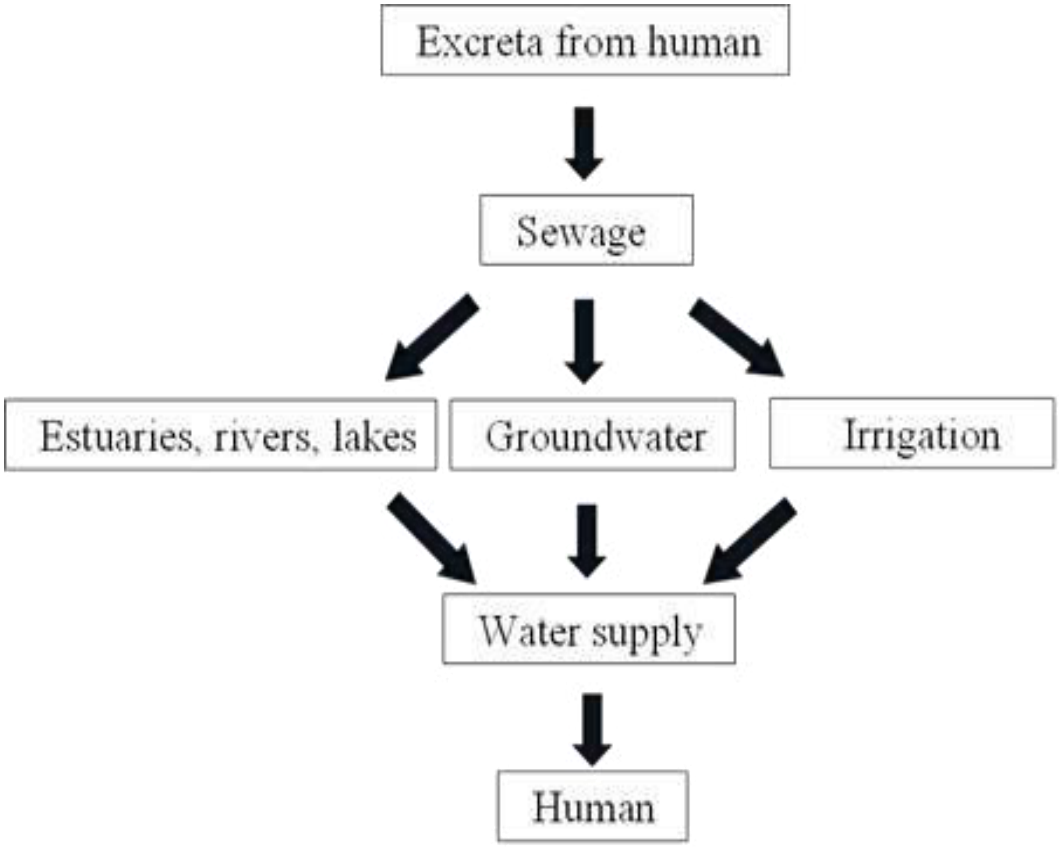


Figure 1.1 Transmission routes of enteric viruses

Human adenoviruses (HAdVs) have been found to be significantly more stable than fecal indicator bacteria and other enteric viruses during UV treatment. United Sates Environmental Protection Agency (USEPA) listed adenovirus as one of nine microorganisms on the Contaminant Candidate List for drinking water (Jiang, 2006) because HAdV removal from water treatment plant is not fully clear and HAdV can easily infect individuals via consumption of contaminated drinking water or inhalation of aerosolized droplets during water recreation. Among waterborne viruses, HAdVs are nowadays described as emerging pathogens and are considered to be highly resistant in water. Adenoviruses have been consistently found in greater numbers than other enteroviruses in raw sewage around the world and are substantially more stable than either poliovirus or hepatitis A virus in tap water and seawater. For example, adenovirus survives three to five times longer in seawater than poliovirus, taking 99 days for a 2-log reduction of infectivity.

1.2 Motivation and Objectives

High quantities of adenovirus DNA have been frequently detected in the natural aquatic environments. This is likely due to contamination of untreated or inefficiently-treated human sewage. Recently, several surveys on HAdVs occurrences in natural water bodies (e.g., river, hot spring, and sea water) have also been conducted for some northern and southern Taiwan areas. Current results suggest presences of HAdVs in such natural water bodies are notable. Several HAdV serotypes (e.g., HAdV 2, 3, 4, 7, 19, 37, 40, 41, etc.) have been identified in the water samples. Adenovirus infection is common in early childhood. Specific serotypes are often associated with particular clinical syndromes.

Adenovirus type 2, 3, and 7 had been associated with severe pneumonia and/or disseminated infection. Hospital, institutional or community outbreaks were also reported. The nationwide surveillance system of Taiwan also detected a similar trend (Tsou et al., 2012). However, a filled database about HAdV presents in environmental waters for entire Taiwan area (including most of metropolitans) has not been well established. Therefore, the overall goal of this study was to gain insight of HAdV diversity and distribution in sewage waters of whole Taiwan area. This would provide strategy maker, environmental engineer, or operator of wastewater treatment plant with further essential information about what pathogenic target needs to be removed in order to meet typical safe water criterion. This is especially important while water-reuse is required to solve the growing problems in reduce of water resources.

Specific objectives of this study were listed as follows:

(1) To indentify HAdV serotype phylogenetically for sewage samples collected from four geologically-separated areas in Taiwan during four seasons.

(2) To assess HAdV distribution in sewages as a function of variations in area and seasons using statistic analyzing tools.

(3) To quantify HAdVs concentration in the studied sewages.

Chapter 2 Literature review

2.1 Human Adenoviruses (HAdVs)

2.1.1 General information of HAdVs

Adenoviruses are members of the Adenoviridae family, which comprises five genera and infects hosts across broad spectrum of vertebrates (Benko et al., 1999). Human adenovirus (HAdV) is a double stranded DNA virus, about 70-90 nm in size, and with an icosahedral capsid (Davison et al., 2003). There are current 52 recognized unique serotypes categorized into seven subgenera (i.e., HAdV-A, -B, -C, -D, -E, -F, and -G) based on their hemagglutination properties, ontogenic potential in rodents, and DNA homology or GC content of their DNA (Jones et al., 2007). They have icosahedral capsids with twelve vertices and seven surface proteins (Benko et al., 2005). The site of entry generally determines the site of infection (Jones et al., 2007; Wigand and Adrian, 1986). Upon infection with adenovirus, one of three different interactions with the cells may occur. The first is lytic infection, which occurs when an adenovirus enters human epithelial cells and continues through an entire replication cycle, which results in cytolysis, cytokine production, and induction of host inflammatory response. The second is chronic or latent infection, the exact mechanism of which is unknown, which frequently involves asymptomatic infection of lymphoid tissue. Lastly, oncogenic transformation has been observed in rats. HAdVs are often also isolated from the pharynx and stool of asymptomatic children and now recognized as the etiologic agent of a variety of clinical syndromes (Wadell, 1999). HAdVs are associated with both sporadic and epidemic disease that results in significant economic losses and morbidity, especially amongst populations in closed settings (Gompf and Oehler, 2005).

2.1.2 Capsid composition

The HAdV particle is composed of more than 2700 polypeptide molecules, with a total mass of approximately 150×10^6 Da (Vellinga et al., 2005). Hexon is the largest and most abundant of the structural proteins in the adenovirus capsid. The 720 hexon monomers present in each virion form 240 hexon homotrimers, which in turn form 20 capsid facets, each consisting of 12 hexon homotrimers (Stewart et al., 1993); on the vertices are also the characteristic fibers (62 kDa), that range from 10 to 30 nm in different serotypes and act as attachment structures (one per cell) (Levy et al., 1994). Other two major capsid proteins, fiber and penton base, form the penton complexes (three subunits of protein IV and five subunits of protein III) at each virion vertex (Fig. 2-1). The position of the minor capsid protein IIIa has been mapped to the icosahedral edges and hexamers of protein VI are positioned underneath the penton base. Protein VI interacts with a core protein, protein V, which is located internally near the penton base in the vertex region of the virion. Here, protein IX with four trimeric densities has been found near the outer surface of the capsid, in the central cavities between the hexon capsomers that form the GONs (groups-of-nine hexons) (Stewart et al., 1991; 1993). The minor capsid protein VIII is presumed to be located internally in the capsid. Protein IIIa: The precursor of protein IIIa has a mass of 67 kDa and is cleaved at the N terminus during maturation of the virion to generate the 63.5 kDa protein IIIa. Imaging techniques have shown that protein IIIa is present as an elongated. Evidence that protein IIIa domains interact with protein VII comes from experiments showing that proteins IIIa and VII can be co-immunoprecipitated. The main density assigned to protein IIIa is located near the outer surface of the capsid (Vellinga et al., 2005), where protein IIIa may stably assemble the facets (Fig. 2-1).

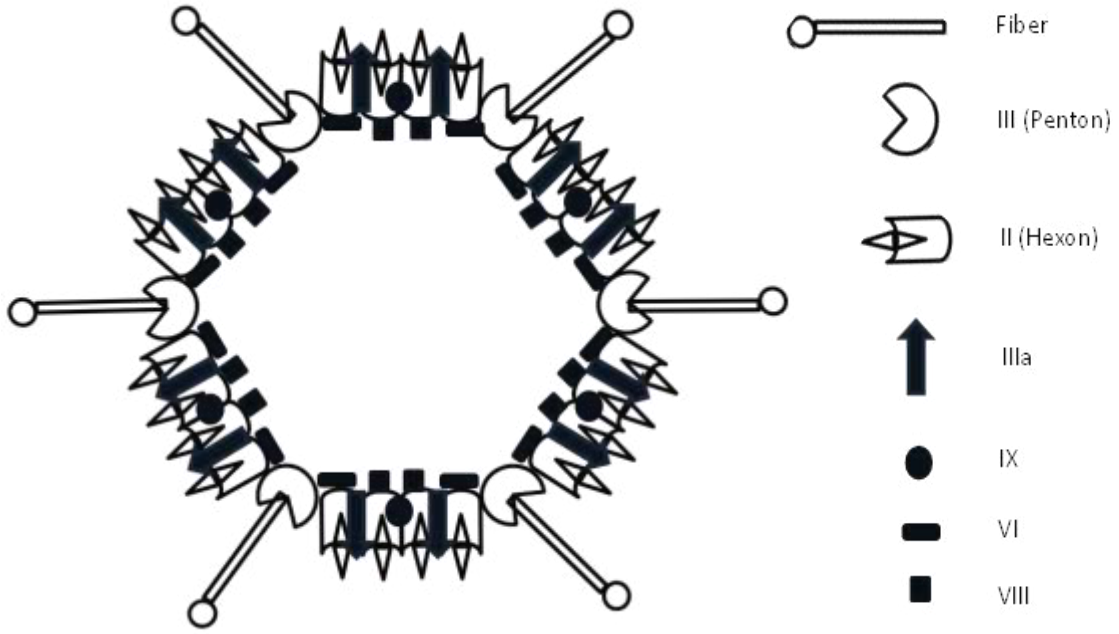


Figure 2.1 Structure of the capsid proteins showed their location within the capsid.

Protein VI: The mature version of protein VI is 22 kDa and is generated by cleavage from a larger precursor. Protein VI is positioned in the interior of the capsid, presumably adjacent to the hexons and it contains two nuclear-localization sequences (NLS) and two nuclear-export sequences (NES) (Greber et al., 1993), they are removed by proteolysis during maturation. An important function of protein VI is to facilitate nuclear import of hexon proteins. In cells infected with H2ts147, a mutant defective in protein VI, no nuclear import of hexon is observed at the non-permissive temperature.

Protein VIII: The least-studied of the minor capsid proteins is the 15.3kDa protein VIII. It is located at the inner surface of the triangular facets as dimers and interacts with hexons of adjacent facets. Whereas the precursor of protein VIII is present in empty capsids, it is undetectable in complete particles.

Protein IX: The 14.3 kDa protein IX is the smallest of the minor capsid proteins. Twelve molecules of protein IX are located at each of the 20 facets of the icosahedral capsid. With nine hexon capsomers, protein IX forms the stable assemblies that are termed GONs. These GONs form the central part of each of the facets of the icosahedral capsid (Rux and Burnett, 2004). The protein IX molecules have been positioned in the cavities between the hexon tops, where they form continuous trimeric densities. This scheme does not represent a real section of the icosahedral virion.

Two major polypeptides are associated with adenovirus cores, polypeptides V and VIII. Polypeptide VII is the major core protein and accounts for about 10% of the protein mass of the virion (Soumitra et al., 1998). It is synthesized as a precursor, polypeptide pVIII, which is cleaved during virion maturation into the mature core protein, which has an estimated molecular weight of 19.4 kDa (Soumitra et al., 1998).

2.1.3 Epidemiology

Adenovirus are a frequent cause of acute upper respiratory tract (URT) infections, they also usually causes a localized infection, but generalized infection can occur in immunocompromised patients, some adenovirus subtypes cause cellular cytopathic effect (rounded, swollen cells and basophilic intranuclear inclusions), HAdV-40 and -41 of subgroup F are unique in being responsible most cases of gastroenteritis in children (Brandt et al., 1969). HAdV- C species (serotype HAdV-1, HAdV-2, HAdV-5, and HAdV-6) cause typically 5% of acute respiratory infections among human hosts in early childhood (Garnett et al., 2002). In addition to acute disease, HAdV-C serotypes establish persistent infections in immune competent hosts characterized by intermittent viral shedding. Human adenovirus serotype 4 was originally isolated from a United States of America army trainee diagnosed with acute respiratory illness during an epidemic in the

winter of 1952 to 1953 (Hilleman, 1956). The HAdV-4a genome type is described as being second only to HAdV-8 as a cause of HAdV-associated eye infections in Japan (Cooper et al., 1993). In contrast, a different pattern of epidemiology and pathogenesis is presented by HAdV-E species (HAdV-4).

Human adenovirus species B has been divided into two sub-species B:1 and B:2, based upon tropism and restriction cleavage patterns (De Jong et al., 1999). HAdV-B species has been further classified into two subspecies, namely, B1 (HAdV-3, HAdV-7, HAdV-16, HAdV-21, and HAdV-50) and B2 (HAdV-11, HAdV-14, HAdV-34, HAdV-35), they are causative agents in epidemic conjunctivitis outbreaks (Table 2.1). Human adenovirus species B infections are reported to be the common cause of outbreaks of febrile respiratory tract infection and pneumonia in military recruits and responsible for 90% of recruit hospitalisations for pneumonia (Booth et al., 2004).

HAdV-D represents the largest class of HAdV species accounting for 35 out of the 56 recognized types, and is therefore the species most amenable to the study of adenoviral evolution. HAdV-D types have been isolated from the upper and lower respiratory, gastrointestinal, and genitourinary tracts, and the eye. Eight types (D43-D49 and D51) were initially characterized after being isolated from AIDS patients, suggesting a possible role for the immunocompromised host in the genesis of new adenoviruses (Lord et al., 2000). Thus, the majority of HAdV infections are self-limited, with severe or disseminated disease occurring sporadically in immunocompromised patients, patients with AIDS, transplant recipients, and those with underlying respiratory or cardiac disease. However, while epidemiology is a valuable tool for determining the spread patterns of diseases and establishing statistically significant associations with potential risk agents, cause and effect of the disease cannot be easily demonstrated.

Table 2.1 Classification of human adenoviruses (Jiang, 2006)

Species	Serotypes (52)	Illness
A	12, 18, 31	Gastroenteritis tract
B	3, 7	Acute febrile pharyngitis
	16	Acute/chronic infection in patients
	21	Haemorrhagic cystitis in children
	11	Follicular conjunctivitis
	14	Acute respiratory disease
	34, 35	Acute/chronic infection in patients with immunosuppression
C	1, 2	Pharyngitis
	5, 6	Acute/chronic infection in patients with immunosuppression
D	8, 9, 10, 13, 15, 17,19,20,22,29,37,42	Epidemic keratoconjunctivitis
	39, 42, 43, 44, 45,46, 47, 48, 49, 51	Acute/chronic infection in patients with immunosuppression and AIDS
E	4	Respiratory tract
F	40, 41	Gastroenteritis tract

Adenovirus attachment is mediated by the interaction between the HAdV fibre and a specific receptor on the target cell. Most HAdVs are reported to utilize the coxsackie-adenovirus receptor as their point of entry (Leen and Rooney, 2004). Adenovirus particles assemble in the nuclei of infected cells and the replication cycle takes 32 to 36 hours (Echavarria, 2009). Although fibre-deficient mutants of HAdV seem to have an altered viral morphogenesis, more recent data suggest that the fiber plays a role in capsid stabilization, rather than in capsid assembly (Hong and Engler, 1996). Three different types of infection are caused by HAdVs: a productive lytic infection predominantly in epithelial cells, and latent or persistent infections in the lymphoid mucosal tissue which may infrequently produce virus, or transformation predominantly in rodent cells. A schematic representation of a HAdV lytic cycle is presented in Figure 2.2.

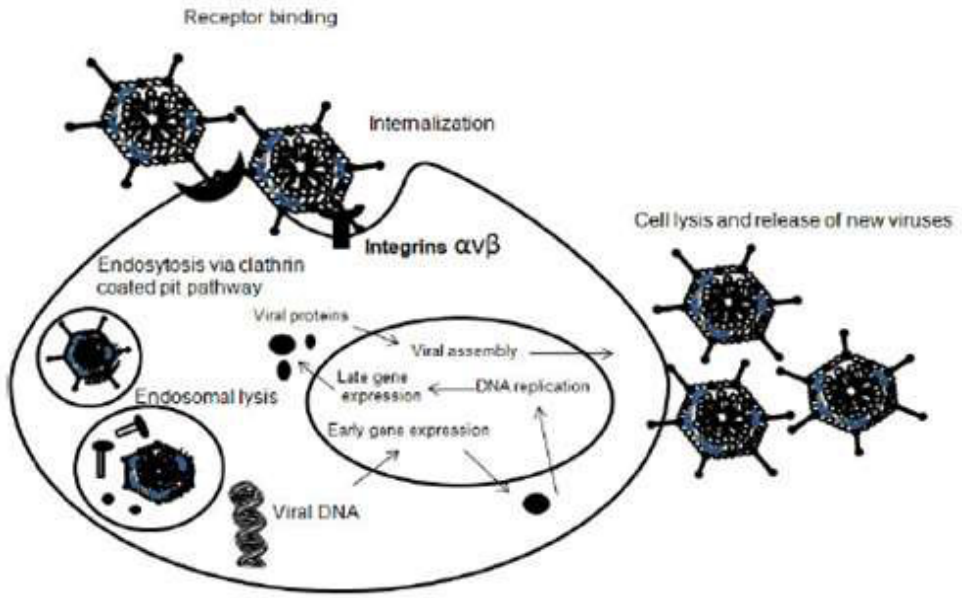


Figure 2.2 Schematic illustration of the adenovirus 5 infection cycle. (Sarkioja, 2008)

2.2 Outbreaks of HAdV

Water virology was born after a large hepatitis outbreak was declared in New Delhi between December 1955 and January 1956. The origin of the outbreak was the contamination by sewage, water virology, actually environmental virology, began with efforts to detect poliovirus in water around half a century ago (Bosch, 1998). The first water-related outbreak of HAdV infection was reported during the summer of 1977 (Dangelo et al., 1979). The waterborne enteric disease outbreak that was only attributed to HAdVs was reported in a Finnish municipality in April 1994. Some 1500-3000 people had symptomatic acute gastroenteritis tract (Kukkula et al., 1997). Outbreaks of HAdVs infection have been described in day care centers, hospitals, military quarters, and swimming pools.

On 7 March 2006, the occurrence of reported outbreaks of HAdV-associated diseases has been described in young people, especially senior high school students, in China. These cases were reported to have an acute respiratory disease that was usually caused by HAdV-3 and HAdV-7 (B1 species) (Zhang et al., 2006).

Outbreaks of pharyngoconjunctival fever caused by HAdV-3 have been studied among male persons aged 11-19 years who reside in boarding schools (Payne et al., 1984). HAdV-3 had isolated since 1954, it has been demonstrated to have a worldwide distribution. HAdV-3 isolates accounted for 13% of the 24.184 HAdV strains reported to the world Health organization in a survey of respiratory virus infections from 1967 through 1976 (Schmitz et al., 1983). Thus, outbreaks of HAdV-7 associated with respiratory disease have been reported worldwide and commonly occur among the military trainees. The occurrence during March-April 2011 and involved new police recruits in the Kuala Lumpur Police Training Centre, approximately 100 trainees were

admitted the Kuala Lumpur Hospital (Yusof et al., 2011). The case affected 851 police trainees and claimed 3 lives. The B1 viruses(HAdV-3,-7,-16,-21, and -50) are usually associated with respiratory tract infections, while the B2 viruses (Li et al., 1991), except for HAdV-11 and HAdV-14, are associated with Kidney and urinary tract infections. Clinical specimens were collected from the patients and healthy people in China of 2006 which concluded that the outbreak of respiratory disease in the Qishan County, Shanxi Province, China, was caused by HAdV-11 that belongs to species B2 (Mei et al., 2003).

Between October 2009 and July 2010, nine cases of HAdV-14 infections were identified at different locations in Ireland. Outbreaks of HAdV-14 associated with acute respiratory disease of variable severity were subsequently detected in US military bases and in civilian populations in Washington, Oregon, Alaska, Wisconsin and Pennsylvania (Metzgar et al., 2007). The community in Oregon resulted in 29 hospitalisations and seven deaths (Lewis et al., 2009). In a military base involved in high rates of transmission of HAdV-14 infection sustained over five months and was associated with 23 hospitalisations and one death, in US (Tate et al., 2009). The Alaskan outbreak in 2008 involved 46 confirmed of HAdV-14 infection, of whom 11 were hospitalized and one died (Centers, 2008).

2.3 Researches of human enteric viruses in Taiwan

2.3.1 Viruses detected in hot spring of Taiwan

Numerous studies have documented the presence of enteroviruses in water environments worldwide (Vivier et al., 2004). A particular risk factor has recently been identified in the use of hot tubs, spas and public baths (Tolentino et al., 2004). In Taiwan,

hot spring facilities have become an increasingly popular form of recreation. Millions of people in Taiwan enjoy bathing in hot spring tubs. Enteroviruses may potentially be transmitted by recreational activities in polluted waters (Hsu et al., 2008). Hsu et al. (2008) detected enteroviruses in 13 of the 34 samples (38.2%), and at four of the six spring recreation areas. The most frequently detected serogroups were Coxsackie virus A2 (26.5% of all samples), followed by echovirus 11 (5.9% of all samples) at hot tubs in Taiwan. The percentage occurrences of enteroviruses in spring recreation areas are higher than the values obtained by the above authors. Hot tubs, swimming pools and spas were generally set up in the spring recreation areas of Taiwan (Hsu et al., 2008).

2.3.2 Viruses identified from drinking water in Taiwan

Numerous studies have documented the presence of enteroviruses in a wide range of water types year round, although the results of these studies have also found that enteroviruses are more prevalent during rainy seasons (Cronin et al., 2003; Percival et al., 2004; Jean et al., 2006). Hsu et al (2009) showed that enteroviruses were detected in various drinking water and seven of all samples (30.4%); therefore, water quality is somewhat affected by human activities in the upper regions of watersheds that contain inadequate sewer systems. The occurrence of enteroviruses in different water types and sampling sites varies. Analysis of our nucleotide sequence data identified three major clusters, which grouped with enterovirus serotypes: Coxsackie virus A2, Coxsackie virus A6, and EV71. From reservoirs of drinking water, enteroviruses were found in collected samples (Hsu et al., 2009).

2.3.3 HAdV identification in clinical specimen of Taiwan

In Taiwan, adenoviruses were the second most common viral agent (4%), after RSV (7%) and followed by EV (3%), found in children hospitalized with respiratory infections

in 1997 (Tsai et al., 2001). A total of 524 adenovirus isolates recovered from children with respiratory infections from 1981 to 2001 were studied. At least seven HAdV serotypes (Ad1, 5, 6, 2, 3, 4, 7, and 14) were observed during this two decade survey of respiratory adenovirus in southern Taiwan (Lin et al., 2004). According to the nationwide surveillance data from the Centers for Diseases Control of Taiwan genotype 3 accounted for 35% to 58% of yearly adenovirus isolates selected for genotyping analysis from 2008 to 2010 (Tsou et al., 2010). Both serotype 3 and serotype 7 were involved in this outbreak, but fatal cases were more frequently infected with serotype 7 (70%). The mortality rate for serotype 7 and non-serotype 7 were 32% (7 out of 22) and 14% (3 out of 22), respectively (Chen et al., 2013). Serotype 3 was identified as the major causative agent responsible for an outbreak of respiratory infection in northern Taiwan from November 2004 to February 2005. This outbreak was likely derived from the community, as the majority of the patients (87.2%) were outpatients (Chang et al., 2008). HAdV-3 of species B is frequently associated with respiratory tract infections in pediatric patients and may cause severe infections in newborns and infants (Cherry et al., 1998).

