

Article

Not peer-reviewed version

# Genome Characterization and Phylogenetic Analysis of Scale Drop Disease Virus Isolated From Asian Seabass (*Lates calcarifer*)

[Putita Chokmangmeepisarn](#) , [Mohammad Noor Amal Azmai](#) , [Jose A. Domingos](#) , [Ronny van Aerle](#) ,  
David Bass , [Pochara Prukbenjakul](#) , [Saengchan Senapin](#) \* , [Channarong Rodkhum](#) \*

Posted Date: 12 June 2024

doi: 10.20944/preprints202406.0836.v1

Keywords: Scale loss; Metagenomics; Evolution; Mutations



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# Genome Characterization and Phylogenetic Analysis of Scale Drop Disease Virus Isolated from Asian Seabass (*Lates calcarifer*)

Putita Chokmangmeepisarn <sup>1</sup>, Mohammad Noor Amal Azmai <sup>2,3</sup>, Jose A. Domingos <sup>4</sup>, Ronny van Aerle <sup>5,6</sup>, David Bass <sup>5,6,7</sup>, Pochara Prukbenjakul <sup>1</sup>, Saengchan Senapin <sup>4,8,9,\*</sup> and Channarong Rodkhum <sup>1,\*</sup>

<sup>1</sup> Center of Excellence in Fish Infectious Diseases (CE FID), Faculty of Veterinary Science, Chulalongkorn University, Bangkok, 10330, Thailand; putita.ch.94@gmail.com channarong.r@chula.ac.th pochara.pruk@gmail.com

<sup>2</sup> Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia; mnamal@upm.edu.my

<sup>3</sup> Aquatic Animal Health and Therapeutics Laboratory, Institute of Bioscience, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

<sup>4</sup> Tropical Futures Institute, James Cook University, Singapore; jose.domingos1@jcu.edu.au

<sup>5</sup> International Centre of Excellence for Aquatic Animal Health, Centre for Environment, Fisheries and Aquaculture Sciences (Cefas), Weymouth, Dorset DT4 8UB, UK; ronny.vanaerle@cefas.gov.uk david.bass@cefas.gov.uk

<sup>6</sup> Centre for Sustainable Aquaculture Futures, University of Exeter, Stocker Road, Exeter EX44QY, UK

<sup>7</sup> Department of Life Sciences, the Natural History Museum, London SW7 5BD, UK

<sup>8</sup> Fish Health Platform, Center of Excellence for Shrimp Molecular Biology and Biotechnology (Centex Shrimp), Faculty of Science, Mahidol University, Bangkok, Thailand; saengchan@biotec.or.th

<sup>9</sup> National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Khlong Nueng, Pathum Thani, Thailand

\* Correspondence: Channarong Rodkhum: channarong.r@chula.ac.th; Saengchan Senapin: saengchan@biotec.or.th

**Simple Summary:** Scale drop disease virus (SDDV) has been identified as a significant pathogen causing scale drop syndrome (SDS), leading to significant economic losses in Asian seabass production in Southeast Asia. This study utilized a metagenomic approach to investigate the bacterial and viral communities associated with SDS, with a particular focus on SDDV and evaluated the potential of metagenomics for retrieving complete SDDV genomes. By characterizing the complete genomes of SDDV strains, we aimed to gain a deeper understanding of the virus. The insights gained from this study are expected to inform the development of comprehensive disease prevention and control strategies for SDDV, mitigating its impact on the aquaculture industry.

**Abstract:** Scale drop disease virus (SDDV), a double-stranded DNA virus in the family *Iridoviridae*, has been reported widely in southeast Asian countries as a causative agent of scale drop syndrome (SDS) in Asian seabass. SDS have resulted in high mortality and significant economic losses to the aquaculture industry. This study demonstrated the use of metagenomic methods to investigate bacterial and viral communities present in infected fish tissues and recovering a complete genome of the causative agent named SDDV TH7\_2019. Characterization of the TH7\_2019 genome revealed a genome size of 131 kb with 134 putative ORFs encoding viral proteins potentially associated with host apoptosis manipulation. A comparative genome analysis showed a high degree of amino acid identity across SDDV strains, with variations in number of repeat sequences and mutations within core genes. Phylogenetic analyses indicate a close relationship among SDDV genomes. This research enhances our understanding of the genetic diversity and evolutionary relationship of SDDV, contributing valuable insights for further development of effective control strategies of SDDV.

**Keywords:** scale loss; metagenomics; evolution; mutations

---

## 1. Introduction

Asian seabass (*Lates calcarifer*) is recognized as a euryhaline species, thriving in both brackish and nearshore marine environments. This species has considerable economic importance within the Asia-Pacific region. It is widely cultured in Australia, Singapore, Malaysia, Thailand, Indonesia, and China [1]. However, intensive aquaculture practices have resulted in a notable impact on Asian seabass cultures, particularly in their susceptibility to infectious diseases. Scale drop syndrome (SDS) represents a significant threat to aquaculture, particularly for Asian seabass cultivation. This disease was initially documented in an Asian seabass farm in Malaysia in 1992 [2]. Typically, affected fish exhibit scale loss over extensive areas, accompanied by skin discoloration, darkened bodies, gill pallor, tail, and fin erosion, as well as pathological features such as vasculitis and tissue necrosis in major internal organs [3]. Several pathogens have been implicated in causing SDS in Asian seabass, including scale drop disease virus (SDDV), *Vibrio harveyi*, and *Tenacibaculum maritimum* [2,4–8]. However, the current scientific consensus is that SDS in Asian seabass is caused by SDDV, resulting in significant economic losses in production of this valuable fish species [2,4,7,8].