2.4 Virus sampling and concentration

Some studies that investigated the filtration of viruses from samples were based on the work of Elford (1931) in which viral adsorption to collodion membranes was investigated. In 1953, Kelly was able to show that ion-exchange resins were successful at concentrating viruses from sewage samples. Chang was able to remove viruses from river water using flocculation procedures (Chang et al., 1958). Later on, the 15th edition of Standard Methods for the Examination of Water and Wastewater was published and included methods for virus concentration and detection from water samples (Greenberg and Taras,

1981). In July of 1997, the U.S. Environmental Protection Agency began an 18-month study as part of the Information Collection Requirement (ICR), a ruling which required the agency to collect data as part of a national research project to support development of national drinking water standards to protect human health (USEPA 1996).

For viral monitoring, the ICR was the largest national occurrence study ever undertaken in the United States. The development of VIRADEL or virus adsorption/elution methods is credited to Wallis and Melnick (1967) and involves the adsorption of viral particles to the filter media by charge interaction and subsequent elution of the virus by a pH-adjusted solution. The most common elution solution typically contains beef extract, solutions containing amino acids and salts have also been used to remove adsorbed virus from the filter (Farrah et al., 1979; Chang et al., 1981). The first electronegative filters to be used for virus concentration were built upon studies using membrane filters to separate viruses from crude cell extracts. Farrah et al. (1976) showed that >90% of seeded virus were adsorbed to filters when the sample pH was 4 or less and that recoveries varied from 40 to 67% in seeded tap water samples. Designed in a cartridge-type format, the 1 MDS filter (Cuno, Meriden, CT, USA) has been a popular electropositive filter choice for many years in the environmental virology field. Early studies on poliovirus with the 1 MDS filter showed recoveries similar to that of the electronegative filter in spiked tap water samples. In an effort to offset the cost of the 1 MDS filter, it was shown that the 1 MDS filter could be washed and re-used (Cashdolla and Dahling, 2006). The advantages to most electropositive filters are that they are easy to use, with no preconditioning of the water samples required, and because of the cartridge format, are able to filter large volumes of water (>1000 liter) at high filtration rates without clogging in most cases (Cashdollar and Wymer, 2013) (Table 2.2).

In spite of the importance of research on wild viral populations and their dependence on concentration methods, existing large-scale concentration methods are inefficient, costly and variably reliable. While it is possible to collect viruses from natural waters using impact filtration onto <0.02 μm pore-size filters (Steward and Culley, 2010), the low filtration speed and rapid clogging of these filters render this approach only useful for filtering smaller sample volumes. Shields et al (1985) focus on chemical techniques to develop a virus concentration method suited to marine virus research applications, by adapting flocculation based wastewater treatment techniques and polyelectrolytes have been used to efficiently flocculate and remove viruses from wastewater.

Table 2.2 Filter type pros and cons

filter type	pros	cons
Electronegative 1MDS	High recoveries for commonly tested enteroviruses	Requires preconditioning of water sample or filter prior to filtration
Electronegative	Has been tested with an array of enteric viruses	Extremely high cost per filter
Electronegative NanoCeram	1. Economical 2. No preconditioning of water sample required	Clogs in more turbid waters
Electronegative Glass Wool	1. Easy to use 2. Field-deployable	Turbid water may cause clogging
Electronegative ViroCap	1. Easy to use 2. Field-deployable	May be limit to volume that can be filtered due to filter`s size
Hollow-Fibre Ultrafiltration	1. Multi-pathogen concentration 2. Economical	Not easily field-deployable Slow filtration rate

2.5 Molecular detection

Multiplex PCR-based assays have been developed for detection and identification by amplicon size of selected HAdV serotypes or species. PCR, however, is rapidly emerging as molecular method for virus detection in environmental samples. Nested PCR amplification was applied to ensure the specificity of detection, eliminate any false-positive results, and increase the amplification signal, providing the method with the highest sensitivity (Puig et al., 1994), which allows us to detect a small number of viral contaminants in environmental samples. The PCR enabled HAdVs to be detected in blood (Kampmann et al., 2005), swimming pool water (Papapetropoulou et al., 1998) and river waters (Castignolles et al., 1998). PCR has been successfully applied for the detection of HAdV in a variety of samples and specimens (Reither et al., 2007). Only the nested PCR showed a higher level of sensitivity of detection of these viruses. One of the most serious problems related to the use of the nested PCR technique is how to avoid the false positives easily obtained by contamination with amplified DNA (Ruuskanen et al., 1994).

The PCR method makes it possible to obtain information about the presence of viral contaminants in environmental samples within a few hours. The methods may provide for the detection of < 10 particles of HAdV, which is 100 to 1000 times higher than the sensitivity of cell culture, gives new comparative data about the presence of adenoviruses and in wastewater, and shows a high prevalence of adenoviruses in sewage and in river water samples (Pina et al., 1998). It must however be emphasized that detection of HAdVs in environmental samples by PCR assay indicates the presence HAdV DNA and does not provide any information on infectivity of the virus, which is directly related to the human health risk (Jiang et al., 2001). PCR is particularly useful for the detection of viruses, such as AMDV, that do not propagate well in cell culture.

2.6 Real-time PCR quantification

Fairly recently, a new method of PCR quantification has been invented. This is called real-time PCR because it allows the scientist to actually view the increase in the amount of DNA as it is amplified. Several different types of real-time PCR are being marketed to the scientific community at this time, each with their advantages. First, we must take a look at the TaqMan probe. The probe consists of two types of fluorophores, which are the fluorescent parts of reporter proteins (Fig 2.3). While the probe is attached or unattached to the template DNA and before the polymerase acts, the quencher (Q) fluorophore reduces the fluorescence from the reporter (R) fluorophore (He et al., 2005). It does this by the use of Fluorescence Resonance Energy Transfer (FRET), which is the inhibition of one dye caused by another without emission of a proton. The reporter dye is found on the 5′ end of the probe and the quencher at the 3′ end. Gu et al., (2003) also reported real-time PCR quantification of adenovirus in a patient`s stool samples. By using control viral samples, sensitivity of detection was demonstrated to be fewer than 10 copies of viral genome per reaction and quantitative linearity was demonstrated to be from 10 to 10^6 copies of input viral DNA. Furuta et al. (2001) described real-time PCR has been demonstrated to be a powerful tool for nucleic acid quantification in water environments, not only to its sensitivity and specificity but also to the fact that, relative to other quantitative tools, it is easy to use.

The use of qPCR is a rapid, sensitive, specific, and quantitative method of detecting viral genomes in low concentrations in water, but it also has a few major drawbacks as follows. First, PCR (including qPCR) alone cannot distinguish between infectious and noninfectious viruses (Niesters et al., 2000).

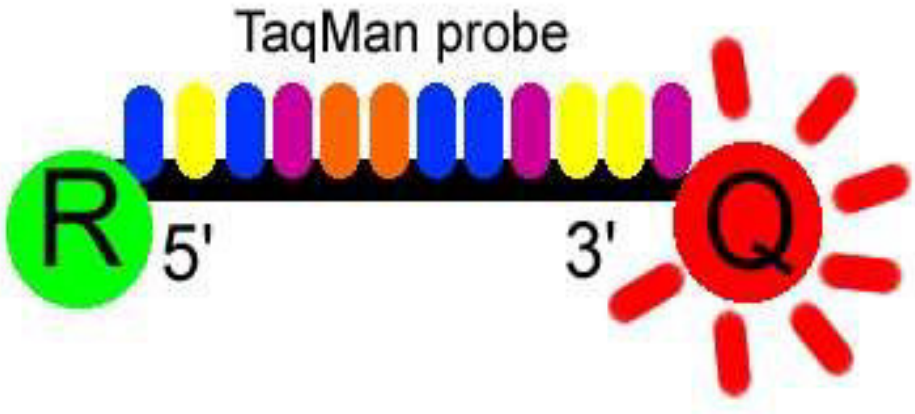


Figure 2.3 The Taqman probe: the right circle represents the quenching dye that disrupts the observable signal from the reporter dye.

2.7 High throughput sequencing

The Sanger method has been the golden standard for DNA sequencing for the last few decades (Sanger, Nicklen & Coulson 1977). The growing interest for sequencing with higher throughput and lower costs in the last decade has led to development of new sequencing applications. The currently leading platforms are Genome Analyzer (Illumina Inc., San Diego, CA, USA), Solid (Applied Biosystems, Foster City, CA, USA) and 454 (454 Life Sciences, a Roche company, Branford, CT, USA). This Master's thesis concentrates on optimizing DNA library preparation for Illumina Genome Analyzer II sequencer with the amount of sample sufficient for sequence capture. However, the sample library steps described here are partially adaptable also for 454 and Solid. Genome Analyzer II can utilize nucleic acids from various resources if it is provided in a compatible format (de Magalhães, Finch & Janssens 2010, Schmidt et al. 2009, Linnarsson 2010).

The library consists of short fragments of double-stranded DNA. These fragments should be approximately 150 to 500 base pairs long to have uniform performance in cluster amplification. The fragments have adapters on both ends of the fragment complementary to both adapters on flow cell and sequencing primers (Linnarsson 2010, Pettersson, Ansorge 2009, Mardis 2008). The next-generation sequencing platforms such as HiSeq 2000 (Illumina) are aiming at the whole genome sequencing in a single run of the instrument. However, the Genome Analyzer II does not yet have such capacity. This compels to the selection of regions of interest and their enrichment in order to keep the sequencing costs and resources in a feasible level. At present selected regions of the genome can be captured using microarray or a solution phase method (Mamanova et al. 2010).

2.8 UniFrac analyses

A phylogenetic tree can be analyzed using multivariate statistical tools for large and complex data sets (Ramette, 2007). The basic aim of multivariate, or exploratory analyses, is to represent the (dis)similarity between objects based on values of multiple variables associated with them. Multivariate analyses are useful to reveal patterns in large data sets with many interacting variables, but do not directly explain why those patterns exist (Ramette, 2007). Traditionally, multivariate analyses are not able to incorporate phylogenetic data. However, UniFrac, a program used in this study, is a suite of tools used that can run multivariate analyses on phylogenetic data.

The UniFrac interface provides a suite of tools for the comparison of microbial communities using phylogenetic information. It takes as input a single rooted phylogenetic tree that contains sequences derived from at least two different

environmental samples and a file describing which sequences came from which sample (Fig 2.4). When comparing microbial communities, researchers often begin by determining whether each pair of communities is significantly different (Lozupone and Knight, 2005). In order to understand how communities differ, it is essential to move beyond pairwise significance tests. For example, we might want to know whether differences between communities stem primarily from particular lineages of the phylogenetic tree, or whether there are environmental factors (such as season, time, or temperature) that group multiple communities together (Lozupone et al., 2007a). Unifrac uses phylogenetic information and multivariate statistical techniques to determine whether microbial communities are significantly different, identify individual lineages that contribute to these differences, and reveal broad patterns relating many environmental samples. The UniFrac web application allows researchers to compare many environments simultaneously using hierarchical clustering and principal coordinates analysis (PCA) (Lozupone et al., 2007b).

UniFrac displays the uploaded data on a page that allows the user to choose among several analyses. A text representation of the input phylogenetic tree is displayed, including the environments to which each sequence was observed. Unifrac can calculate whether two communities differ significantly using either the UniFrac significance test or P test. The P test estimates similarity between communities using the number of changes from one environment to another along a branch that is required to explain the distribution of sequences between the different environments in the tree (Martin, 2002). The significance test measures similarity between communities as the fraction of branch length in the tree that is unique, meaning that it leads to descendants in one environment or the other but not both (Fig 2.5).

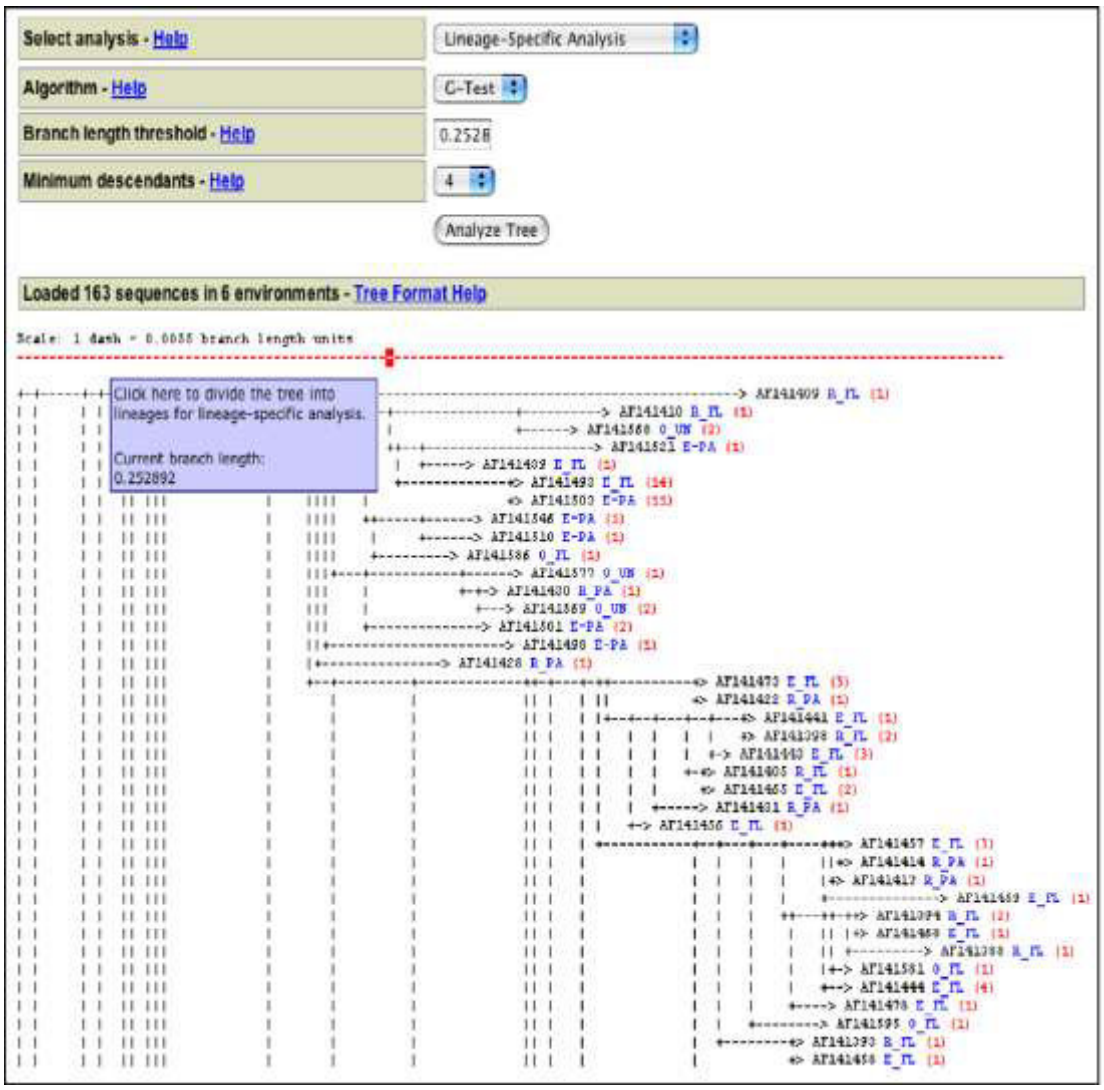
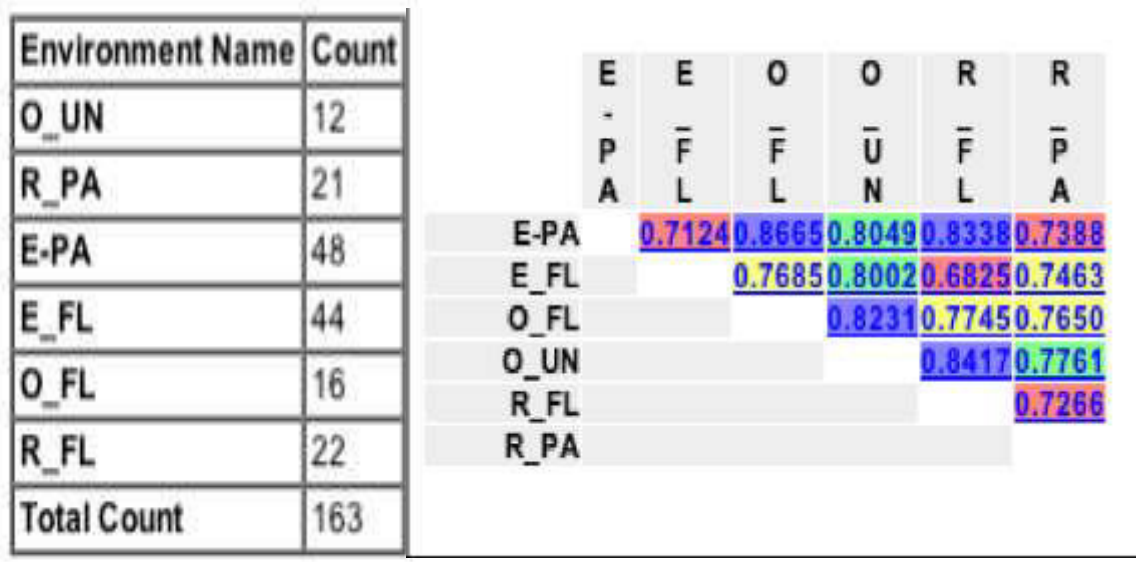


Figure 2.4 Analysis showed that is displayed after loading phylogenetic tree and environment file.

The PCA option performs Principal Coordinates Analysis. PCA is a multivariate statistical technique for finding the most important axes along which your samples vary. Distances are converted into points in a space with a number of dimensions one less than the number of samples. The principal components, in descending order, describe how much of the variation each of the axes in this new space explains. The first principal component separates out the data as much as possible (Dunfield et al., 2004; Stach et al., 2003;

Schadt et al., 2003), the second principal component provides the next most separation, and so forth. The UniFrac interface returns information on all principal component axes in a data table. It also allows easy visualization of that data in scatterplots that compare all pairs of the first three principal components (Fig 2.6). The points are either marked with text or colored symbols. Symbols representing environments with similar names are assigned the same color, providing a convenient way to group environments visually by changing the names in the uploaded file.

The web interface also provides an environment familiar to most potential users, and allows the results to be visualized as trees or graphs directly rather than requiring that the data be exported to other programs. Unlike our earlier command-line implementation, the user need not install additional software to run the web version of UniFrac, and the analysis is not limited to a single CPU.



Environment Name	Count
O_UN	12
R_PA	21
E-PA	48
E_FL	44
O_FL	16
R_FL	22
Total Count	163

	E-PA	E_FL	O_FL	O_UN	R_FL	R_PA
E-PA		0.7124	0.8665	0.8049	0.8338	0.7388
E_FL			0.7685	0.8002	0.6825	0.7463
O_FL				0.8231	0.7745	0.7650
O_UN					0.8417	0.7761
R_FL						0.7266
R_PA						

Figure 2.5 Analysis results for sequence and comparison of the UniFrac significance test and P test

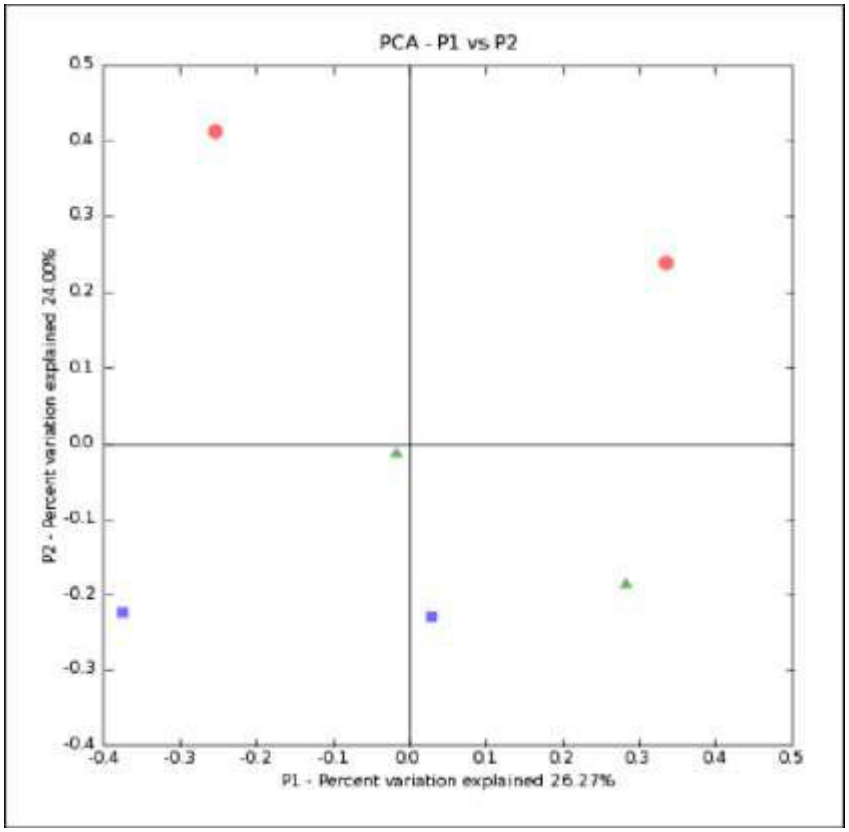


Figure 2.6 **Scatter Plots**: Will plot the following principal components against each other for each environment: P1 vs P2, P1 vs P3, and P2 vs P3.

Chapter 3 Materials and Methods

This study was aimed to assess HAdVs distribution in sewage samples that were collected from four water recycle centers during four seasons. Target HAdVs hexon gene was PCR amplified from extracted viral DNA and further sequencing via both cloning and high-throughput approaches. Sequencing results were analyzed using phylogenetic and statistic tools. Overall research approach was showed in Figure 3.1.

3.1 Sample collection

The aim of this study was to survey HAdVs geological dispersal in Taiwan. This study chose untreated domestic wastewater (i.e., raw sewage) for investigation of HAdV distribution in Taiwan. A total of 16 raw sewage samples were collected from four wastewater treatment plants (North-Neihu, Central-Futien, South-Anping, and East-Hualien, Fig 3.2) for autumn, winter, spring, and summer seasons during years 2011 and 2012 based on virus absorption-elution method suggested by US EPA. The viral sample collection device was set up according to US EPA Manual of Methods for Virology (Fig 3.3). It was recommended that one liter of highly contaminated waters such as sewage to 1,800 liters of drinking waters should be passed through a 1MDS cartridge filter. Each of the wastewater samples was pumped through a 1MDS filter at a flow rate of 3 gal/min and a back-up pressure less than 30 PSI. When the sampling pumped pressure reached 30 PSI, the raw sewage pumping was stopped. Sampling volume together with basic sewage conditions (pH, temperature, MLSS, and BOD) were recorded on the sample data sheet. All filters with viral particles were held at 4ºC and eluted within 72 hours of the start of the sample collection.

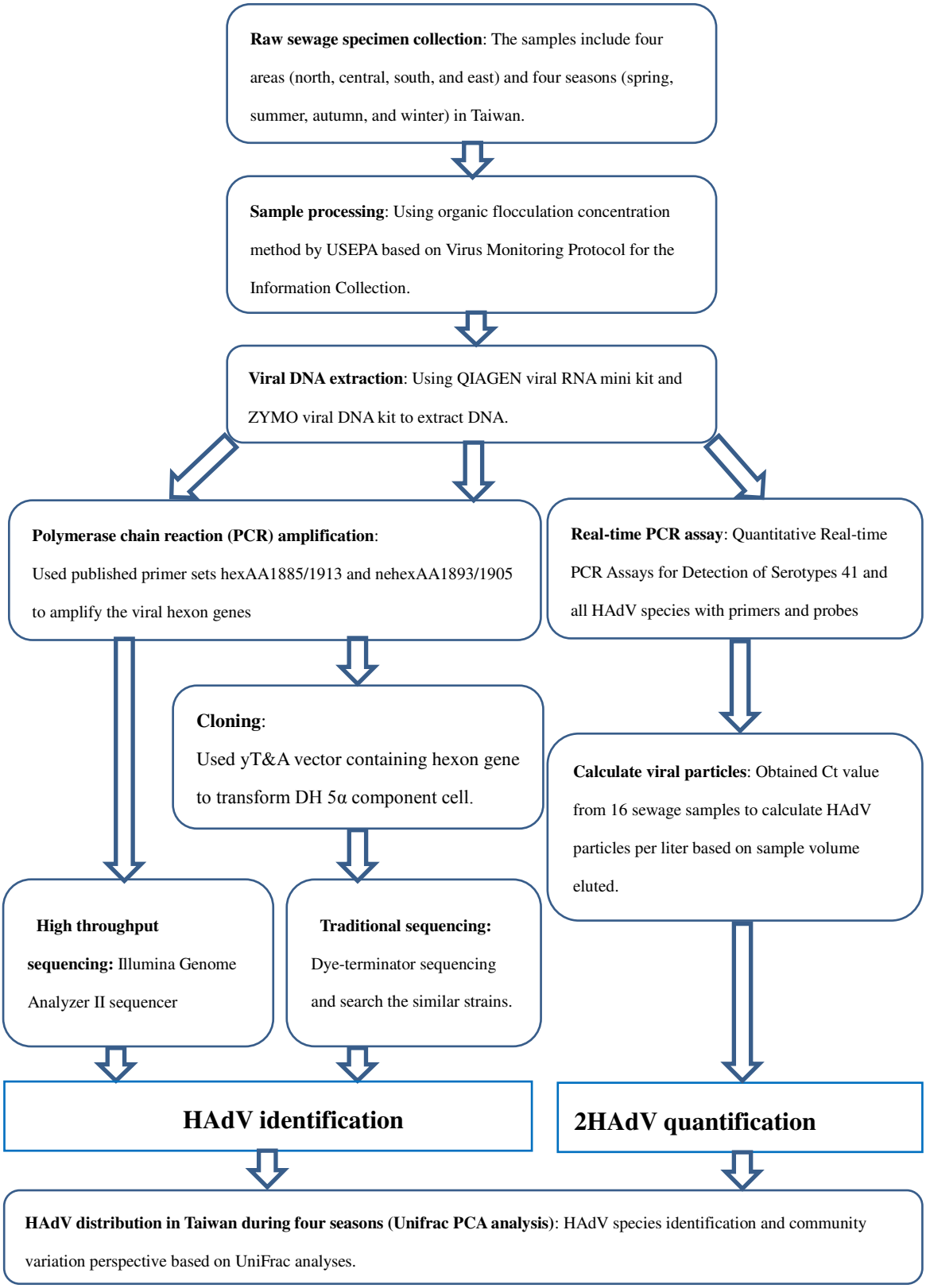


Figure 3.1 Research approach

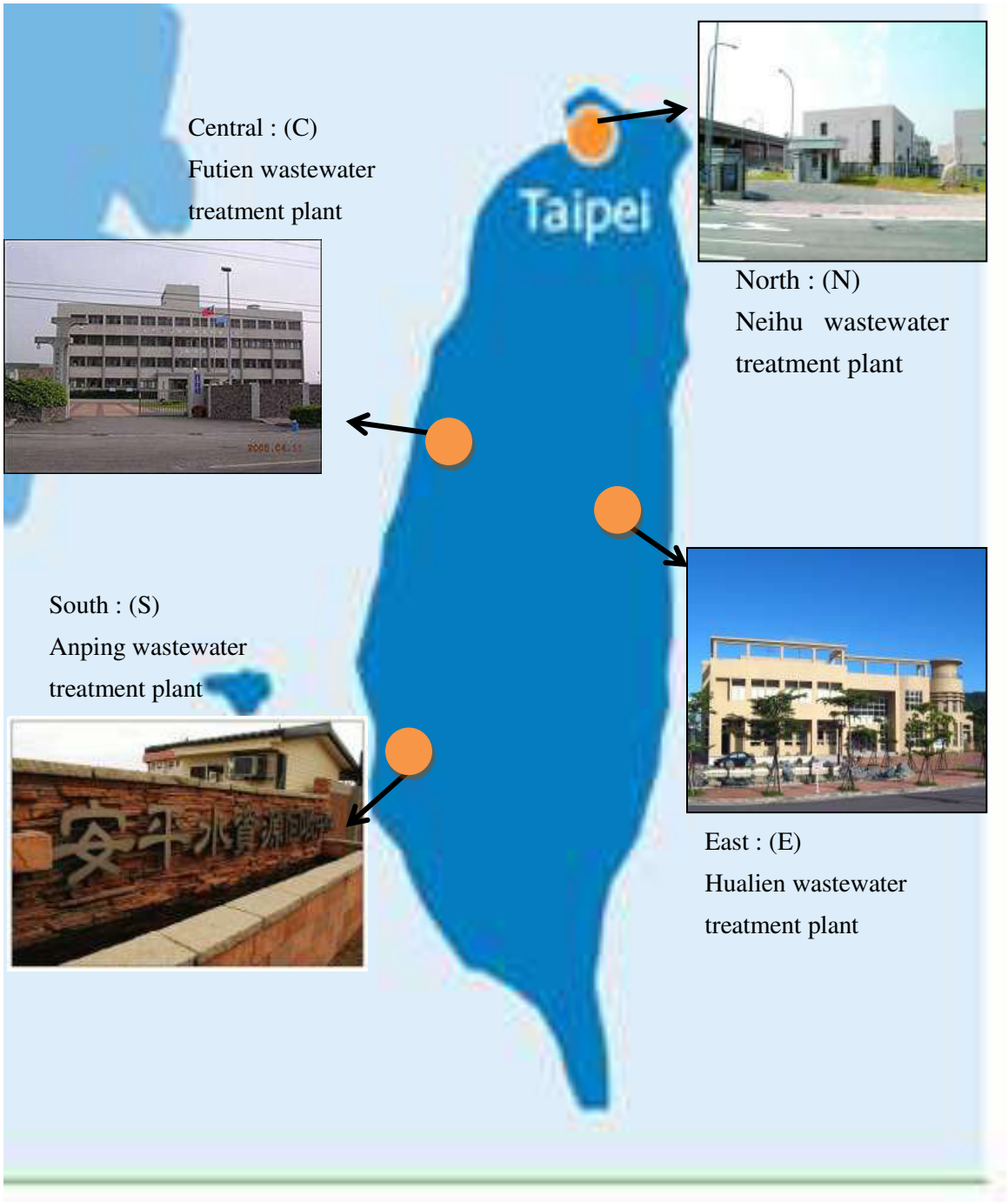


Figure 3.2 Sampling from four wastewater treatment plants

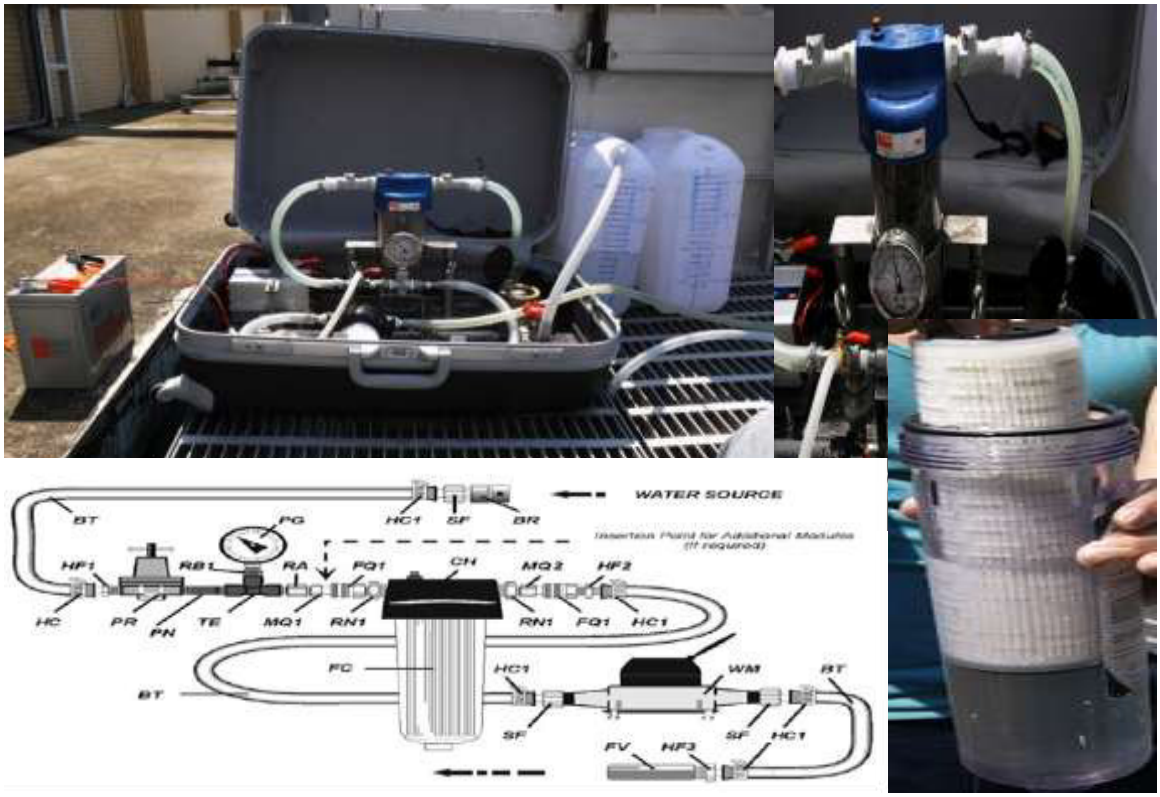


Figure 3.3 Sample collection device and 1MDS filter based on the standard filter apparatus

Figure 3.4 Picture showed the processing of organic flocculation concentration

3.2 Sample processing

Sample elution was based on Virus Monitoring Protocol for the Information Collection Requirements Rule from US EPA. Further concentration of elutes was critical since the viruses might present at low numbers. It was necessary to reduce the volume of the sample to be assayed to a few milliliters. The 1 MDS filters was merged of with 1 liter beef extract (30g of beef extract powder, 7.5 g of glycine, pH9.0-9.5) to elute the viral particles for 2 minutes (Fig 3.4). The concentration solution was flocculated by lowering the solution pH to 3.5 for 30 minutes using 1N HCl. The precipitates flocculated were collected by centrifuged with 5000 g at 4℃ for 15 minutes. Suspended the precipitates by 30 ml sodium phosphate (final pH 9.0-9.5), the precipitates were centrifuged again at 7500 g for 10 minutes. In the end, the supernatants were obtained and discarded the precipitates, adjusted pH to 7.0-7.5 by 1N HCl, filtered through sterilized 0.45 um and 0.22 um filters, and stored in tubes at -80℃.

3.3 Nucleic acid extraction

Viral nucleic acid extraction kits provide the fastest and easiest way to purify viral DNA for reliable use in amplification technologies. Viral DNA can be purified from plasma (treated with anticoagulants other than heparin), serum, and other cell-free body fluids. Samples may be fresh or frozen, but if frozen, should not be thawed more than once. Repeated freeze-thawing of plasma samples will lead to reduced viral titers and should be avoided for optimal sensitivity.

In this study, the viral nucleic acid was directly extracted from the concentrated specimens by using a QIAamp mini-viral RNA extraction kit (Qiagen) and a Zymo viral

DNA extraction kit (Zymo research). Briefly, the concentrated sample was added to Buffer AVL by vortexing 15 seconds and incubated at room temperature for 10 minutes. Briefly Centrifuged to remove drops from the inside of the liquid. The ethanol (96-100%) was added 560 µl, and mix by vortexing for 15 seconds. After mixing, briefly centrifuge the tube to remove drops from inside the lid. After centrifugation of 1 minute at 8000 g for spin column, the supernatant was transferred to a fresh tube and an equal volume of Buffer AW1 and AW2 was added for washing and mixed well. Finally, centrifugation at 13000 g for 3 minutes, the aqueous phase was transferred to a fresh micro-centrifuge tube and 60 µl Elution Solution stored at -20ºC.

3.4 Polymerase chain reaction (PCR) assessment

3.4.1 Oligonucleotide primers

Published degenerate primer pairs hexAA1885/1913, and nehexAA1893/1905, which successfully amplified a conserved region of HAdV hexon gene fragment and which reportedly detect all known serotypes of HAdVs (Allard et al. 1990; Allard et al. 1992) were used in nested PCR. The primers used for the detection of HAdV DNA in a conventional PCR, and their sequences were presented in Table 3.1

3.4.2 PCR amplification of viral hexon genes

Polymerase chain reaction (PCR) is a method that used a thermostable DNA polymerase enzyme to amplify a DNA sequence of interest. It is necessary to know the base pair composition of the short regions flanking the sequence to be amplified in order to design the primers that initiate DNA synthesis. Nested PCR was used to amplify the hexon genes using hexAA1885/1913 and nehexAA1893/1905 published primer sets. The PCR reaction mixture contained 1 µl of each forward and reverse primer, master mix (1.5 mM of 25

mM 2'-deoxynucleotide 5'-triphosphate (dNTP) mix, 1.5 mM of MgCl2, 1X Taq Buffer with (NH4)2SO4, 0.5 U Taq DNA polymerase, (Fermentas Life Sciences) , and approximate 2ul of extracted DNA. RNase free water was added to bring the final volume to 25 µl. PCR products were amplified using hexAA1885/hexAA1913 primer sets. The following conditions: 94ºC for 5 minutes; followed by 30 cycles at 94ºC for 1 minute, 55ºC for 1 minutes, 72ºC for 1 minute; and finally 5 minutes at 72ºC. The nested PCR with nehexAA1893/nehexAA1905 primer sets amplified the templates. Thermocycling used the following condition: 94ºC for 3 minutes; 35 cycles (94ºC, 30 s; 55ºC, 30 s; 72 ºC, 1 min); 72ºC for 5 min. For distinguishing results from high throughput sequencing, this study used tru barcoding sequences (Table 3.2) at 5' end of the nehexAA1893/1905 primers to amplify the viral hexon genes. The thermal conditions were the same as described for nested PCR. The PCR amplification mixture was concentrated from PCR clean-up Kit for high-throughput sequencing using Illumina.

Table 3.1 A list of oligonucleotide primers for PCR

Name	Sequence	Reference
hexAA1885	GCCGCAGTGGTCTTACATGCACAGC	Allard et al. (1990)
hexAA1913	CAGCACGCCGCGGATGTCAAAGT	
nehexAA1893	GCCACCGAGACGTACTTCAGCCTG	Allard et al. (1992)
nehexAA1905	TTGTACGAGTACGCGGTATCCTCGCGGTC	

Table 3.2 TruSeq index sequences

Index	Sequences	Index	Sequences
Autumn - north	ATCACG	Spring - north	GATCAG
Autumn - central	CGATGT	Spring - central	TAGCTT
Autumn - south	TTAGGC	Spring - south	GGCTAC
Autumn - east	TGACCA	Spring - east	CTTGTA
Winter - north	ACAGTG	Summer - north	AGTCAA
Winter - central	GCCAAT	Summer - central	AGTTCC
Winter - south	CAGATC	Summer - south	ATGTCA
Winter – east	ACTTGA	Summer - east	CCGTCC
Binding primers	Sequences+ nehexAA1893 CGAGACGTACTTCAGYMTG nehexAA1905 ACGCGGTATCCTCGCGRTC		

3.4.3 Agarose gel electrophoresis

DNA molecules of different sizes can be separated on an agarose gel due to the negative charge of the phosphate groups in the backbone of the nucleic acids. Agarose is a gelatinous polysaccharide that forms a porous matrix after dissolving in hot water and cooling down. The PCR products were subjected to agarose gel electrophoretic analysis. 1, 1.5, or 2% of 1×TAE-agarose gel supplemented with 1 μl Nucleic Acid Staining (Yeastern, Taiwan) per 20 ml was prepared, depending on the expected size(s) of the amplicon(s).

5 to 10 of each post-PCR DNA sample was mixed with loading dye (Fermentas) and then loaded into the sample well of the agarose gel. To evaluate the size of each sample, the 100-3000 bps ladder (Fermentas) was loaded as the DNA size maker. Afterwards, the gel was electrophoresed in the horizontal electrophoresis device containing 1X TAE buffer at 50 V until the adequate DNA migration was achieved, which usually took 30 min. Finally, the DNA fragments in the gel was visualized in a long wave UV light box and then

documented as an image file by the Alpha Imager system (MS, major science).

3.5 TA cloning and plasmid preparation

3.5.1 yT&A vector ligation

Cloning is detrimental because of the uncertainty of science and technology, the loss of genetic diversity and the possibility of extinction of a particular species. The PCR products were subsequently cloned into the yT&A vector (Fig 3.5) using the protocol provided in the yT&A Cloning kit (Yeastern Biotech, Taipei, Taiwan). The ligation mixture consisted of 1X ligation buffer A, 1X ligation buffer B, yT&A cloning vector: PCR product (1 : 3), 0.2 unit of YEA T4 DNA ligase and topped up with distilled water to a final volume of 10 µL. The ligation mixture was mixed by pipetting gently and incubated overnight at 4ºC (Fig 3.6).

All of the ligation mixture was added to ice thawed Escherichia coli strain DH5α competent cells. The treatment of E.coli cells with Ca^{2+} and some other divalent or multivalent cations induces competence and makes these cells to take up DNA from the environment (Glover and Hames, 1995). The mixture was incubated on ice for 10 minutes, then heat shocked at 42ºC for 30 seconds and immediately placed on ice for 1 minute to promote the binding of DNA to the cells membranes and the uptake of the DNA into the cells. Meanwhile, 20 µl IPTG (0.1M isopropyl β d 1 thiogalactopyranoside) and 10µl Xgal (40 µg/mL 5-bromo-4 chloro-3-indolyl-beta-d-galacto-pyranoside) was prepared and spread onto LB agar supplemented with 50 µg/mL ampicillin for blue/white selection screening for recombinants. The cell pellets were resuspended in the remaining supernatant and plated on two LB agar plates each supplemented with 50 µg/mL ampicillin, IPTG and X-gal. The plates were incubated overnight at 37ºC.

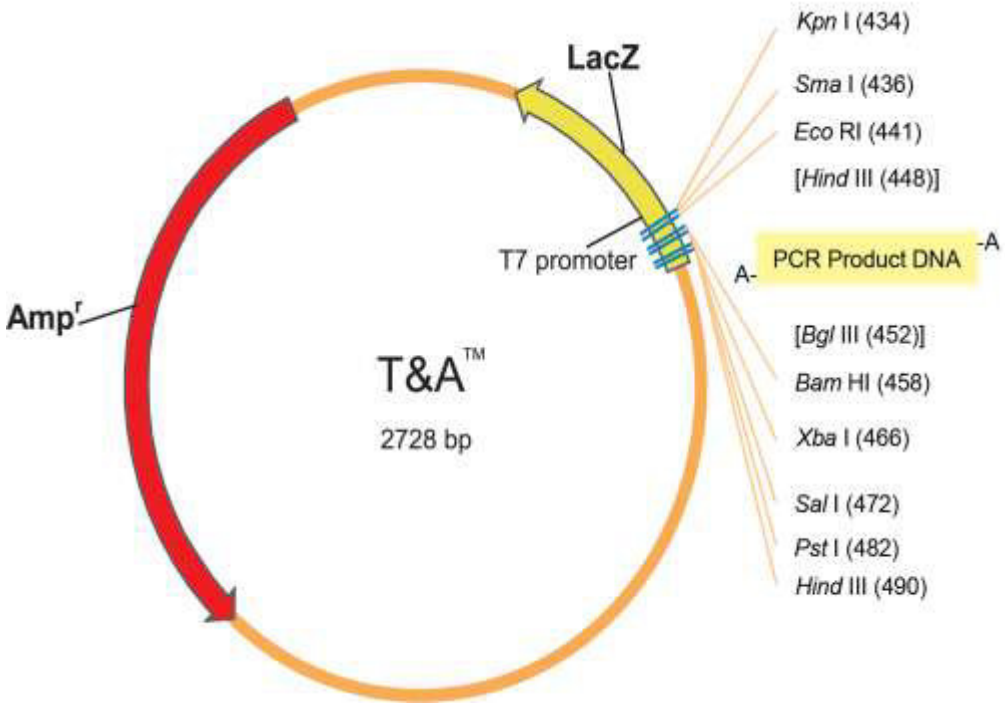


Figure 3.5 Map of yT&A^{TM} cloning vector used in this study

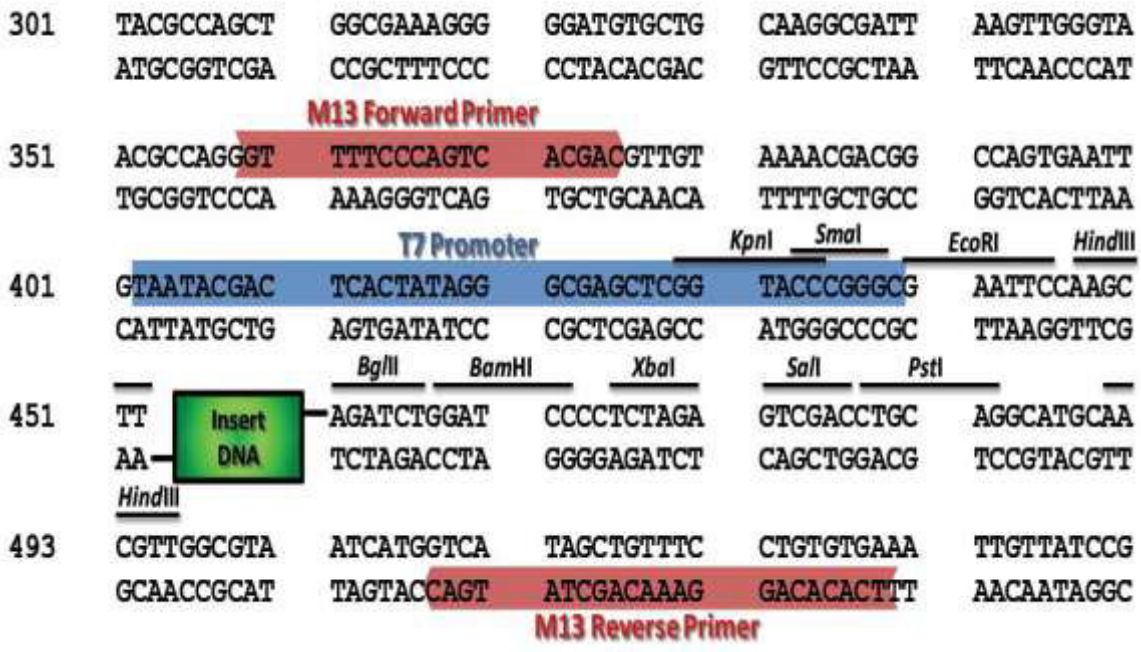


Figure 3.6 DNA cloning region in vector

3.5.2 Colony PCR assessment

Bacteria originating from a single colony can be used directly as template for PCR because the initial denaturating heat disrupts the cell walls and make DNA accessible (Fuchs and Podda, 2005). Even though both ends of the open yT&A vector are extended with 3′ terminal thymidines to avoid self re-ligation these ends are not dephoshorylated and can be religated by T4 ligase. Colony PCR was done to verify the presence and size of inserts of targeted PCR products after ligation reaction. Each colony chosen for colony PCR was transferred to a new ampicillin plate with a pipette tip and remains of the cells from the tip were mixed into the PCR reaction. The PCR reaction mix was set up as described in chapter 3.4.2, except that the total volume was adjusted to 25 µl without a DNA template. In each screen, the primers with one of the insert specific sequences for the yT&A vector containing hexon genes were used (nehexAA1893/1905). The program for the PCR was the same as used before (chapter 3.4.2). The amplified products were separated from the template DNA with 1.5% agar gel containing TAE buffer and visualized as described in agarose gel electrophoresis protocol.

3.5.3 Plasmid preparation

Plasmid DNA can be rapidly isolated from bacterial cultures using differential alkaline denaturation. Single colony harboring desirous plasmid was inoculated in 5ml LB broth (Lysogeny-Broth, also called Luria-Bertani) containing 1% peptone, 0.5% yeast extract, 0.5% NaCl, 150 µg/mL of ampicillin and cultured with vigorous shaking for 16 hours at 37ºC. Plasmid was purified from plasmid Miniprep purification Kit (GeneMark, Taiwan). The culture was centrifuged at 13000 g for 1 minute at room temperature and the supernatant removed. The pellet was resuspended in 200 µl Plasmid solution I (50 mM glucose, 25 mM Tris-HCl (pH 8), 10 mM EDTA) by vortex. After 200 µl Plasmid

solution II (0.2M NaOH, 1% SDS) was added, the mixture was mixed by inverting the tube 5 times. After Plasmid solution III (3M sodium acetate, pH 4.8) was added, the mixture was mixed by inverting the tubes 5 times. After centrifugation of 5 minutes at 13000 g in spin column, the supernatant was transferred to a fresh tube and an equal volume of Wash Solution was added and mixed well. After centrifugation at 10000 g for 15 minutes, the aqueous phase was transferred to a fresh micro-centrifuge tube and 60 µl Elution Solution stored at -20ºC.

3.6 Restriction enzyme reaction

The PCR amplicons were treated with 5U HpyCH4III (would only cut partial hexon sequence of HAdV 41) at 37ºC for 4 hours. Enzyme was selected using the NEB cutter computer program (Fig 3.7). The colony PCR products and restriction fragments were separated and visualized on a 1.5% agarose gel (100V, 30 min) to screen colony with different digesting pattern in order to find addition HAdV serotype.

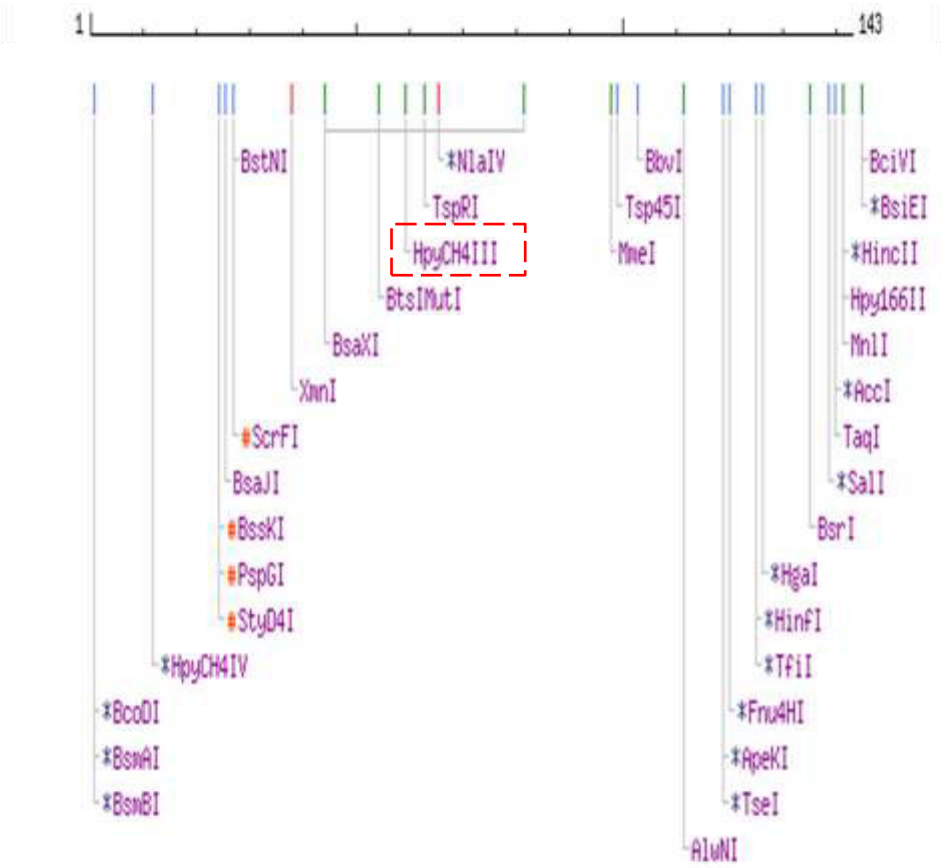


Figure 3.7 Position of HpyCH4III enzyme digested in colony PCR amplicons.

3.7 DNA sequencing

The colony was incubated in 5 ml LB broth (Lysogeny-Broth, also called Luria-Bertani) containing 1% peptone, 0.5% yeast extract, 0.5% NaCl, 150 µg/mL of ampicillin) after cloning and cultured with vigorous shaking for 16 hours at 37℃. The cultures were incubated for sequencing. The exact order of base pairs in a certain DNA sequence can be determined by dye-terminator sequencing. The double stranded DNA fragment to be analyzed is mixed with a single primer binding to a strand complementary to the strand to be sequenced. DNA polymerase and nucleotide mixture containing ordinary dNTPs and ddNTPs labeled with a fluorescent dye but missing 3′-OH group at the polymerization site. In a reaction based on the same principle as PCR, insertion of ddNTP terminates the

extension of that particulate strand. The insertion of these nucleotides is random but after a series of cycles the sequencing reaction contains mixture of copies of one strand ended at each base. There are four different dye-labels specific for each nucleotide (Reece 2004). Sequence data from both strands were edited in Alignment Editor (ClustalX). The sequences of representative isolates were compared with HAdV sequences present in GenBank by using BLAST program to search for the most similar strains available.

3.8 High-throughput sequencing

After PCR cycling, the 25 µl aliquots were pooled to obtain a total of 320 µl of PCR product. This solution was cleaned up using the PCR Clean-Up Kit (GeneMark). For a 320 µl reaction volume, 1000 µl of Buffer PB was added, vortexed briefly to mix, and applied to a column with collection tube. The tubes were centrifuged for 1min at 14000 rcf, and the flow-through was discarded. Then, 750 µl of PE was added to each column, the tubes were centrifuged for 1 min at 14000 rcf, and the flow-though was discarded. Next, each column was dried by centrifuging again for 30 s at 14000 rcf. The columns were then transferred to a fresh 1.5 ml microcentrifuge tube, and left open in a laminar flow hood to dry for 2 min. Following the air-drying step, 10 µl of elution buffer was added directly onto the center of the column and incubated for 1 min at room temperature. The follow-through from this step contains the DNA from the PCR reaction and was saved and stored at -20 ºC. The mixture was performed for illumine sequencing (Fig 3.8). The illumine platform was capable of performing paired-end (mate-paired) reads in which sequencing is conducted from both ends of the target DNA. To date, the illumine methods is capable of sequencing ~140 bps from each end of an insert, although the quality of the reads past 100 bps drops significantly.

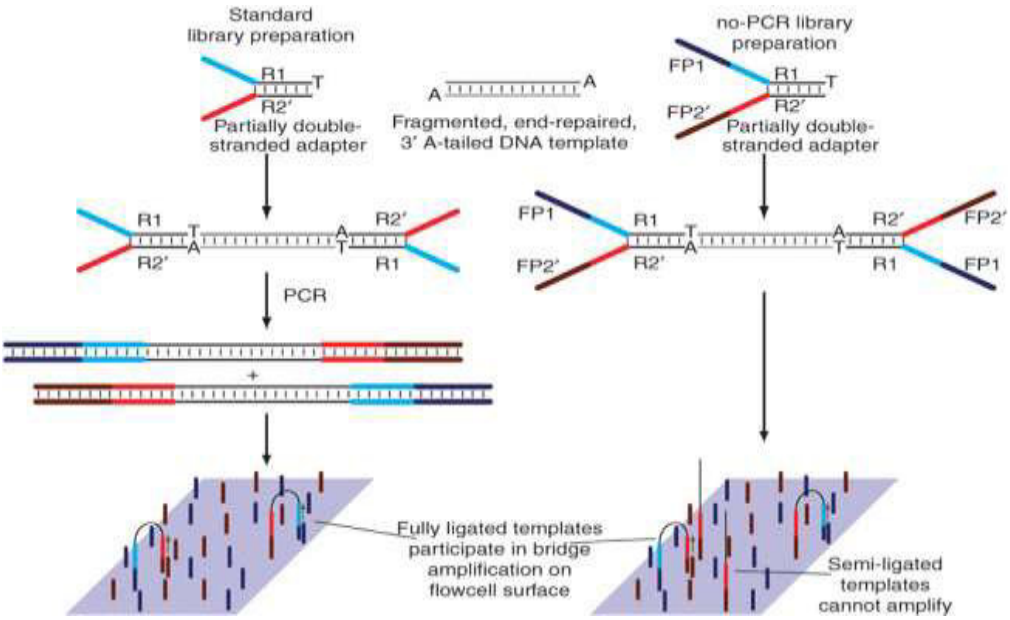


Fig 3.8 Mechanism for high-throughput sequencing

During illumina sequencing, both standard and no-PCR library preparations (partially complementary adapters with a 3′ thymine, T) overhang are ligated onto fragmented, end-repaired, 3′ adenine (A)-tailed DNA (Fig 3.8). Whereas standard adapters consist only of sections to which read 1 and read 2 sequencing primers hybridize (R1 and R2'), no-PCR adapters also contain sequences that facilitate hybridization to oligonucleotides attached to the flowcell surface (FP1 and FP2'). The standard library preparation uses PCR to add these sections and to enrich for fully ligated templates, which then amplify on the flowcell surface. In the no-PCR approach, the flowcell itself is used to select for fully ligated template molecules. All no-PCR templates hybridize to the flowcell in the same orientation because only the FP2' sequence is reverse-complementary to a flowcell oligonucleotide.

3.9 Statistical analysis- UniFrac

The Unifrac web application allowed researchers to compare many environments simultaneously using hierarchical clustering and principal coordinates analysis (PCA). Weighted Unifrac can be used to compare many communities simultaneously using standard multivariate methods. We use a hierarchical clustering method called unweighted pair group method with arithmetic averages (UPGMA) to cluster the community samples. This study analyzed 16 HAdVs-DNA sequences generated in 16 different of marine environments. Unifrac requires two input files: a single rooted phylogenetic tree that contains sequences derived from at least two different environmental samples, and a file describing which sequences came from which sample. The phylogenetic tree can be in Nexus format, and can be generated by ClustalX. The environment information must be provided in a text file in which each line contains a sequence name and an environment name separated by a tab (Fig 3.9).

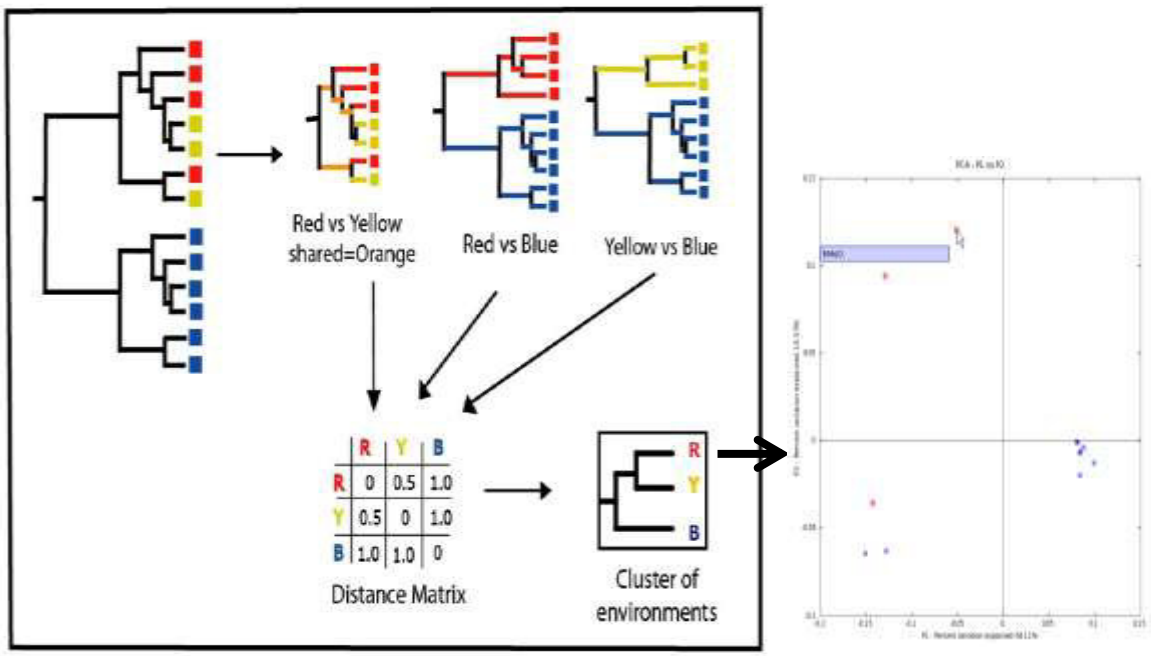


Figure 3.9 Unifrac distance metric.

Scatter plots will plot the following principal components against each other for each environment: P1 vs P2, P1 vs P3, and P2 vs P3. The names of the two principal components that contribute to the graph are written in the title. Environment with similar names will be assigned the same color, providing a convenient way to group environments visually by changing the names in the file you upload.

3.10 Real-time PCR quantification

3.10.1 Preparation of real-time PCR standard

The real-time PCR standard solution was prepared according to the following procedure.

● Calculate the molar weight (M = mole/L) of the purified plasmid..

● Determine copy numbers per microliter of the purified plasmid based on molar concentration. 1M is equivalent to about 6.0221415×10^{23} copies.

● Make the first dilution, typically with a final concentration of 1×10^{10} copies per 5 µl*(2× copies/µl). Use the C1V1=C2V2 relationship, where V1 is the unknown volume, C1 is the stock linearized plasmid concentration (copies/µl), V2 is the final volume (typically 100 µl) and C2 is the final concentration.

● Carry out a serial dilution of 1×10^{10} copies/µl of the linearized standard. Use nuclease-free water with plasmid DNA as diluent. A volume of 90 µl of diluent is used per dilution. See the diagram below as reference. Discard the last 10 µl. Use the 1×10^{7} copies/µl through the 10 copies/µl dilutions as standards, stored at -20ºC. (Fig 3.10)

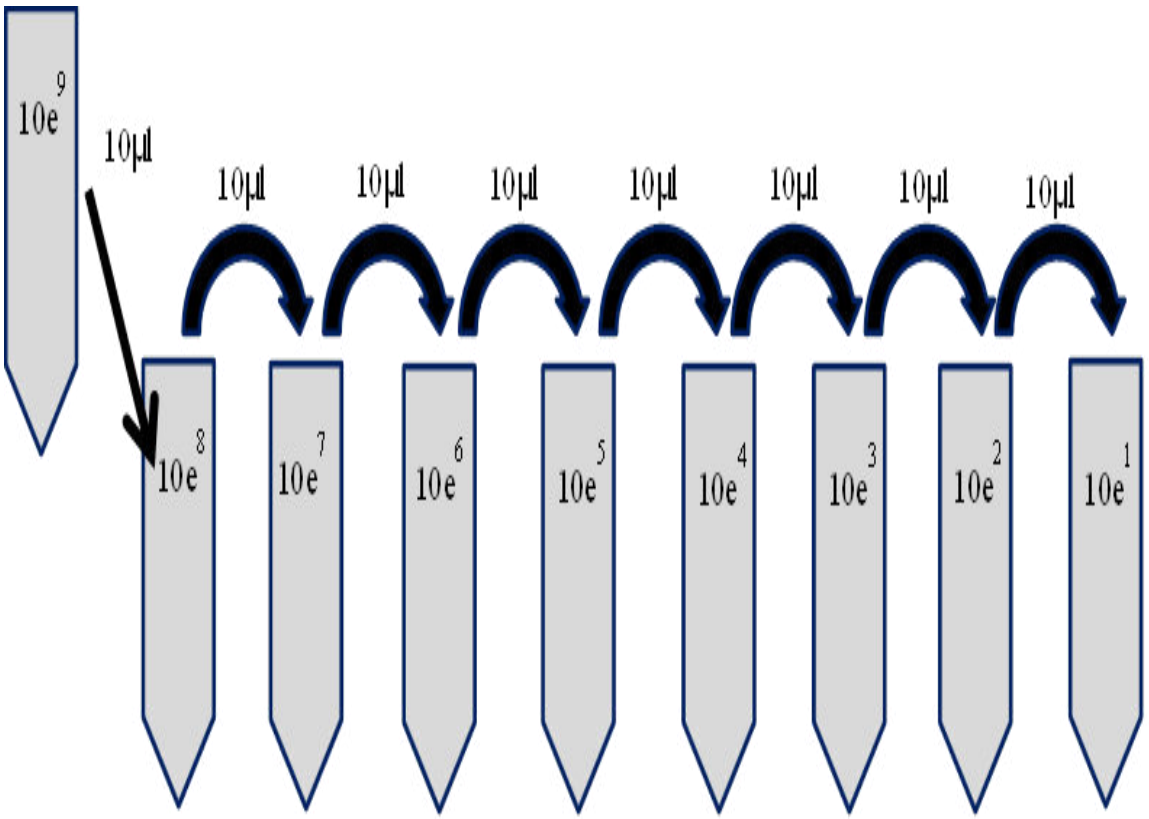


Figure 3.10 Standards preparation for real-time PCR

3.10.2 Real-time PCR assessment with probe and primer

In this study we used TaqMan® chemistry Fast Real-Time PCR System (Applied Biosystems Foster city, California). TaqMan® chemistry is based on measuring increase of fluorescence. The target probe has a fluorescent dye (reporter) bound in the 5′- position and a quenching dye (quencher) in the 3′ position. If the probe anneals to a specific target sequence, it is cleaved by the 5′ nuclease activity in the Taq® DNA polymerase as the primer is extended (Fig 3.11). And during accumulation of the PCR product the fluorescence intensity from the probe increases. If the probe does not anneal the reporter dye is quenched by the quenching dye. The increase in fluorescence is proportional to the amount of amplicons. This specific sequence is then used to design primers and probe. It is recommended to design primers and probe in a region with low Guanine/Cytosine

content (20-80 %). The primers should be close to the probe and amplify short sequences within the target sequence, 50 to 150 bps in amplicon size. The melting temperature (Tm) for the primers should be 58-60°C and 68-70°C Tm for the probe.

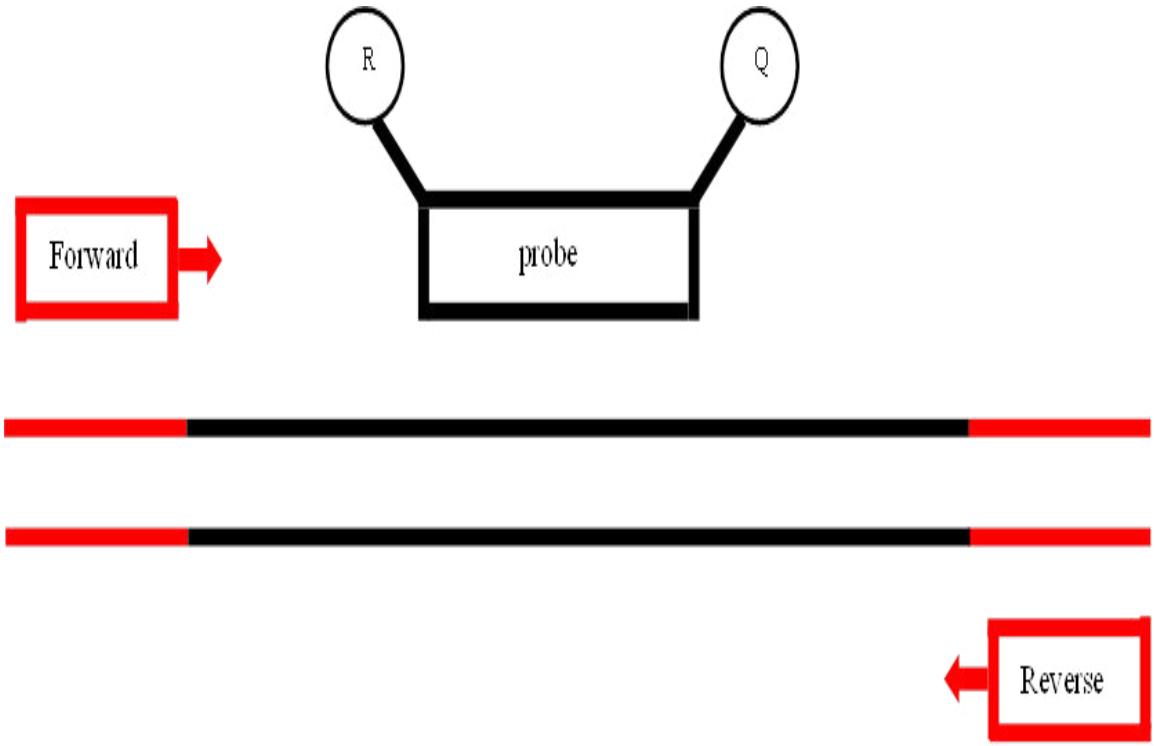


Figure 3.11 TaqMan® Assay of probe and primer

However, the overall species of HAdV primers and probe were designed as a TaqMan® assay according to paper's guidelines. The primers and probe sequences and the amplicon were compared to species F41 sequences by using Ad41 primers and probe from a previous study. Table 3.3 shows the sequences of the primers and probe used for targeting the all species and F41 in this study.

The reaction solution contained TaqMan® Fast Universal PCR Master Mix (2×), in addition RNase Free water, primers, and probes. The final concentration of 900 nM of each primer and 600 nM of HAdVs probe for F41 detected was found to be optimal, each of primer and probe for overall species in a final concentration of 150 nM. For detection

of all species and the species F41, the reaction volume was 20 µl, with 18 µl reaction solution and 2 µl templates. All templates were run in parallel, and in addition each PCR analysis was run with a known negative template control containing water to test the reaction solution.

Table 3.3 Primers and Probes for quantitative real-time PCR

Name	Sequence	Reference
JTVXF	GGACGCCTCGGAGTACCTGAG	Jothikumar et al. 2005
JTVXR	ACIGTGGGGTTTCTGAACTTGTT	
JTVXP	FAM-CTGGTGCAGTTCGCCCGTGCCA-BHQ	
HAdV-F4041-hex157f	ACCCACGATGTAACCACAGAC	Jiang et al., 2005
HAdV-F41-hex346r	CACTTTGTAAGAATAAGCGGTGTC	Xagoraraki et al., 2007
HAdV-F4041-hex214p	FAM-CGACKGGCACGAAKCGCAGCGT-BHQ	Xagoraraki et al., 2007

The real-time PCR thermal step for total HAdV species was as follows:

Step 1: 15 minutes at 95°C to activate the Taq DNA polymerase.

Step 2: 10 seconds in 45 cycles of 95°C to denaturate DNA.

Step 3: 30 seconds at 55°C to anneal and extend primers.

Step 4: 15 seconds at 72°C to elongate and finally at a temperature transition rate of 20°C.

The real-time thermal step for HAdV 41 serotype was as follows:

Step 1: 15 minutes at 95°C to activate the Taq DNA polymerase.

Step 2: 10 seconds in 50 cycles of 95°C to denaturate DNA.

Step 3: 40 seconds at 58°C to anneal and extend primers.

Step 4: 20 seconds at 70°C and finally 30 seconds at 40°C.

Chapter 4 Results and discussion

In this study, partial hexon genes (approx. 143 bps) of human adenoviruses (HAdVs) were chose as target regions that were PCR amplified and further sequenced to identify HAdV serotypes in sewages from northern, central, southern, and eastern areas of Taiwan. Sequencings were carried out based on two approaches i.e., Sanger 4 color terminators using ABI 3730 Sequencer (for cloned hexon genes) and paired-end 4 color-reversible terminators using Illunima Genome Analyzer (for mixture of PCR amplicons). Additionally, HAdVs concentrations in the sewage samples were quantified using real-time PCR assays. Results of HAdVs identification (serotyping) and quantification were presented in the following sections. Moreover, geological distribution of HAdVs and seasonal variations were further discussed.

4.1 Collection of viral samples from municipal sewages

Efficient virus collecting methods were crucial for detecting low number of viral particles in water samples. In this study, viral samples were concentrated from large amount of raw sewages using filtration, elution, and flocculation procedures suggested by USEPA. Large volume of raw sewages (approximately 18 ~ 200 liters depending on the MLSS concentration, Table 4.1) were pumped through 1-MDS filters. Typically, higher MLSS in the raw sewage sample would lead to smaller sampling volume (e.g., raw sewages of north, central, and south areas) because suspended solids would clog the filter within a shorter period of time. Meanwhile, backup pressures measured between pump and filter would build up faster for the sewage sample with high MLSS. Each of sampling events was stopped when the backup pressure reached 30 psi in order to collect high quantity of viral particles.

Table 4.1 Records of sampling dates and water qualities for 16 sampled sewages.

Seasons	Areas	Date	Temp. (ºC)	pH	Volume (L)	SS (mg/L)	BOD/COD ratio
Autumn	N	2011/10/02	25	6.9	20	123	0.4
	C	/10/12	25	6.7	20	151	0.5
	S	/10/16	27	6.9	20	173	0.3
	E	/10/22	24	6.8	60	100	0.3
Winter	N	2012/02/17	25	7.0	20	124	0.4
	C	/01/07	25	6.7	20	163	0.5
	S	/02/16	21	6.9	20	168	0.3
	E	/02/17	24	6.8	60	64	0.4
Spring	N	2012/04/26	25	6.9	18	428	0.4
	C	/04/03	25	6.7	20	166	0.4
	S	/04/11	26	6.9	20	127	0.3
	E	/04/26	24	6.8	200	52	0.2
Summer	N	2012/07/30	26	6.5	40	93	0.4
	C	/07/05	24	7.0	35	120	0.4
	S	/07/17	29	6.2	35	163	0.3
	E	/08/20	26	7.2	120	54	0.3

(Water quality data for the influent raw sewage was recorded from the monitoring results done by the four WWTPs located at northern, central, southern, and eastern Taiwan)

In addition, pH and temperature of sampled sewages were ranged from 6.2 to 7.2 and from 21 to 29℃, respectively. The values for southern Taiwan were varied more significantly. Besides, the BOD/COD ratios were less than 0.5 suggesting wastewater sources other than municipal sewage (e.g., industrial and agricultural wastewaters) could somehow be collected to the WWTP.

All 16 filtrated samples were eluted and concentrated in laboratory within 72 hours. The final concentrated viral samples (~30 mL) were aliquoted into 5-mL tubes and stored in -80℃ freezer for further viral DNA/RNA extraction. The concentrated samples should contain most of viruses and phages with size less than 0.22 μm.

4.2 Viral DNA extraction

Viral DNAs were extracted from the concentrated sewage samples using two commercial kits. Since the RNase was not applied during extraction, the extracts should also contain viral RNAs. Purified viral DNA/RNA concentrations and 260/280 ratios were about 13~25 ng/μL and 1.6~3.3 (Table 4.2) suggesting quantity and quality of the extracted DNA/RNA were rather passable. The extracted viral DNA/RNA concentrations of 16 concentrated sewage samples ranged from 13 to 25 ng/μL implying certain amount of viruses including HAdVs might present in the collected samples. Additionally, four extracts had 260/280 ratios higher than 3. Although high 260/280 purity ratios were not indicative of an issue, it was surmised some detection variations might occur at both wavelengths of 260 and 280 nm. All 16 viral DNA extracts were then PCR amplified for further cloning, sequencing, and enumerating.

Table 4.2 Concentrations (ng/μL) and 260/280 ratios of the extracted viral DNA

Areas	Seasons			
	Autumn	Winter	Spring	Summer
North	Conc.: 17	Conc.: 16	Conc.: 13	Conc.: 18
	260/280 2.1	260/280: 2.0	260/280: 3.3	260/280: 1.6
Central	Conc.: 21	Conc.: 19	Conc.: 16	Conc.: 19
	260/280: 1.7	260/280: 1.6	260/280: 3.1	260/280: 2.1
South	Conc.: 14	Conc.: 21	Conc.: 24	Conc.: 25
	260/280: 2.1	260/280: 3.2	260/280: 2.1	260/280: 2.1
East	Conc.: 14	Conc.: 15	Conc.: 15	Conc.: 13
	260/280: 2.1	260/280: 1.8	260/280: 3.1	260/280: 1.7

4.3 PCR amplification of target hexon genes

Results from nested PCR amplification showed that the target HAdV hexon genes were successfully amplified using the two published primer sets of hexAA1885/1913 (approx.. 301 bps for the first PCR run) and nehexAA1893/1905 (approx.. 143 bps for the second PCR run) as images of gel electrophoresis showed the final PCR amplicons (~ 143 bps) had a same size nearby the 100-bps marker (Fig 4.1). Nevertheless, during the first PCR trial carried out by using a commercial master mix (i.e., Dream-Taq, Thermo, USA), target amplicons could not be seen for three (out of the 16) samples (i.e., spring-south, summer-central, and summer-east). This problem was overcome after replacing DNA taq with another commercial master mix (i.e., TaKaRa taq, TaKaRa, JP) while dilution of viral DNA extracts did not improve PCR amplifications for the target hexon genes. The 16 PCR products were further cloned for sequencing.

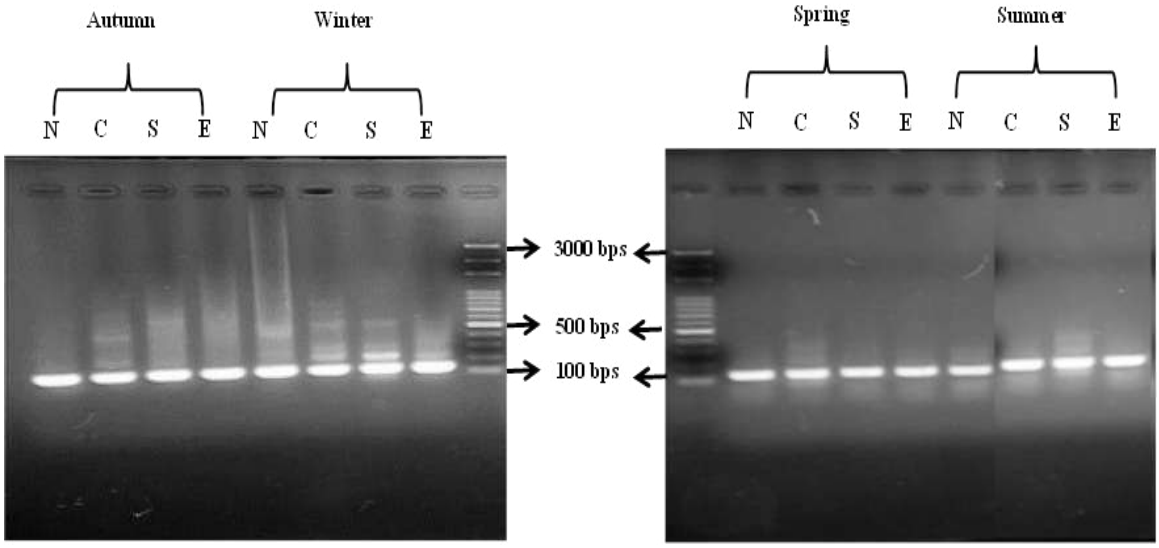


Fig 4.1 PCR-amplified partial hexon genes (143 bps) for the 16 viral DNA extracts were observed in agarose gel after electrophoresis. (N) North, (C) Central, (S) South, (E) East.

Besides, mixture of all 16 PCR products (together with other 16 activated sludge samples for another study) was directly sequenced based on high-throughput paired-end sequencing method using 4 color reversible terminators. In order to distinguish sequenced results, the PCR runs were performed using 16 special primer pairs (each for one of the 16 samples) that were labeled with different 6-mer barcode tag. The results showed target hexon genes were successfully amplified from all 16 viral DNA extracts with an amplicon size around 155 bps (Fig. 4.2). The 16 PCR products were mixed together and further cleaned-up and concentrated. The prepared mixture had a 260/280 ratio of 2.0 and a concentration of 82.7 ng/μl. The mixture seemed very clean and adequate when observed in gel electrophoresis (Fig. 4.3). It was directly sequenced using the Illumina Genome Analyzer without cloning procedure. A total of approximately 3.3×10^5 reads were obtained (i.e., about 2×10^4 reads for each of the 16 samples) to provide large amount of sequences for serotyping HAdVs in sewages.

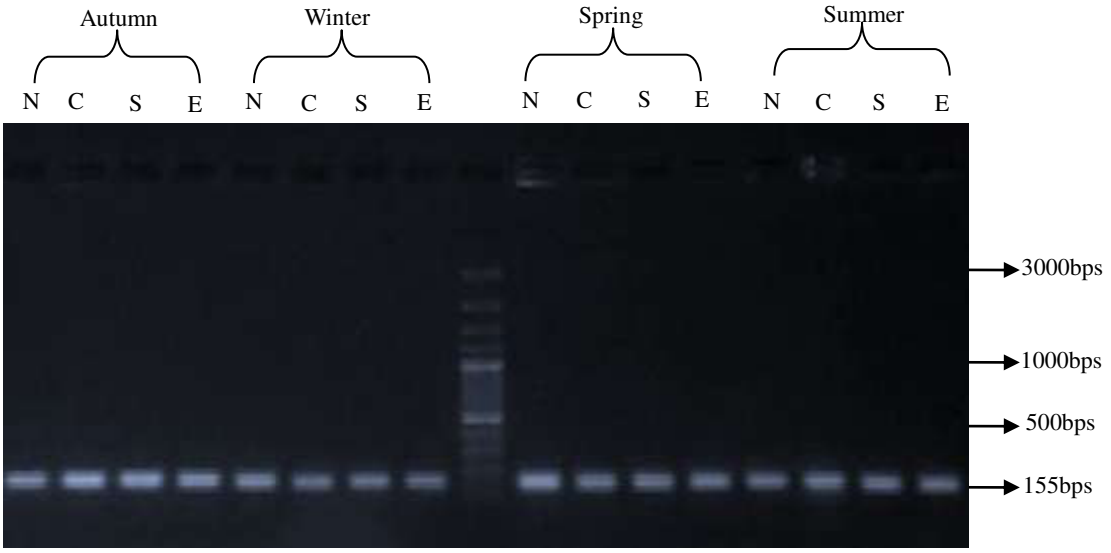


Fig 4.2 The gel image of PCR amplification with tru primer pairs for high-throughput sequencing. (N) North, (C) Central, (S) South, (E) East.

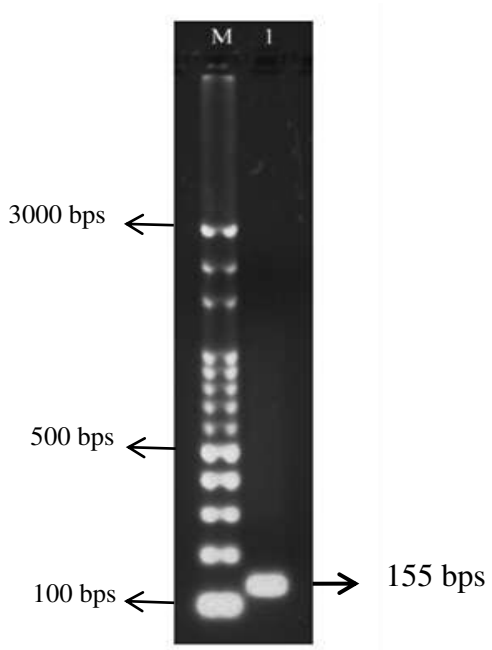


Fig 4.3 Image of gel electrophoresis showed the prepared PCR mixture was clean and adequate (M: DNA size marker; 1: the concentrated PCR mixture).

4.4 Cloning of PCR-amplified partial hexon genes

The PCR-amplified partial hexon genes were subsequently ligated into yT&A vector. Prior to transformation of ligated vectors into host cells, the ligates were checked with PCR tests using a primer pair of M13F/R to confirm insertions of partial hexon genes into vectors. When partial hexon genes were inserted into vectors, amplicons with a size of approx. 315 bps would be seen in images of gel electrophoresis (Fig 4.4); otherwise, approx. 172-bps amplicons would show up. The ligated vector containing target hexon genes were then transformed into competent cells (i.e., Escherichia coli DH5α). The transformed host cells were grew overnight on LB agar medium containing ampicillin (50 mg/L); and, overnight-grew white colonies (Fig 4.5) were numbered and picked-up to another LBamp plate for further screening and sequencing.

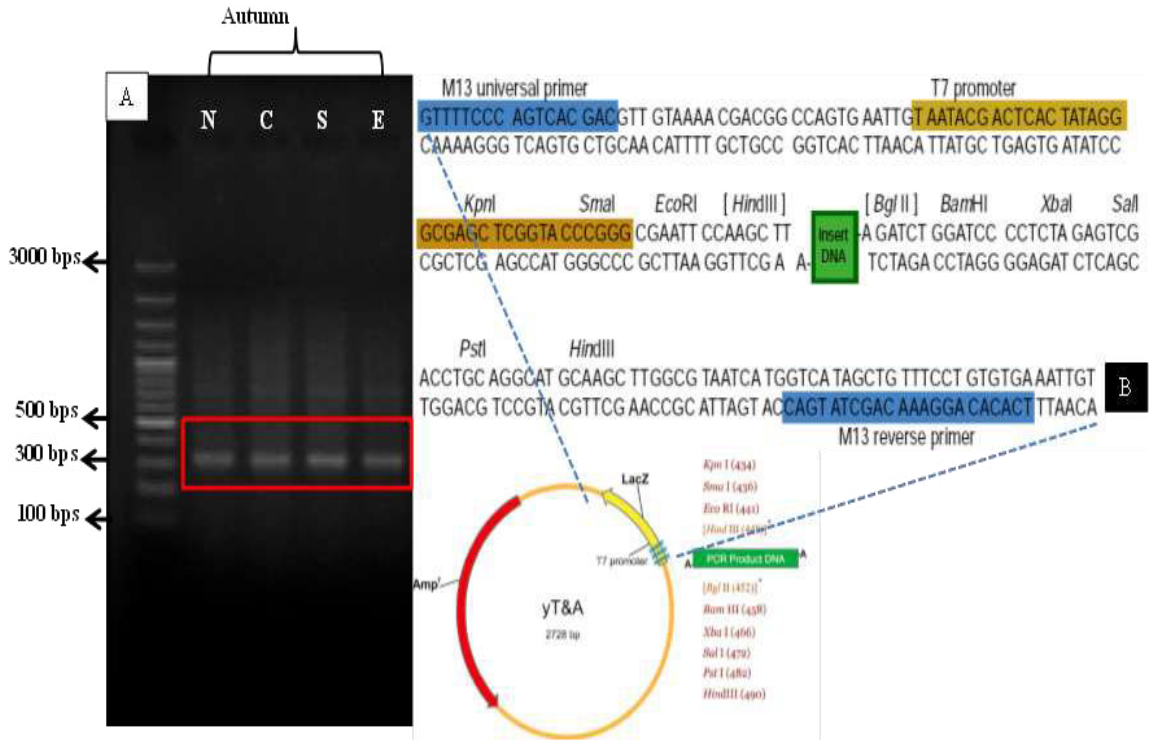


Fig 4.4 An example of gel image showed PCR amplicons from ligated vector were about 315 bps (i.e., 143 bps of partial hexon gene + 172 bps of vector fragment) in size (A) confirmed the partial hexon genes (~143 bps) were ligated into the vector (B).

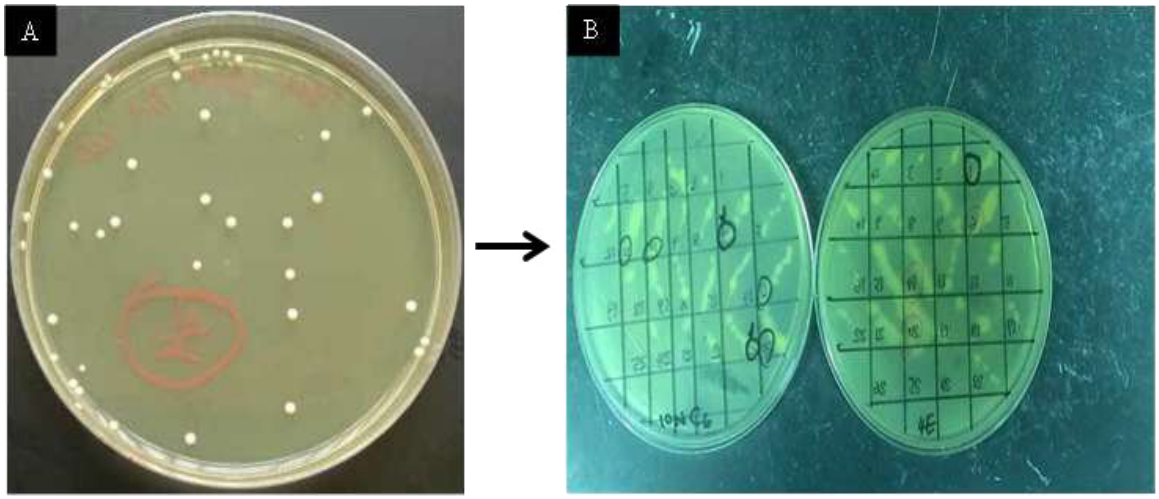


Fig 4.5 Examples of (A) overnight-grew colonies and (B) numbered/selected colonies on LBamp plates.

A total of 50-60 colonies were picked-up for each of the 16 samples (Table 4.3). Approximately half of them were screened by colony PCR using nehexAA1893/1905 primer set to check whether target hexon genes were presented in host E. coli cells or not. Results of colony-PCR-screening demonstrated that more than 40 % of screened colonies were positive (i.e., contained target hexon genes) (Table 4.3, Fig 4.6). For each of the 16 samples, about 8-35 positive colonies were sent for sequencing; and, for most of the samples, about 20 sequences (17-22) were obtained for phylogenetic analyses (Table 4.3).

Table 4.3 Numbers of colonies that had been picked-up (P), PCR-screened (S), screened to be positive colonies (PC), sequenced (SC), and used for phylogenetic analyses (UC).

Areas	Colonies	Seasons				Total
		Autumn	Winter	Spring	Summer	
North	P	60	60	60	60	
	S	30	30	30	40	
	PC	20	30	30	40	
	SC	20	30	30	20	
	UC	**20**	**20**	**20**	**20**	**80**
Center	P	60	60	60	30	
	S	40	30	35	20	
	PC	20	25	30	20	
	SC	20	25	30	20	
	UC	**17**	**20**	**19**	**7**	**63**
South	P	60	60	50	60	
	S	40	40	20	40	
	PC	32	35	8	40	
	SC	32	35	8	20	
	UC	**22**	**13**	**3**	**20**	**58**
East	P	60	60	50	30	
	S	25	40	30	20	
	PC	20	30	20	20	
	SC	20	30	20	20	
	UC	**15**	**22**	**10**	**7**	**54**
Total						**255**

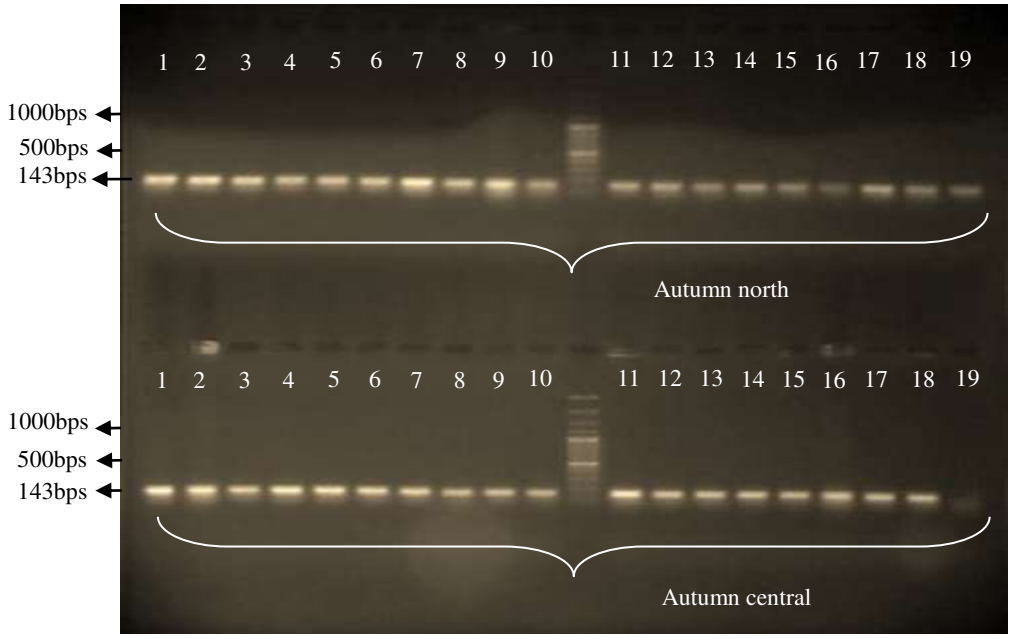


Fig 4.6 An example of gel image for colony PCR showed most of colonies contained partial hexon genes (~143 bps).

4.5 Restriction enzyme digestion

Restriction enzyme HpyCH4III was used to digest partial hexon genes that were amplified from another colony PCR using M13F/R primer pair to gain various fragment patents for predicting sequences variety before sequencing. Based on hexon gene sequences of seven HAdV species, the HpyCH4III could digest partial hexon genes at different positions resulting in various fragment profiles for each distinct species. Results showed that, most of time, the HpyCH4III did not cut the partial hexon genes as expected (Fig 4.7). Few trials did show different digestion patents; however, the sequencing results demonstrated the sequences were identical and belonged to same HAdV species. Therefore, positive colonies were randomly selected for follow-up sequencing by ABI 3730 Sequencer using a method of Sanger 4 color terminators.

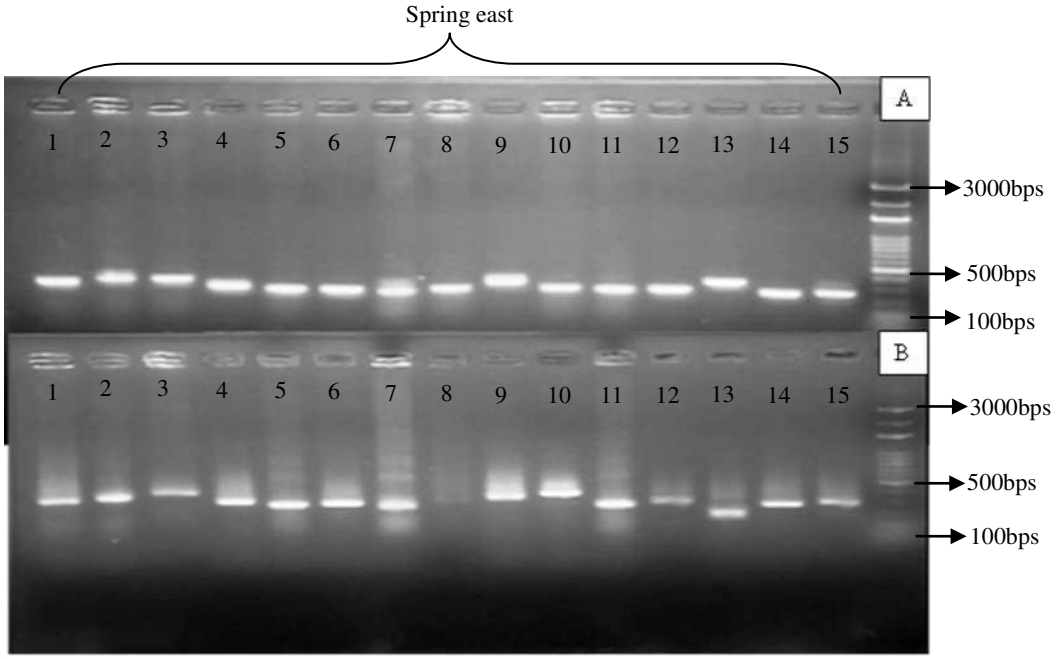


Fig 4.7 Examples of gel images of (A) colony PCR product using M13F/R primer set and (B) restriction enzyme HpyCH4III digested product.

4.6 Nucleotide sequence and phylogenetic analysis of HAdVs

In this study, the PCR-amplified partial hexon genes were revealed by using two sequencing approaches including the Sanger method using 4 color terminators and ABI 3730 Sequencer for cloned hexon genes (conducted at Mission Biotech, Taipei, Taiwan) and a pair-end method using color-reversible terminators and Illunima Genome Analyzer for mixture of hexon genes (conducted at Yourgene Bioscience, Taipei, Taiwan).

Each of sequencing results for the cloned partial hexon genes were compared individually with reference sequences in the NCBI database using BLAST method. Sequence comparison results indicated the cloned partial hexon genes were matched sequences of HAdV serotypes F41, C2, C6, and D8 (similarities ≥ 98%). The HAdV F41 was identified as a sole serotype group in 12 samples (out of 16 of total viral samples collected from raw

sewages) (Table 4.4). Besides, both serotypes F41 and C2 were found for a southern sample collected during winter; and, both serotypes F41 and D8 were discovered in an eastern sample taken during spring season. Additionally, HAdV C6 and C2 were recognized as an only serotype group for the autumn-north and spring-south samples, respectively.

Additionally, all the revealed partial hexon sequences and reference sequences from NCBI genbank were aligned to construct a phylogenetic tree (Fig. 4.8). All the cloned and sequenced hexon genes were highly affiliated (≥98%) to the hexon sequences of C, D, and F species and clustered into four groups within C2, C6, D8, and F41 serotypes. The majority of cloned hexon sequences (~87.8%, i.e., 224 hexon sequences were related to F41 out of 255 total obtained sequences) were related to serotype F41 (Genbank accession number: X51783). Many cloned hexon sequences that belonged to HAdV F41 were the same. In contrast, only few sequences i.e., 9 (~3.5%), 20 (7.8%), and 2 (~0.8%) belonged to C2, C6, and D8, respectively.

These results suggested that F41 was the most dominant HAdV serotype presented in the studied raw sewages while other serotypes (i.e., C2, C6, and D8) only could be detected occasionally. Besides, it was out of anticipation that only a sole serotype was found in 14 samples. These may be owing to different fates between various serotypes about their releasing into sewage, survival rate within sewage, withdraw effectiveness during sampling and concentrating, enlargement efficiency during PCR amplification of hexon genes, and/or chance for PCR-amplified hexon genes to be cloned and sequenced. In order to minimize possibility of not been cloned and sequenced for target hexon genes, a high-throughput sequencing approach was used to reveal large amount of PCR-amplified hexon genes (approx. 1.56×10^5 reads per sample).

Table 4.4 Matched HAdV serotypes and species for 16 viral samples collected from sewages by two sequencing approaches.

Areas	Seasons							
	Autumn		Winter		Spring		Summer	
	C*	H†	C	H	C	H	C	H
North	**C6:**20	**C6**:17	---	**C2**:12	---	---	---	---
	---	**F41**: 35,467	**F41**:20	**F41**:33,259	**F41**:20	**F41**:48,790	**F41**:20	**F41**:45,754
Center	---	---	---	**A**: 7	---	---	---	---
	F41:17	**F41**: 2,555	**F41**:20	**F41**:40,113	**F41**:19	**F41**:24,449	**F41**:7	**F41**:11
South	---	---	**C2**:6	**C2**:7	**C2**:3	**C1,2,6**:38	---	---
	F41:22	---	**F41**:7	**F41**:2,303	F:rt‡	F:rt	**F41**:20	**F41**:2,193
East	---	**C2**:450	---	**C**:26	---	**C2**: 154	---	---
	---	---	---	**D**:12	**D8**:2	**D**:8	---	---
	F41:15	**F41**:2,418	**F41**:22	**F41**:13,352	**F41**:8	**F41**:72,185	**F41**:7	**F41**:2,308

* C: cloning-sequencing; †H: high throughput sequencing; ‡ rt: real-time PCR found; number listed after the matched HAdV serotype represented number of sequence.

A total of 325,888 sequences obtained from high throughput sequencing were matched HAdVs partial hexon gene sequences from species A, C, D, and F. These results quite matched those from cloning-sequencing approach (Table 4.4). Nine out of the 16 samples showed completely same HAdV identification results while other five samples were somewhat similar (i.e., north-winter, center-winter, east-autumn, east-winter, and east-spring). Only two samples (north-autumn and south-autumn) showed significantly different results (i.e., F species HAdV was only identified in one sequencing method). Additionally, F41 was still the most dominant HAdV serotype based on high-throughput sequencing results while many revealed sequences matched F41 after BLAST comparison. The overall results demonstrated two sequencing approaches did gained similar identification outcomes about HAdV occurrence in the 16 sewage samples.

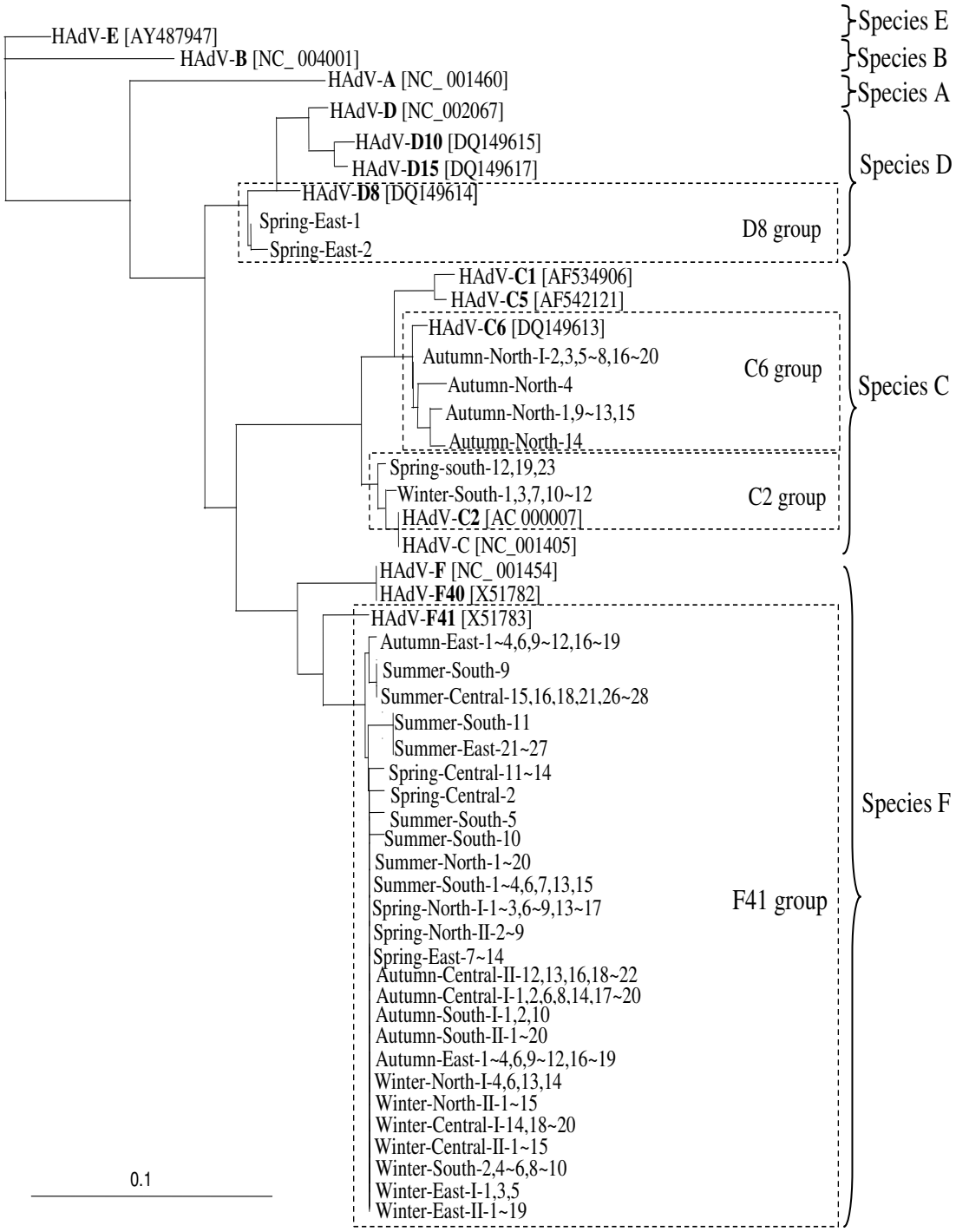


Fig 4.8 Phylogenetic tree based on sequence comparisons of a 143-bps hexon fragment been cloned and sequenced showing HAdV from raw sewage samples were clustered into three groups belonging to species C, D, and F.

Additionally, among large number of sequences were identical to particular HAdV serotypes, sequence variations were still noticed. The top five most abundant sequences from high throughput sequencing were aligned with reference sequences from NCBI Genebank to construct a phylogenetic tree (Fig 4.9). All the high throughput sequenced hexon genes were highly affiliated ($\geq$ 98%) to the hexon sequences of C and F species and clustered into four groups within C1, C2, C6 and F41 serotypes. In addition, species C was discovered for a southern sample during spring season by two sequencing approaches. But, the quantification results (Section 4.9) also showed that serotype F41 was existed. These may be owing to specific probe successfully amplified hexon genes by real-time PCR quantification. Species C and D were relatively few found in 5 samples. Owing to sequence quantities were insufficient, it could not identify by cloning sequencing easily.

Some distinct difference in this study incidence related serotypes 3 and 7 from species B was not apparent by two sequencing approaches. In the case of species B, serotype 3 was very uncommon in 1985 and 1986 and underrepresented in 1983, 1989 and 1995 although there were no years showing a significant peak of isolates. On the other hand, the serotype 7 was particularly prevalent in 1984 and significant numbers were isolated in 1990 and 1992 but relatively few were found in the other years (Cooper et al., 1993). In clinical categories, general species B was also occurred frequently with acute febrile pharyngitis from the hospitalized patients and a small number of non-hospital patients who had symptoms severe enough to warrant the taking of diagnostic specimens. However, the HAdV serotypes from species B were not detected in raw sewage samples possible due to the transmission route of air and contact would reduce changes for HAdV B to be introduced into sewages.

Since gastroenteritis may cause releasing of pathogens in excreta, HAdV F41 would occur frequently in sewage as demonstrated from this study. However, another gastroenteritis-related HAdV species A was rarely found in this study like due to low affinity of PCR primers (hexAA1885/1913 and nehexAA1893/1905 had poor detection limit for HAdV A and B species, Kuo et al., 2009). Besides, HAdV species D, includes the epidemic conjunctivitis serotypes 8, 19 and 37 (Cooper et al., 1993), was seldom found in sewage of Taiwan.

This study showed that accurate positive cloning results. By the operational taxonomic units method, it has not been applied in cloning database. If a set of samples robustly clusters into a particular set of OTU, no matter what the addition order, this suggests that these OTU have some congruence with cloning taxa. The variability of positive colonies observed OTU is not a unique feature in this study.

4.7 HAdV diversity

HAdV F41 has been suggested as one of the most common serotypes causing gastroenteritis; thus, many of F41 would be released from infected individuals, collected by sewage system, and ended-up in influent sewages of treatment plants. In contrast, HAdV species C and D were generally the pathogenic agents causing pharyngitis, pneumonia, and conjunctivitis making them seldom been released into the sewages.

It had been suggested that HAdV species D and F were more resistant to natural environments and/or water treatment processes than others making them had commonly been detected even in drinking water (Heerden et al., 2005). Accordingly, serotypes from these species may have higher opportunity to be remained in sewage thus had been identified in this study.

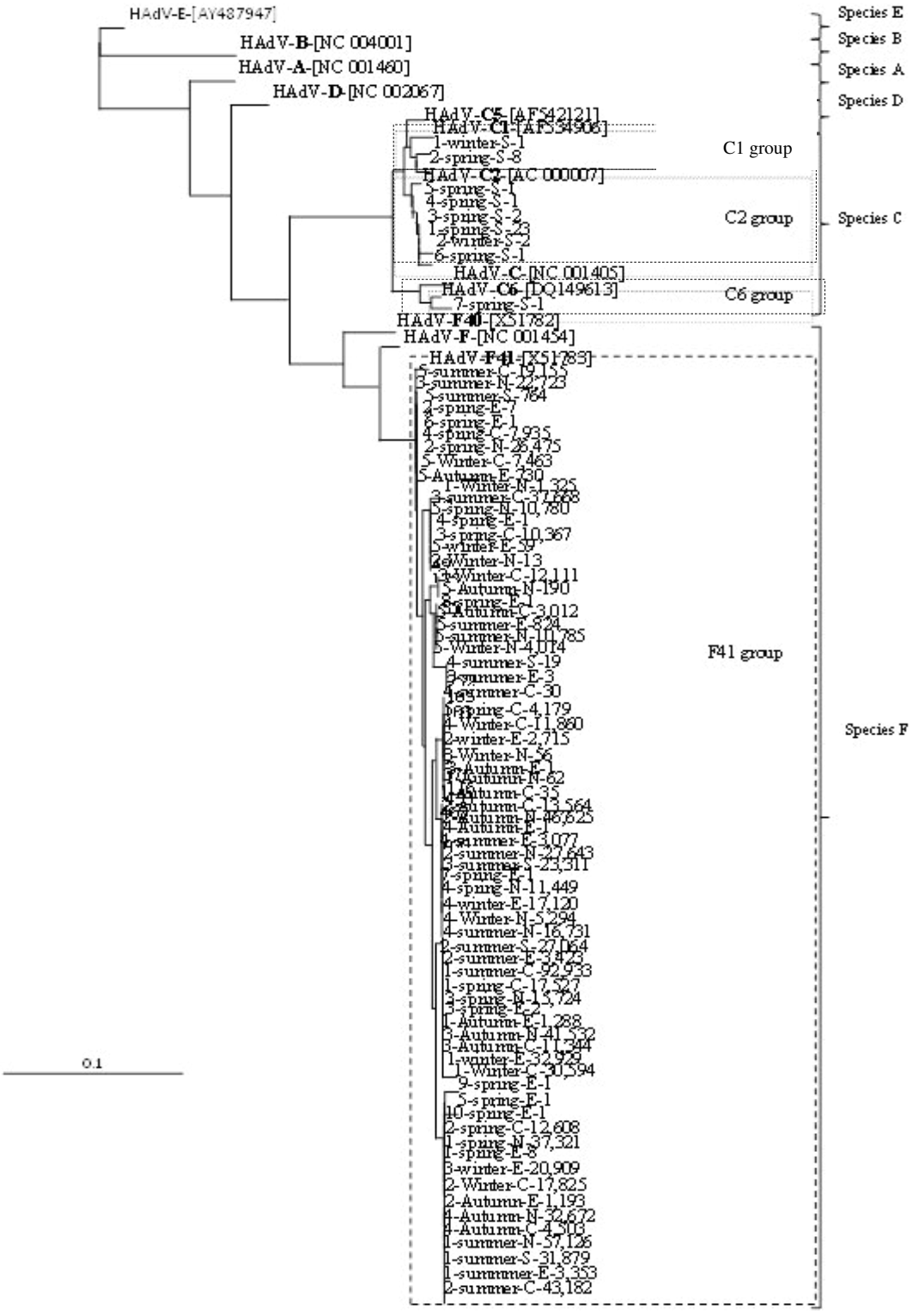


Fig 4.9 Phylogenetic tree based on sequence comparisons of a 143-bps hexon fragment been high throughput sequenced showing HAdV from raw sewage samples were clustered into three groups belonging to species C and F.

In a previous study, HAdV from A, C, and F species were for a municipal wastewater in northern America (Kuo et al., 2010). Interestingly, HAdVs from A species were not found in this study based on cloning-sequencing approach. This might be caused by different affinity between primers and templates. Another study had shown that, the nested PCR using primer sets of hexAA1885/1913 and nehexAA1893/1905 had poor amplification efficiencies for HAdV species A and B (Kuo et al., 2009). When target hexon gene was 10^4 copies for species A or B, the nested PCR still could not amplify it; while the same nested PCR amplified target hexon genes of C, D, E, and F species even when the levels were as low as 10^2-10^3 copies. Thus, hexon genes of HAdV-A species might not be able to be amplified and identified in this study. Practices of some serotype- specific primer sets for PCR amplification may allow some other HAdV serotypes been discovered (Nadan et al., 2003).

4.8 Seasonal and geological variation for HAdV distribution

To investigate geological distribution and seasonal dynamics of HAdV in sewage samples was one of the main objectives in this study. An unweighted-UniFrac method using principal coordinate analysis (PCA) with branch length normalization was used to determine seasonal and geological variations for the 16 sewage samples. A phylogenetic tree and distance matrix were used as input for the UniFrac analyses. In PCA, a distance matrix was used to plot certain amount of samples (n) in n dimensional spaces (Lozupone et al., 2007b). The factor through the space that describes as much variation as possible was principal coordinate 1. Orthogonal axes were subsequently assigned to explain as much of the variation not yet explained by previously assigned axes as possible. When few independent factors cause most of the variation, the first two or three principal coordinates often explain most of the variation in the data.

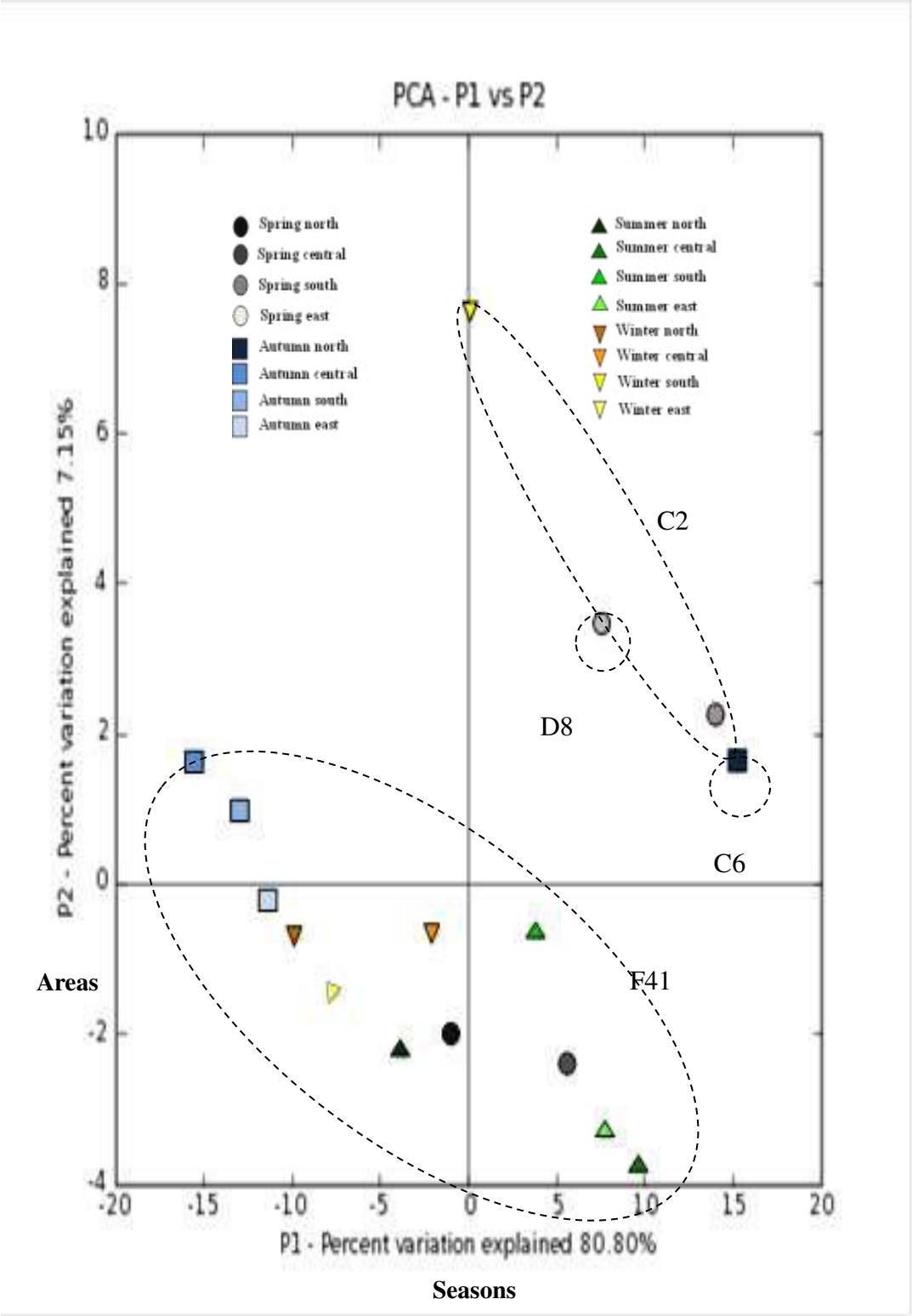


Fig 4.10 Principal coordinate analyses (PCA) for HAdV distribution during four seasons (the first principal component, P1) at four geological areas (P2).

In this study, the principal coordinates 1 describe 80.8% of the variation suggesting that HAdV distribution correlated strongly with season. The principal coordinates 2 explained 7.1% of the variation demonstrating HAdV distribution was not affected significantly by geological difference (Fig 4.10). This result suggested that HAdVs distribution in Taiwan would varied considerably because of seasonal difference. Furthermore, this study demonstrated that the HAdV distributions for the four geological areas during spring season were significantly different ($P \leq 0.05$) from the other three seasons suggesting diverse HAdV serotypes (e.g., F41, C2, and D8) may occur during spring in Taiwan. Besides, geological variation showed less significant impact on HAdV distribution may be caused by small data set (a total of 254 hexon sequences) obtained from cloning and sequencing approach. Larger data set may allow such statistic analyses to be improved and gain more precise results.

Depend on high throughput sequencing data, we also used UniFrac method by principal coordinate analysis (PCA) to determine seasonality and regional variations for 16 sewage samples. In this PCA program, the principal coordinates 1 described 68.26% of the variation suggesting that HAdV distribution correlated strongly with season. The principal coordinates 2 explained 10.05% of the variation indicating HAdV distribution was not affected significantly by geological difference (Fig 4.11). There was an variation evidence for HAdV seasonality variation result using UniFrac method compared with previous PCA program (cloning sequencing data) was consistent.

However, the test for seasonality was not carried out for clinical HAdV distribution due to different transmitted sources, this study clearly demonstrated annual seasonality of HAdV variations in the water environment. Environmental epidemiological studies would be required to support this hypothesis. In stark contrast, some meteorological factors

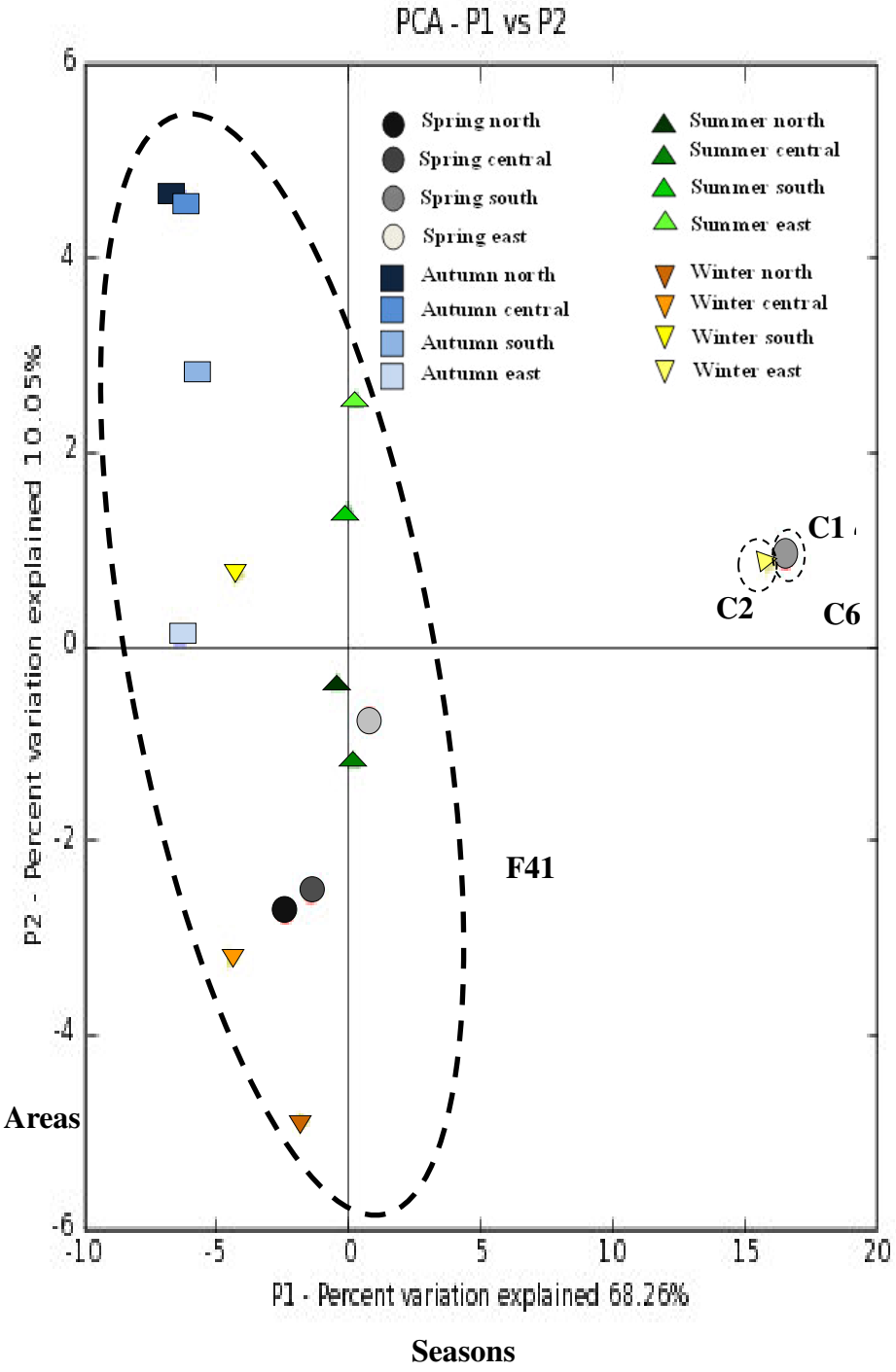


Fig 4.11 Principal coordinate analyses (PCA) for HAdV distribution during four seasons (the first principal component, P1) at four geological areas (P2)

probably associated with the HAdVs trends. Similar association between daily maximal fluctuations in temperature and humidity for respiratory infected viruses was reported previously (Chew et al., 1998). The reasons for this are not clear but it could be possible that greater fluctuations in seasonality may act on the HAdV distribution.

The observed association between influenza B virus and rainfall was also reported in northeast Brazil (De Arruda et al., 1991). But, we did not find a significant association between HAdVs trends and rainfall. This was likely due to variety conditions of meteorology in our study. Low relative humidity has also been shown to favour the survival of certain viruses while others have been shown to lose their infectivity after drying (Buckland et al., 1962). In these studies, respiratory tract viruses were found to occur year round during the rainy. In temperature communities, respiratory tract peaks during the winter months, whether in the southern or northern hemisphere (Gilchrist et al., 1994, Hierholzer et al., 1994, Thomas et al., 1994). A few serotypes showed evidence of seasonal variation. There was obvious seasonal variation with species C which showed incidences in three seasons except in summer. Although, this study analysed data by Unifrac for relatively few HAdV sequences, there was discernible seasonality in the more commonly environmental phenomenon (Buckland et al., 1962).

4.9 Real-time PCR quantification for HAdV

Quantities of overall HAdV and F41 in the 16 sewage samples were determined by two real-time PCR assays using TaqMan probes. The JTVXF/R primer set and JTVXP TaqMan probe were used for quantifying overall HAdV while the hex157f/246r primer set and hex214 TaqMan probe were used for F41. Two standard curves were prepared by running series-diluted plasmid DNA (i.e., 10^1-10^7 copies/μL) containing partial hexon

genes of C2 and F41 and plotting log concentrations with each of particular Ct values (Fig 4.12 and Fig 4.13). Each concentration level was amplified in triplicate and plotted as mean value with error bars representing standard deviation. All triplicate amplifications had very similar Ct values (Std < 0.5) and the linear regressions had R-square values close to 0.99 indicating the standard curves were fairly drivable.

The viral nucleic acid extracts from the sewage samples were quantified to determine concentrations of total HAdV and serotypes F41 using two quantitative real-time PCR assays. For the purpose of quantification, the data points (ΔRn) demonstrated during the extension step for each cycle was analyzed. The ΔRn was averaged for each cycle in amplification plot (Fig 4.14). The ΔRn value was plotted on the y-axis, and the x-axis represented cycle number. During the early cycles of the PCR amplification, the ΔRn value remained at base line. Most PCR amplifications reached a plateau phase as the reaction was carried out to high cycle numbers. The point crossed the threshold level was defined as Ct value. The HAdV concentrations were calculated using gained Ct values and the previous prepared standard curves.

The real-time PCR results showed that overall HAdV and serotype F41 were quantified in 16 and 15 raw sewage samples, respectively. The overall HAdV concentrations for the raw sewage were ranged from 8.8×10^{2} to 1.7×10^{8} viral particles/L (Table 4.5; Figs 4.15~18). Serotypes F41 concentrations were from 7.9×10^{2} to 1.8×10^{8} viral particles/L (Table 4.6; Figs 4.15~18). The serotype F41 had been quantified for spring-south sample but did not find in previous cloning-sequencing results suggesting some serotype might present in the sewage samples but had not been discovered by cloning-sequencing method. No clear profile was observed for the HAdV concentrations present across four seasons.

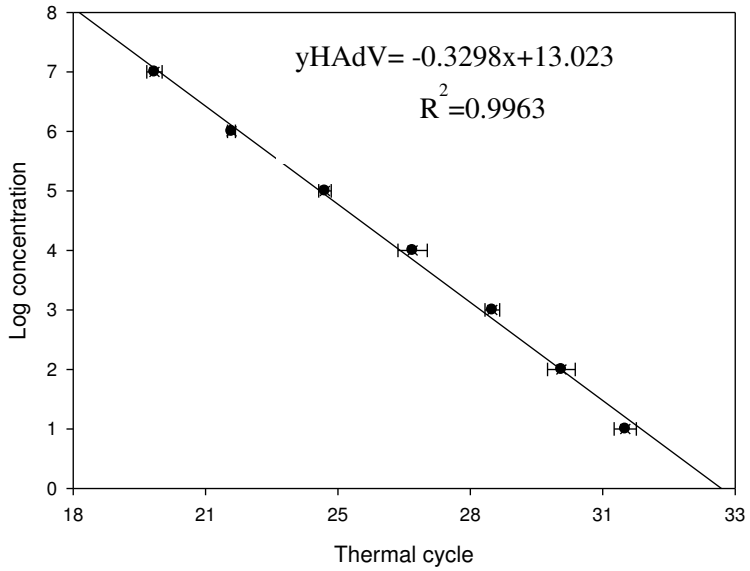


Fig 4.12 Standard curve prepared using quantified C2 hexon gene cloned in plasmid vector for the overall HAdV quantification.

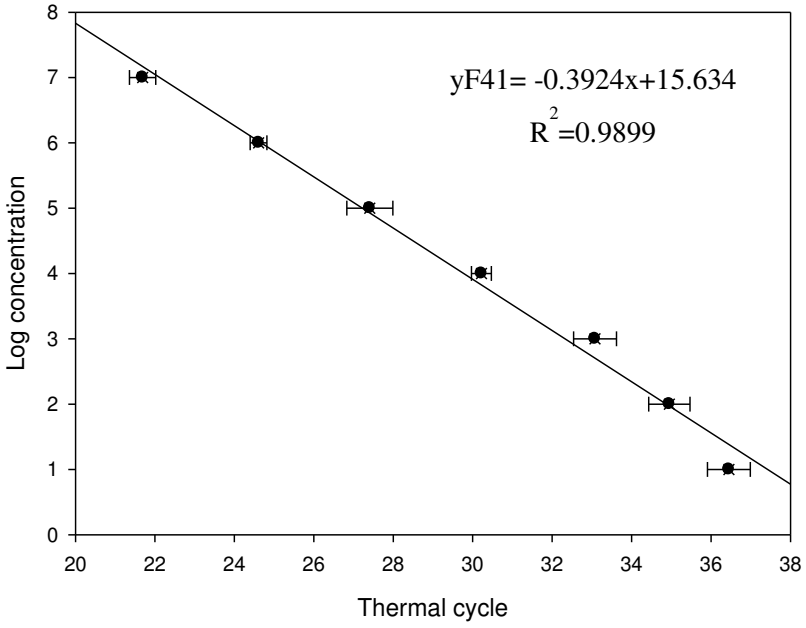


Fig 4.13 Standard curve prepared using quantified F41 hexon gene cloned in plasmid vector for the HAdV F41quantification.

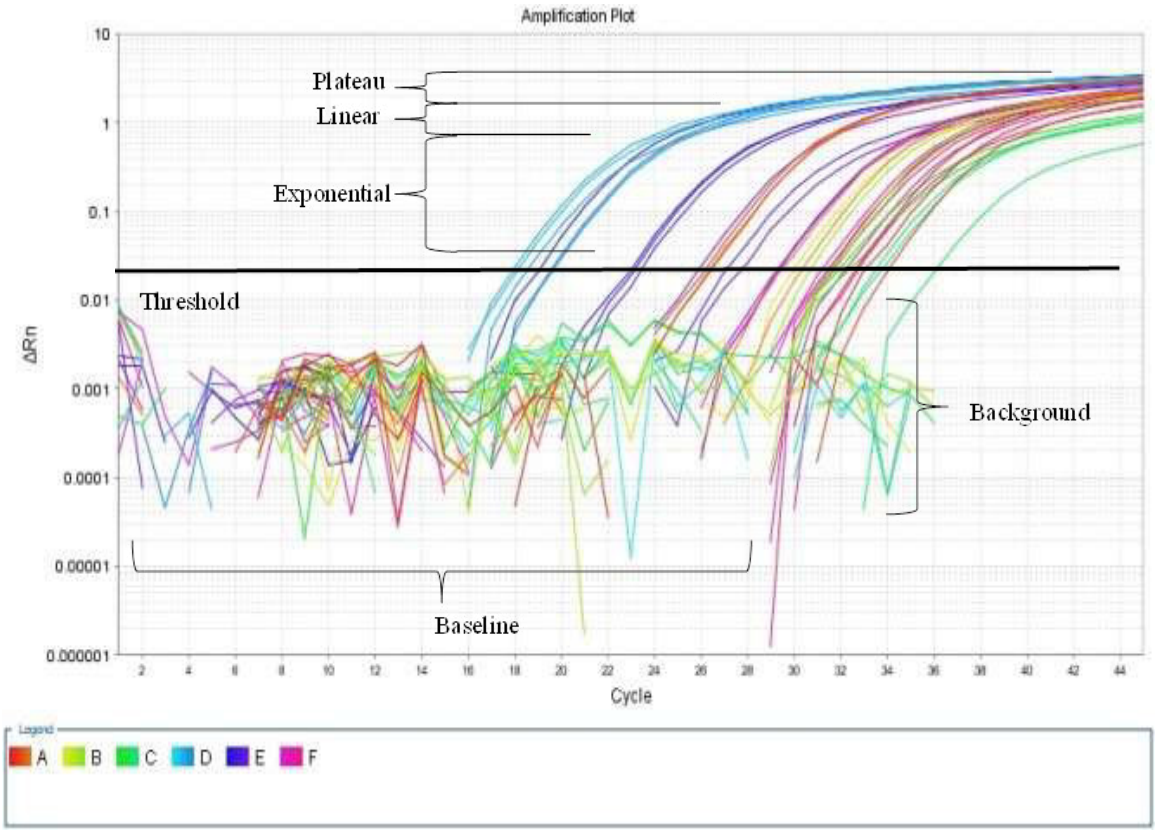


Fig 4.14 Amplification plots from the extension phase fluorescent emission data collected from the base line of the amplification plot. CT values are calculated by determining the point at which the fluorescence exceeds a threshold limit.

Table 4.5 Quantitative of overall species for HAdV in 16 sewage samples (particles/L)

	Seasons			
Areas	Autumn	Winter	Spring	Summer
North	1.78×10^{8}	1.89×10^{5}	3.95×10^{6}	8.24×10^{7}
Center	1.85×10^{6}	8.65×10^{5}	1.85×10^{6}	8.82×10^{2}
South	3.92×10^{6}	4.15×10^{6}	4.15×10^{4}	3.95×10^{6}
East	4.05×10^{5}	3.95×10^{6}	9.08×10^{7}	4.15×10^{5}

Table 4.6 Quantitative of serotype F41 for HAdV in 16 sewage samples (particles/L)

	Seasons			
Areas	Autumn	Winter	Spring	Summer
North	0	4.42×10^{4}	1.64×10^{6}	1.88×10^{8}
Center	1.79×10^{5}	2.94×10^{4}	2.47×10^{6}	7.92×10^{2}
South	4.69×10^{5}	7.69×10^{4}	1.19×10^{4}	6.65×10^{5}
East	2.69×10^{6}	2.42×10^{6}	4.05×10^{7}	4.42×10^{5}

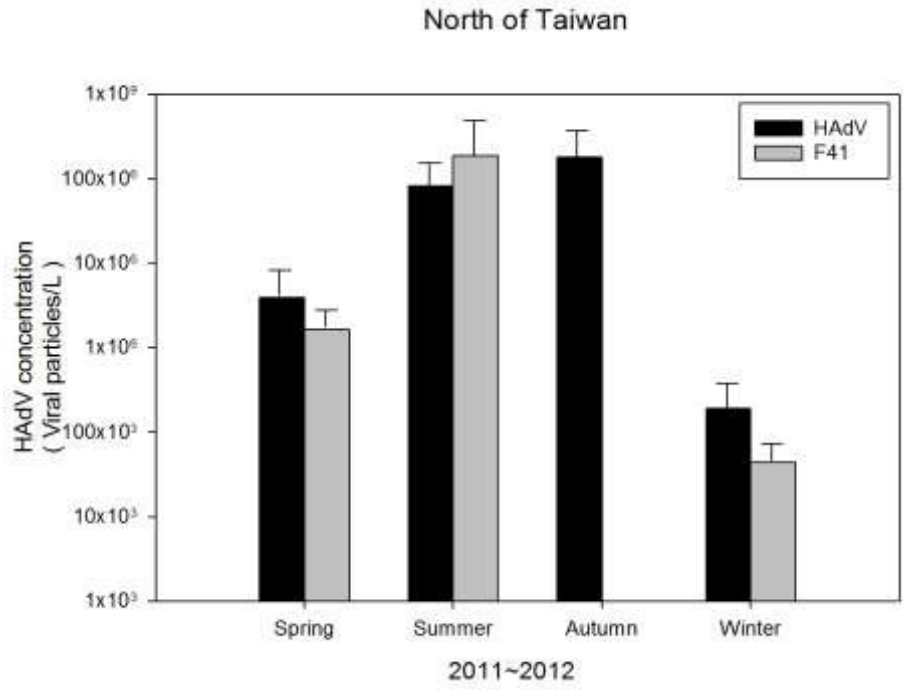


Fig 4.15 Concentrations of overall HAdV and serotypes F41 for the raw sewage samples by real-time PCR compared during four seasons in northern Taiwan.

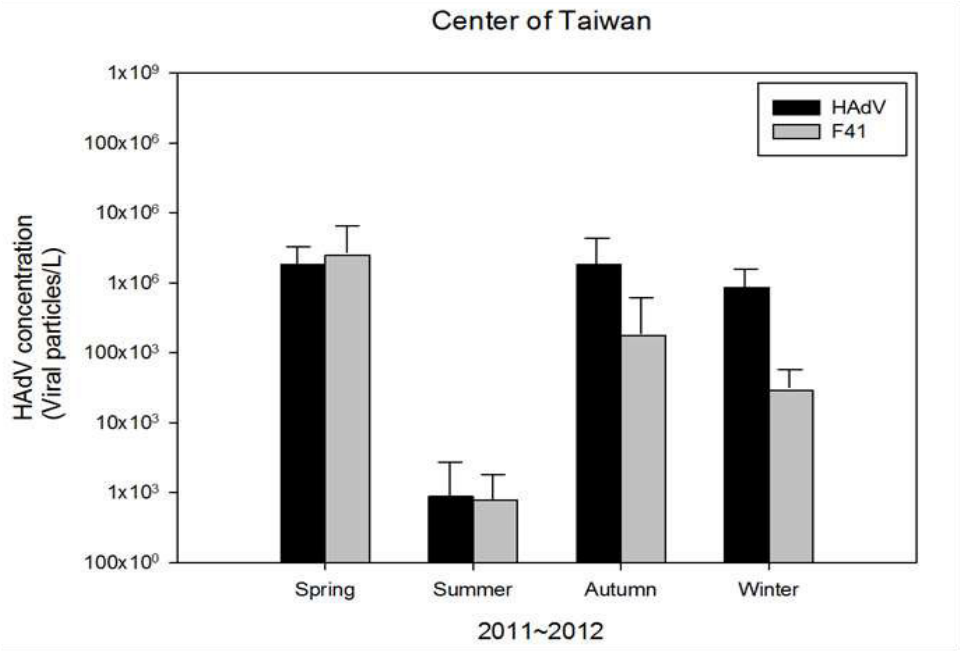


Fig 4.16 Concentrations of overall HAdV and serotypes F41 for the raw sewage samples by real-time PCR compared during four seasons in central Taiwan.

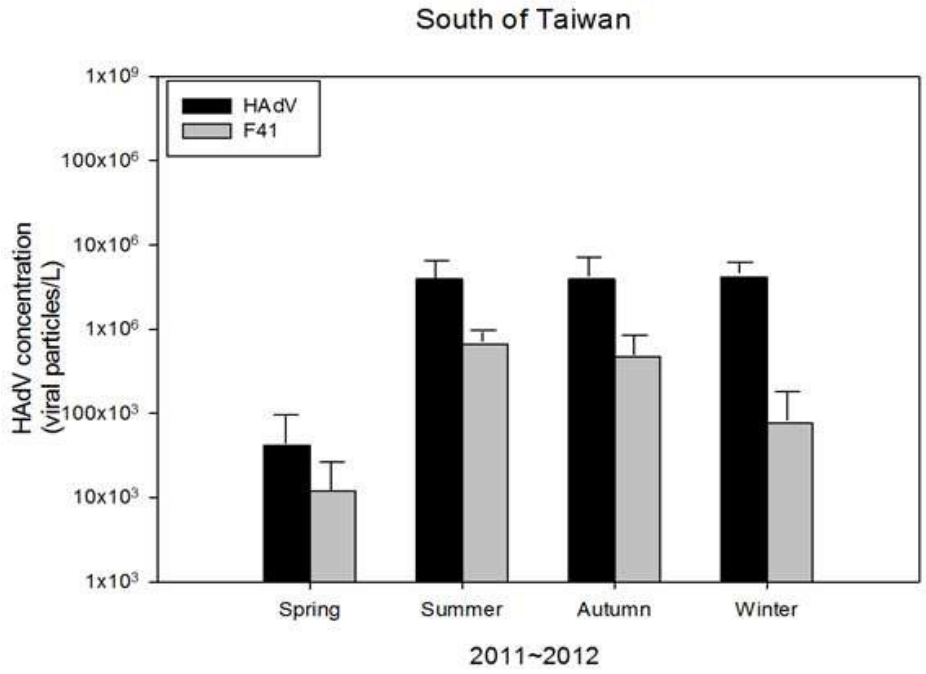


Fig 4.17 Concentrations of overall HAdV and serotypes F41 for the raw sewage samples by real-time PCR compared during four seasons in southernTaiwan.

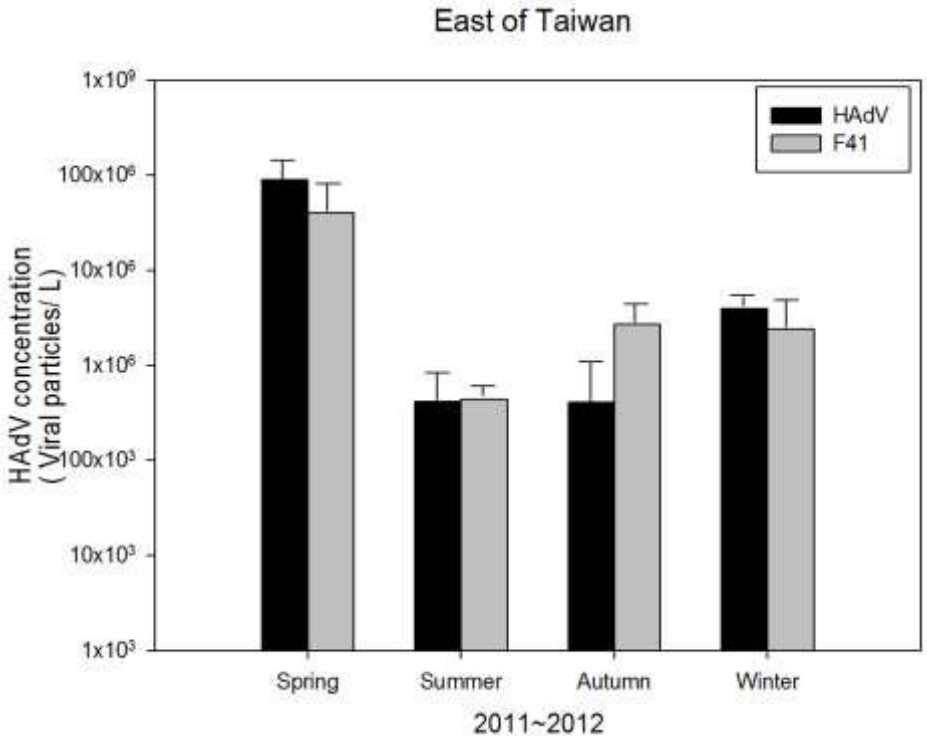


Fig 4.18 Concentrations of overall HAdV and serotypes F41 for the raw sewage samples by real-time PCR compared during four seasons in eastern Taiwan.

Chapter 5 Conclusions and suggestions

5.1 Conclusions

The major objectives of this study were to identify and quantify HAdV in 16 sewage samples collected from four areas during four seasons and to further assess HAdV distribution in these sewages as a function of differences in areas and seasons. Significant conclusions are summarized below.

In this study, HAdV serotypes F41, C2, C6, and D8 had been identified in 16 sewages of Taiwan area. HAdV-F41 was the most dominant serotype found in the sewages.

Serotype F41 was detected in 14 sewage samples. Serotype C2 was discovered in south of Taiwan during winter and spring seasons while serotypes C6 and D8 were detected in autumn-north and spring-east samples, respectively.

Species A was rarely found in winter of center Taiwan by high throughput sequencing and the common species B for sewages was not found in this study.

Based on Unifrac analyses, the seasonal difference had more significant effects on HAdVs distribution in Taiwan (80% variation) than geological variation.

Quantitative results of real-time PCR assay for all six adenovirus species (A-G) were ranged from 8.8×10^{2} to 1.7×10^{8} viral particles/L; and, concentrations of serotypes F41 were from 7.9×10^{2} to 1.8×10^{8} viral particles/L.

5.2 Suggestions

1. Studying the stability of HAdVs in treated water by analyzing different treated efficient of adenovirus samples.

2. Studying the mechanisms of adenovirus latency need further investigation such as miRNAs.

3. Studying other possible factors that could down regulate the viral gene expression affect DNA amplification.

4. Adenovirus could be cultured and purified antigen to produce antibody would be applied in sampling.

References

Abd-Jamil, J., Teoh, B.T., Hassan, E.H., Roslan, N., Abubakar, S., 2010.Molecular identification of adenovirus causing respiratory tract infection in pediatric patients at the University of Malaya Medical Center. BioMed Central pediatrics 10,46.

Allard, A., Albinsson, B., Wadell G., 1992. Detection of adenovirus in stools from healthy persons and patients with diarrhea by two-step polymerase chain reaction. Journal of Medical Virology 37,149–157.

Allard, A., Girones, R., Juto, P., Wadell, G., 1990. Polymerase chain reaction for detection of adenoviruses in stool samples. Journal of Clinical Microbiology 28, 2659–2667.

Aminu, M., Ahmad, A.A., Umoh, J.U., Beer, M.C., Esona, M.D., Steele, A.D., 2007. Adenoviruses infection in children with diarrhea disease in northwestern Nigeria. Annals of African Medicine 6, 168-173.

Ansorge, W.J., 2009. Next-generation DNA sequencing technique. N Biotechnol. 25,195-203.

Benko, M., Harrach, B., Both, G.W., Russell, W.C., Adair, B.M., Adam, E., de Jong, J.C., Hess, M., Johnson, M., Kajon, A., Kidd, A.H., Lehmkuhl, H.D., Li, Q.G., Mautner, V., Pring-Akerblom P., Wadell, G., Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (editors). Virus taxonomy: classification and nomenclature of viruses (8th report of the International Committee on the Taxonomy of Viruses). New York: Elsevier Academic Press, 213-228.

Benko, M., Harrach, B., Russell, W.C., 1999. Family adenoviridae in virus taxonomy.

Academic Press, 227-238.

Hsu, B.M., Chen, C.H., Wan, M.T. 2008. Prevalence of enteroviruses in hot spring recreation areas of Taiwan. FEMS immunology and medical microbiology 52, 253-259.

Hsu, B.M., Chen, C.H., Wan, M.T., Chang, P.J., Fan, C.W., 2009. Detection and identification of enteroviruses from various drinking water sources in Taiwan. Journal of Hydrology 365, 134-139.

Booth, J.L., Coggesall, K.M., Gordon, B.E., Metcalf, J.P., 2004 Adenovirus type 7 induces interleukin-8 in a lung slice model and requires activation of Erk. Journal of Virology 78, 4156-4164.

Bosch, A., 1998. Human enteric viruses in the water environment: a minireview. Internatl Microbiol 1, 191-196.

Brandt, C.D., Kim, H.W., Vargosko, A.J., Jeffries, B.C., Arrobio, J.O., Rindge, B., Parrott R.H., Chanock R.M., 1969. Infections in 18,000 infants and children in a controlled study of respiratory tract disease .I. Adenovirus pathogenicity in relation to serologic type and illness syndrome. Am. J. Epidemiol 90, 484-&.

Buckland, F.E., Tyrrell, D.A.J., 1962. Loss of infectivity on drying various viruses. Nature 195, 10634.

Cashdollar, J.L., Dahling, D.R., 2006. Evaluation of a method to re-use electropositive cartridge filters for concentrating viruses from tap and river water. Journal of Virology 132, 13–17.

Cashdollar, J.L. and Wymer, L., 2013. Methods for primary concentration of viruses from water samples: a review and meta-analysis of recent studies. Journal of Applied

Microbiology, 1364-5072.

Castingnolles, N., Petit, F., Mendel, I., Simon, L., Cattolico, L., 1998. Detection of adenovirus in the waters of the Seine river estuary by nested-PCR. Journal of Virology 12, 175–180.

Centers for Disease Control and Prevention (CDC), 2010. Outbreak of adenovirus 14 respiratory illness-Prince of Wales Island, 2008. MMWR Morb Mortal Wkly 59, 6-10.

Chang, L.T., Farrah, S.R., Bitton, G., 1981. Positively charged filters for virus recovery from wastewater treatment plant effluents. Applied Environmental Microbiology; 42: 921–924.

Chang, S.L., Stevenson, R.E., Bryant, A.R., Woodward, R.L., Kabler, P.W., 1958. Removal of Coxsackie and bacterial viruses in water by flocculation. II. Removal of Coxsackie and bacterial viruses and the native bacteria in raw Ohio River water by flocculation with aluminum sulfate and ferric chloride. Am J Public Health Nations Health 48, 159–169.

Chang, S.Y., Lee, C.N., Lin, P.H., Huang, H.H., Chang, L.Y., Ko, W., Chang, S.F., Lee, P.I., Huang, L.M., Kao, C.L., 2008. A community-derived outbreak of adenovirus type 3 in children in Taiwan between 2004 and 2005. Journal of Medical Virology 80, 102-112

Chapron, C.D., Ballester, N.A., Fontaine, J.H., Frades, C.N., Margolin, A.B., 2000. Detection of astroviruses, enteroviruses, and adenovirus type 40 and 41 in surface waters collected and evaluated by the Information Collection Rule and an integrated cell culture-nested PCR procedure. Applied Environmental Microbiology 66, 2520–5.

Charles, P., 2004. Why the concern about pathogens in water? Southwest Hydrology, 14-27.

Chen, H.L., Chiou, S.S., Hsiao, H.P.,Ke, G.M., Lin, Y.C., Lin, K.H., Jong, Y.J., 2013. Respiratory adenoviral infections in children: A study of hospitalized cases in southern Taiwan in 2001–2002. J Trop Pediatr 50, 279–284.

Cherry, J.D., 1998. Adenovirus. In: Feigin RD, Cherry JD, editors. Textbook of Pediatric Infectious Diseases, 4th edition. Philadelphia: W.B. Sannders Company, 1666–1684.

Chew, F.T., Doraisingham, S., Ling, A.E., Kumarasinghe, G., Lee, B.W., 1998. Seasonal trends of viral respiratory tract infections in the tropics. Epidemiology and infection 121, 121–128.

Cooper, R.J., Bailey, A.S., Killough, R., Richmond, S.J., 1993. Genome analysis of adenovirus 4 isolated over a six year period. Journal of Medical Virology 39, 62-66.

Cooper, R.J., Hallett, R., Tullo, A.B., Klapper, P.E., 2000. The epidemiology of adenovirus infections in Greater Manchester, UK 1982–96. Epidemiology and infection 125, 333–345.

Cox, P., Griffith, M., Angles, M., Deere, D., Ferguson, C., 2005. Concentrations of pathogens and indicators in animal feces in the Sydney watershed. Applied Environmental Microbiology 71, 5929-5934.

Cronin, A.A., Taylor, R.G., Powell, K.L., Barrett, M.H., Trowsdale, S.A., Lerner, D.N., 2003. Temporal variations in the depth-specific hydrochemistry and sewagerelated microbiology of an urban sandstone aquifer, Nottingham, United Kingdom. Journal of Virology 11, 205 – 216.

Dangelo, L. J., Hierholzer, J. C., Keenlyside, R. A., Anderson, L. J., Martone, W. J., 1979.

Pharyngoconjunctival fever caused by adenovirus type-4sreport of a swimming pool-related outbreak with recovery of virus from pool water. Journal of Infectious Diseases 140, 42-47.

Dangelo, L. J., Hierholzer, J. C., Keenlyside, R. A., Anderson, L.J., Martone, W. J., 1979. Pharyngoconjunctival fever caused by adenovirus type-4sreport of a swimming pool-related outbreak with recovery of virus from pool water. Journal of Infectious Diseases140, 42-47

Davison, A. J., Benko, M., Harrach, B., 2003. Genetic content and evolution of adenoviruses. Journal of Virology 84, 2895-2908.

De Arruda, N.E., Hayden, F.G., McAuliffe, J.F., 1991. Acute respiratory viral infections in ambulatory children of urban northeast Brazil. Journal of Infectious Diseases; 164: 252-8.

De Jong, J.C., Wermenbol, A.G., Verweij-Uijterwaal, M.W., Slaterus, K.W., Doornum, G.J.J., Khoo, S.H., Hierholzer, J.C., 1999. Adenoviruses from human immunodeficiency virus-infected individuals, including two strains that represent new candidate serotypes Ad50 and Ad51 of species B1 and D, respectively. Journal of Clinical Microbiology 37, 3940-3945.

Dunfield, K.E., King, G.M., 2004. Molecular analysis of carbon monoxide-oxidizing bacteria associated with recent Hawaiian volcanic deposits. Applied Environmental Microbiology 70, 4242–4248.

Echavarria, M., Zuckerman, A.J., Banatvala, J.E., Schoub, B.D., Griffiths, P.D., Mortimer, P., 2009. Principles and Practice of Clinical Virology, 6th ed. Wiley-Blackwell: John Wiley& Sons,463-488.

Elford, W.J., 1931. A new series of graded collodion membranes suitable for general bacteriological use, especially in virus studies. Journal of Pathology 34, 505–521.

Farrah, S.R. Bitton, G., 1979. Low molecular weight substitutes for beef extract as eluents for poliovirus adsorbed to membrane filters. Journal of Pathology 25, 1045–1051.

Farrah, S.R., Gerba, C.P., Wallis, C., Melnick, J.L., 1976. Concentration of viruses from large volumes of tap water using pleated membrane filters. Applied Environmental Microbiology 31, 221–226.

Fong, T., Lipp, EK., 2005. Enteric viruses of humans and animals in aquatic environmental: health risks, detection, and potential water quality assessment tools. Microbiology and Molecular Biology Reviews 69, 357-371.

Fuchs, J., Podda, M., 1998. Encyclopedia of medical genomics and proteomics, Marcel Dekker, USA.

Furuta, Y., Ohtani, F.M., Sawa, H., Kukuda, S., Inuyama, Y., 2001. Quantitation of varicella-zoster virus DNA in patients with Ramsay Hunt syndrome and zoster sine herpete. Journal of Clinical Microbiology 39, 2856-2859.

Garnett, C.T., Erdman, D., Xu, W., Gooding, L.R., 2002. Prevalence and quantitation of species C adenovirus DNA in human mucosal lymphocytes. Journal of Virology 76, 10608-10616.

Gibson, U.E.M., Heid, C.A., Williams, P.M., 1996. A novel method for real time quantitative competitive RT-PCR. Genome Res.

Gilchrist, S., Torok, **T.J.,** Gary, H.E., Alexander, J.P., Anderson, L.J., 1994. National surveillance for respiratory syncytial virus - United States, 1985-90. Journal Infectious

Diseases **170,** 986-90.

Glover, D.M., Hames, B.D., 1995. DNA cloning : Core techniques, 2nd edition, Oxford University Press, New York.

Gompf, S.G., Oehler, R.L., Adenoviruses. EMedicine Journal 2005. Available at: URL: http://www.emedicine.com/med/topics57.htm

Grabow, W.O.K., Puttergill, D.L., Bosch, A., 1992. Propagation of adenovirus types 40 and 41 in the PLC/PRF/5 primary liver carcinoma cell line. Journal of Virological Methods 37, 201-208.

Grabow, W.O.K., Taylor, M.B., de Villiers, J.C., 2001. New methods for the virological analysis of drinking water supplies. Water Science and Technology; 43, 1-8.

Greber, U.F., Willetts, M., Webster, P., Helenius, A., 1993. Stepwise dismantling of adenovirus 2 during entry into cells. Cell. Journal of Pathology 75, 477-486.

Greenberg, A., Taras, A.M.J., 1981. Standard Methods for the Examination of Water and Wastewater. Washington, DC: American Public Health Association.

Gu, Z., Belzer, S.W., Gibson, C.S., Bankowski, M.J., Hayden, R.T., 2003. Multiplex, real-time PCR for quantitative detection of human adenovirus. Journal of Clinical Microbiology 41, 4636-4641.

He, J.W., Jiang, S., 2005. Quantification of enterococci and human adenoviruses in environmental samples by real-time PCR. Applied and Environmental Microbiology 71, 2250-2255.

Heim, A., Ebnet, C., Harste, G., Pring-Akerblom, P., 2003. Rapid and quantitative detection of human adenovirus DNA by real-time PCR. Journal Medical Virology 70,

228-239.

Hierholzer, J.C., 1992. Adenoviruses in the immunocomprosmised host. Clinical Microbiology Reviews 5, 262-274.

Hierholzer, J.C., Tannock, G.A., Hierholzer, C.M., 1994. Subgrouping of respiratory syncytial virus strains from Australia and Papua New Guinea by biological and antigenic characteristics. Archives of virology 136**,** 133-47.

Hilleman, M.R., 1956. Acute respiratory illness caused by adenoviruses: a military problem. Armed Forces Medical Journal 7, 1717-1725.

Hong, J.S., Engler, J.A., 1996. Domains required for assembly of adenovirus type 2 fiber trimmers. Journal of Virology 70, 7071-7078.

Hsieh, W.Y., Chiu, N.C., Chi, H., Huang, F.Y., Hung, C.C., 2009. Respiratory adenoviral infections in Taiwanese children: a hospital-based study. Journal of microbiology 42, 371–377.

Jean, J.S., Guo, H.R., Shen, S.H., Liu, C.C., Chang, W.T., Yang, Y.J., Huang, M.C., 2006. The association between rainfall rate and occurrence of enterovirus epidemic due to a contaminated well. Journal of Applied Microbiology 101, 1224 – 1231.

Jiang, S., 2006. Human adenoviruses in water: occurrence and health implications: a critical review. Environmental Science Technology 40, 7132-7140.

Jiang, S., Noble, R, Chui, WP., 2001. Human adenoviruses and coliphages in urban runoff-impacted coastal waters of Southern California. Applied and Environmental Microbiology 67, 179-184.

Johansson, M.E., Andersson, M., Gudnadottir, S., 1993. Molecular epidemiology of

adenoviruses isolated in Iceland during 1988–1990. Journal of virological methods 42, 193–205.

Jones, M.S., Harrach, B., Ganac, R.D., Gozum, M.M.A., Cruz, W.P., Riedel, B., Pan, C., Delwarrt, E.L., Schnurr, D.P., 2007. New adenovirus species found in a patient presenting with gastroenteritis. Journal of Virology 81, 5978-5984.

Jothikumar, N., Cromeans, T.L., Hill, V.R., Lu, X., Sobsey, M.D., Erdman, D.D., 2005. Quantitative real-time PCR assays for detection of human adenoviruses and identification of serotypes 40 and 41. Applied Environmental Microbiology 71, 3131–3136.

Kajon, A.E., Wadell, G., 1992. Molecular epidemiology of adenoviruses associated with acute lower respiratory disease of children in Buenos Aires, Argentina (1984–1988). Journal of medical virology 36, 292–297.

Kampmann, B., Cubitt, D., Walls, T., Naik, P., Depala, M., Samarasinghe, S., Robson, D., Hassan, A., Rao, K., Gaspar, H., Davies, G., Jones, A., Cale, C., Gilmour, K., Real, M., Foo, M., Bennett-Rees, N., Hewitt, A., Amrolia, P. & Veys, P., 2005. Improved outcome for children with disseminated adenoviral infection following allogeneic stem cell transplantation. British Journal of Haematology 130, 595– 603.

Kidd, A.H., Jonsson, M., Garwicz, D., Kajon, A.E., Wermenbol, A.G., Verweij, M.W., De Jong, J.C., 1996. Rapid subgenus identification of human adenovirus isolates by a general PCR. Journal of Clinical Microbiology 34, 622-627.

Kukkula, M., Arstila, P., Klossner, M. L., Maunula, L., Bonsdorff, C. H. V., Jaatinen, P., 1997. Waterborne outbreak of viral gastroenteritis. Scand. Journal Infectious Diseses 29, 415-418.

Kuo, D.H.W., Simmons, F.J., Blair, S., Hart, E., Rose, J.B., Xagoraraki, I., 2010. Assessment of human adenovirus removal in a full-scale membrane bioreactor treating municipal wastewater. Water Research 44, 1520-1530.

Kuo, D.H.W., Simmons, F., Xagoraraki, I., 2009. A new set of PCR assays for the identification of multiple human adenovirus species in environmental samples. Journal of Applied Microbiology, 1364-5072.

Lee, J., Choi, E.H., Lee, H.J., 2010. Comprehensive serotyping and epidemiology of human adenovirus isolated from the respiratory tract of Korean children over 17 consecutive years (1991–2007). Journal of medical virology 82, 624–631.

Leen, A.M., Rooney, C.M., 2004. Adenovirus as an emerginh pathogen in immunocompromised patients. British Journal of Haematology 128, 135-144.

Levy, J.A., Fraenkel-Conrat, H., 1994. Owens RA. Virology, 3rd edition. Englewood Cliffs, NJ: Prentice-Hall, 185-192.

Lewis, P., Schmidt, M.A., Lu, X., Erdman, D.D., Campbell, M., Thomas, A., 2009. A community-based outbreak of severe respiratory illness caused by human adenovirus serotype 14. Journal Infectious Diseases 199, 1427-34.

Li, Q.G., Hambraeus, J., Wadell G., 1991. Genetic relationship between thirteen genome types of adenovirus 11, 34, and 35 with different tropisms. Intervirology 32, 338–350.

Lin, K.H., Lin, Y.C., Chen, H.L., Ke, G.M., Chiang, C.J., Hwang, K.P., Chu, P.Y., Lin, J.H., Liu, D.P., Chen, H.Y., 2004. A two decade survey of respiratory adenovirus in Taiwan. Journal Medical Virology 73,274–279.

Linnarsson, S., 2010. Recent advances in DNA sequencing methods – general principles of sample preparation. Experimental cell research 316, 1339-1343.

Lord, A., Bailey, A.S., Klapper, P.E., Snowden, N., Khoo, S.H., 2000. Impaired humoral responses to subgenus D adenovirus infections in HIV-positive patients. Journal of Medical Virology 62l, 405-409.

Lozupone, C.A. & Knight, R., 2007b. Global patterns in bacterial diversity. Processding of National Academy of sciences 104, 11436–11440.

Lozupone, C.A., Hamady, M., Kelley, S.T. & Knight, R., 2007a. Quantitative and qualitative {beta} diversity measures lead to different insights into factors that structure microbial communities. Applied Environment Microbiology 73, 1576–1585.

Lozupone, M. and Knight, R., 2005. UniFrac: a new phylogenetic method for comparing microbial communities. Environmental Microbiology 71, 8228-8235.

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J. & Turner, D.J., 2010. Target-enrichment strategies for next-generation sequencing. Nature Methods 7, 111-118.

Mandelboim, M., Dror, P., Azar, R., Bromberg, M., Mendelson, E., 2011. Adenovirus infections in hospitalized patients in Israel: epidemiology and molecular characterization. Journal of clinical microbiology 49, 597–601.

Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. Trends in Genetics. Journal Medical Virology 24, 133-141.

Martin, A.P., 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. Applied Environmental Microbiology 68, 3673–3682.

Mei, Y. F., Skog, J., Lindman, K., and Wadell, G., 2003. Comparative analysis of the genome organization of human adenovirus 11, a member of the human adenovirus

species B, and the commonly used human adenovirus 5 vector, a member of species . Journal of general virology 84, 2061–2071.

Metzgar, D., Osuna, M., Kajon, A.E., Hawksworth, A.W., Irvine, M., Russell, K.L., 2007. Abrupt emergence of diverse species B1 and B2 adenoviruses in US military recruit training centers. Journal Infectious Diseses 196, 1465-73.

Metzgar, D., Osuna, M., Yingst, S., Rakha, M., Earhart, K., 2005. PCR analysis of egyptian respiratory adenovirus isolates, including identification of species, serotypes, and coinfections. Journal of Clinical Microbiology 43, 5743–5752.

Nadan, S., Walter, J.E., Grabow, W.O.K., Mitchell, D.K., Taylor, M.B., 2003. Molecular characterization of astroviruses by reverse transcriptase PCR and sequence analysis: comparison of clinical and environmental isolates from South Africa. Applied and Environmental Microbiology 69, 747-753.

Niesters, H. G., van Esser, J., Fries, E., Wolthers, K. C., Cornelissen, J., and Osterhaus, A. D., 2000. Development of a real-time quantitative assay for detection of Epstein-Barr virus. Journal of Clinical Microbiology- 38,712–715.

Palomino, M.A., Larranaga, C., Villagra, E., Camacho, J., Avendano, L.F., 2004. Adenovirus and respiratory syncytial virus-adenovirus mixed acute lower respiratory infections in Chilean infants. The Pediatric infectious disease 23, 337–341.

Papapetropoulou, M., Vantarakis, A.C., 1998. Detection of adenovirus outbreak at a municipal swimming pool by nested PCR amplification. The Journal of Infection 36, 101-103.

Payne, S.B., Grilli, E.A., Smith, A.J., Hoskins, T.W., 1984. Investigation of an outbreak of

adenovirus type 3 infection in a boys' boarding school. Journal of general virology 93, 277–83.

Percival, S., Chalmers, R., Embrey, M., Hunter, P., Sellwood, J., Wyn-Jones, P., 2004. Microbiology of waterborne diseases. In: Enterovirus: Poliovirus, coxsackievirus, echovirus. Elsevier Academic Press, Amsterdam, 401 – 418.

Piatak, M.J., Saag, M.S., Yang, L.C., Clark, S.J., Kappes, J.C., Luk, K.C., Hahn, B.H., Shaw, G.M., and Lifson, J.D., 1993. High levels of HIV-a in plasma during all stages of infection determined by competitive PCR. Scence 259, 1749-1754.

Pina, S., Puig, M,, Lucena, F., Jofre, J., Girones, R.., 1998. Viral pollution in the environment and in shellfish: human adenovirus detection by PCR as an index of human viruses. Applied and Environmental Microbiology 64,3376-3382.

Puig, M., Jofre, J., Lucena, F., Allard, A., Wadell, G., Girones, R., 1994. Detection of adenoviruses and enteroviruses in polluted water by nested PCR amplification. Applied and Environmental Microbiology 60, 2963-2970.

Ramette, A., 2007. Multivariate analyses in microbial ecology. FEMS Microbiology Ecology 62, 142-160.

Reece, R.J., 2004. Analysis of genes and genomes, Chichester, West Sussex, England; Published by John Wiley&Sons.

Reither, K., Ignatius, R., Weitzel, T., Seidu-Korkor, A., Anyidoho, L., Saad, E., Djie-Maletz, A., Ziniel, P., Amoo-Sakyi, F., Danikuu, F., Danour, S., Otchwemah, R.N., Schreier, E., Bienzle, U., Stark, K., Mockenhaupt, F.P., 2002. Acute childhood diarrhea in northern Ghana: epidemiological, clinical and microbiological characteristics. BMC

Infectious Diseases 7, 104.

Ruuskanen, O., Metcalf, J.P., Meurman, O., Akusjarvi, G., Whitley, R.J., Hayden, F.G., 2009. Clinical Virology, 3th ed. American Society for Microbiology, 559-574.

Rux, J.J., Burnett, R.M., 2004. Adenovirus structure: review. Human Gene Therapy 15, 1167-1176.

Sanger, F., Nicklen, S. & Coulson, A.R., 1977. "Dna Sequencing with Chain-Terminating Inhibitors, Proceedings of the National Academy of Sciences of the United States of America 74, 5463-5467.

Sarkioja, M., 2008. Adenoviral gene therapy for non-small cell lung cancer [Dissertation]. University of Helsinki and Helsinki University Central Hospital Finland, Helsinki.

Schadt, C.W., Martin, A.P., Lipson, D.A. & Schmidt, S.K., 2003. Seasonal dynamics of previously unknown fungal lineages in tundra soils. Science 301, 1359–1361.

Schmidt, D., Wilson, M.D., Spyrou, C., Brown, G.D., Hadfield, J. & Odom, D.T., 2009. Using high-throughput sequencing to discover protein–DNA interaction, Methods,. Journal of general virology 48, 240-248.

Schmitz, H.,Wigand, R., Heinrich,W., 1983. Worldwide epidemiology of human adenovirus infections. Journal of Epidemiology 117, 455–66.

Selvaraju, S.B., Kovac, M., Dickson, L.M., Kajon, A.E., Selvarangan, R., 2011. Molecular epidemiology and clinical presentation of human adenovirus infections in Kansas Citychildren. Journal of clinical virology 51, 126–131.

Shields, P.A., Berenfeld, S.A. and Farrah, S.R., 1985. Modified membrane-filter procedure for concentration of enteroviruses from tap water. Applied Environment

Microbiology 49, 453–455.

Soumitra, R., Shirley, P.S., McClelland, A., Kaleko, M., 1998. Circumvention of immunity to the adenovirus major coat protein hexon. Journal of Virology 72, 6875-6879.

Stach, J.E., Maldonado, L.A., Masson, D.G., Ward, A.C., Goodfellow, M. & Bull, A.T., 2003. Statistical approaches for estimating actinobacterial diversity in marine sediments. Applied Environment Microbiology 69, 6189–6200.

Steward, G.F., and Culley, A.I., 2010. Extraction and purification of nucleic acids from viruses. Journal of general virology 92, 154–165.

Stewart, P.L., Fuller, S.D., Burnett, R.M., 1993. Difference imaging of adenovirus bridging the resolution gap between X-ray crystallography and electronmicroscopy. EMBO Journal 12- 2589-2599.

Stewart, PL, Burnett, RM, Cyrklaff, M, Fuller, SD., 1997. Image reconstruction reveals the complex molecular organization of adenovirus. Cell 67, 145-154.

Tate, J.E., Bunning, M.L., Lott, L., Lu, X., Su, J., Metzgar, D., 2009. Outbreak of severe respiratory disease associated with emergent adenovirus 14 at a US Air Force training facility in 2007. Journal Infectious Diseases 199, 1419-26.

Thomas, E., Margach, M.J., Orvell, C., Morriaon, B., Wilson, E., l994. Respiratory syncytial virus subgroup B dominance during one winter season between 1987 and 1992 in Vancouver, Canada. Journal Clinical Microbiology 32, 23842.

Tolentino, A., Ahkee, S. & Ramirez, J., 1996, Hot tub legionellosis. Journal of Medical Association 94, 393–394.

Tsai, H.P., Kuo, P.H., Liu, C.C., Wang, J.R., 2001. Respiratory viral infections among pediatric inpatients and outpatients in Taiwan from 1997 to 1999. Journal Clinical Microbiol 39,111–118.

Tsou, T.P., Tan, B.F., Chang, H.Y., Chen, W.C., Huang, Y.P., 2012. Community Outbreak of Adenovirus, Taiwan, 2011. Emerg Infect Dis 18,1825–1832.

USEPA, Information Collection Rule Technical.1996.

Van Heerden, J., Ehlers, M.M., Zyl, W.B., Grabow, W.O.K., 2003. Incidence of adenoviruses in raw and treated water. Water Research 37, 3704-3708.

Van Heerden, J., Ehlers, M.M., Vivier, J.C., Grabow, W.O.K., 2005. Risk assessment of adenoviruses detected in treated drinking water and recreational water. Journal of Applied Microbiology 99, 926-933.

Vellinga, J., Heijdt, V.D.S., Hoeben, R.C., 2005. The adenovirus capsid: major progress in minor proteins. Journal of General Virology 86, 1581-1588.

Vivier, J.C., Ehlers, M.M. & Grabow, W.O.K., 2004. Detection of enteroviruses in treated drinking water. Water Research 38, 2699–2705.

Wadell, G., 1999. Adenoviruses (Adenoviridae): general features of adenoviruses. Encyclopedia of Virology, 1-7.

Wallis, C. and Melnick, J.L., 1967. Concentration of enteroviruses on membrane filters. Journal of Virology 1, 472–477.

Wigand, R., Adrian, T., 1986. Classification and epidemiology of adenoviruses. Martinus Nijhoff Publishing Boston MA.

Xagoraraki, I., Kuo, D.H.-W., Wong, K., Wong, M. and Rose, J.B., 2007. Occurrence of

human adenoviruses at two recreational beaches of the Great Lakes. Applied Environmental Microbiology 73, 7874–7881.

Yeung, R., Eshaghi, A., Lombos, E., Blair, J., Mazzulli, T., 2009. Characterization of culture-positive adenovirus serotypes from respiratory specimens in Toronto, Ontario. Journal of virology 6, 11.

Yusof, M.A., Abdul- Rashid, T.R., Thayan, R., Othman, K.A., Abu-Hasan, N., Adnan, N., Saat, Z., 2012. Human adenovirus Type 7 Outbreak in Police Training Center, Malaysia. Emerging Infectious Diseases 18, 852-4.

Zhang, Q., Su, X., Gong, S., Zeng, Q., Zhu, B., Wu, Z., Peng, T., Zhang, C., and Zhou, R., 2006. Comparative genomic analysis of two strains of human adenovirus type 3 isolated from children with acute respiratory infection in southern China. Journal General. Virology 87, 1531–1541.

Zhu, Z., Zhang, Y., Xu, S., Yu, P., 2009. Outbreak of acute respiratory disease in China caused by B2 species of adenovirus type 11. Journal Clinical Microbiology 47, 697–703.

Appendix A Cloning-sequencing

Seasons	Areas	Sequences
Autumn	N	**C6**: 1~20
	C	**F41**:(I)1~4,6~8,10,11;(II)12,13,16,18~22
	S	**F41**: (I) 1,2,10; (II):1~15,17,18,20,21
	E	**F41**:1~4,6,7,9~12,15-19
Winter	N	**F41**:(I*)2,4,6,13,14;(II*) 1~15
	C	**F41**:(I)1,4,6,13,14;(II) 1~15
	S	**F41**:2,4,5,6,8~10 **C2**: 1,3,7,10~12
	E	**F41**: (I) 1,3,5; (II) 1~19
Spring	N	**F41**:(I)1~3,6-9,13~17; (II) 2~9
	C	**F41**:(I)2,3,5,6,8,11~15;(II)1~9
	S	**C2**: 12,19,23
	E	**F41**:7~14 **D8**: 1,2
Summer	N	**F41**: 1~20
	C	**F41**:15,16,18,21,26~28
	S	**F41**: 1~20
	E	**F41**: 21~27

Appendix B 1-19: high repeat sequence numbers

	Autumn				Winter				Spring				Summer			
	N	C	S	E	N	C	S	E	N	C	S	E	N	C	S	E
1	62,433	35,463		1,288	13,252	30,594	2	32,929	37,321	17,527	23	8	57,126	92,933	31,879	3,553
2	46,625	13,564		1,193	13,175	17,825	1	27,152	26,475	12,608	8	7	23,643	43,182	27,064	3,423
3	41,532	11,344		1,116	5,637	12,111		20,909	15724	10,367	1	2	23,537	37,668	23,311	3,165
4	32,672	4,503		1,020	5,294	11,860		17,120	11,449	7,935	1	2	16,731	30,103	19,752	3,077
5	19,055	3,012		730	4,014	7,463		5,949	10,780	4,179	1	1	11,119	19,155	7,644	824
6	14,994	1,639		726	4,012	5,839			8,074	3,024	1	1	7,137	14,199	6,175	814
7	14,877	1,397		688	1,084	3,412			5,112	2,452	1	1	6,655	13,510	5,991	812
8	12,846	1,208		622	981	2,455			3,513	1,807	1	1	5,785	6,550	5,021	803
9		565			940	2,408			344		1	1	3,242	1,761	2,042	
10					898						1	1	2,528	1,277	1,865	
11					647							1	2,069	890	1,475	
12					614							1	1,845	731	1,404	
13					589							1	988	683	1,330	
14					547								864	593	1,238	
15													683	579	1,217	
16													584	559	1,167	
17															1,027	
18															1,025	
19															518	
>500	24,503	72,695		7383	51684	93967	3	104059	118792	59899	39	28	165114	264373	141145	16471
<500	61,181	32,293		5170	61630	70612		43915	33671	20866			58994	61665	56944	6966
Total	306215	104988		12553	113314	164579	3	147974	152463	80765	39	28	224108	326038	198089	23437
	80.02%	69.24%		58.81%	45.61%	57.10%		70.32%	77.92%	74.16%			73.68%	81.09%	71.25%	70.28%