SDDV is a double-stranded DNA virus belonging to the genus *Megalocytivirus*, family *Iridoviridae*. SDDV infections have been primarily observed in Asian seabass and reported in Southeast Asian countries including Thailand, Singapore, Malaysia, and Indonesia [1,3,6,7]. More recently, yellow seabream (*Acanthopagrus latus*) infected with SDDV were observed in China, exhibiting distinct clinical signs and pathological characteristics (swollen abdomen and ascites) compared to those observed in Asian seabass [9]. Understanding the genomic diversity of SDDV across Southeast Asia could potentially facilitate disease control and inform protective strategies against the disease. However, to date, only few SDDV genome sequences have been deposited to public sequence databases. Understanding the genomic diversity could greatly enhance disease control and inform protective strategies against the disease [10,11]. The first partial genome sequence of SDDV was reported from Asian seabass in Singapore [4], followed by complete genomes from the same fish species in Thailand [12]. Additionally, an SDDV genome sequence was reported from yellow seabream in China [9]. A genome comparison between the first SDDV isolated in Singapore and the Thai SDDV revealed a high degree of sequence identity (99.97%), along with some notable variations, such as genome size and mutations [12]. Despite the regional significance of SDDV, there is limited information available regarding the genomic characteristics, strain diversity, and phylogenetic relationships of SDDV, particularly strains originating from different regions. In this context genome characterization of different SDDV strains could elucidate factors influencing host susceptibility, virulence, and geographical distribution.

Metagenomics offers a powerful approach for characterizing viral and microbial communities, overcoming the limitations of traditional culturing methods [13]. This technique is particularly advantageous when prior knowledge of the organisms is scarce, co-infections are suspected, or comprehensive pathogen identification is required [14]. Several studies have successfully employed metagenomics to retrieve complete genomes of pathogens from diverse sample types [12,15,16].

This study aimed to use a metagenomic approach to investigate bacterial and viral communities associated with SDS in diseased Asian seabass isolated from a selection of Southeast Asian countries. We particularly focused on identifying potential pathogens co-infecting SDS affected fish. Additionally, this study evaluated the potential of metagenomics for retrieving complete SDDV genomes and elucidating their phylogenetic relationships. By characterizing the complete genome of different SDDV strains, we aimed to gain a deeper understanding of the virus, which could ultimately inform the development of comprehensive disease prevention and control strategies for SDDV.

## 2. Materials and Methods

### 2.1. Fish Samples and DNA Extraction

A total of eight Asian seabass (*Lates calcarifer*) were collected from farms in Thailand, Singapore, and Malaysia during disease outbreaks from 2016 to 2019 from our previous studies (Table 1) [7,8]. A single fish collected from Singapore in 2019 was a juvenile (length 30 cm, weight 293 g), while the two fish collected from Malaysia in 2019 were adults (length 30 - 40 cm, weight 450 -700 g). Fish from both countries were cultured in sea cages with water salinities ranging from 28 – 31 ppt. Fish collected from Thailand during 2016-2019 were juveniles (weights ranging from 60 - 280 g) cultured in floating cages with water salinity of 10 ppt. All fish displayed typical clinical signs of SDS including lethargic, darkened bodies, fin, and tail rot, easily detached scales, severe scale loss, and hemorrhages, were collected for tissue collection including brain, kidney, liver, spleen, and fin. Tissues were preserved in RNAlater™ reagent (Thermo Fisher Scientific, Massachusetts, USA) and kept at -20 °C until use. Genomic DNA was extracted using a conventional phenol/chloroform method [17]. All fish samples were screened for SDDV infection using SDDV-specific semi-nested conventional PCR as described in [18] to confirm SDDV infection prior to sequencing.

**Table 1.** The information of metagenomic samples collected from diseased fish.

| Metagenomic sample no. | Geographical origin   | Year | Organ                                            |
|------------------------|-----------------------|------|--------------------------------------------------|
| 2                      | Chanthaburi, Thailand | 2016 | Liver                                            |
| 3                      | Chanthaburi, Thailand | 2016 | Liver                                            |
| 4                      | Chanthaburi, Thailand | 2017 | Liver                                            |
| 6                      | Chanthaburi, Thailand | 2018 | Liver                                            |
| 7                      | Chanthaburi, Thailand | 2019 | Fin                                              |
| 12                     | Singapore             | 2019 | Kidney                                           |
| 21                     | Selangor, Malaysia    | 2019 | Pool internal organs (liver, spleen, and kidney) |
| 23                     | Selangor, Malaysia    | 2019 | Pool internal organ (liver, spleen, and kidney)  |

2.2. Library Preparation and Sequencing

Genomic DNA extracted from fin (sample 7), kidney (sample 12) and pooled internal organs (kidney, spleen, and liver, sample 21 and 23) were subjected to metagenome sequencing. DNA concentrations of all samples were quantified using a Qubit™ fluorometer (Thermo Scientific, USA) and adjusted to 100 ng/μL prior to sequencing. Sequence library preparation was carried out using NEBNext® Ultra DNA Library Prep Kit for Illumina (NEB, MA, USA) following the manufacturer’s instructions. The libraries were sequenced using the Illumina HiSeq 1000 in 150 bp paired-end mode. The sample sequencing was carried out by Suzhou GENEWIZ Biotechnology company (Suzhou, China).

2.3. Sequence Analysis

2.3.1. Reads Quality Control and Host Removal

After obtaining the metagenomic sequencing datasets, low quality reads (Q<20) and adapter sequences were removed using Trimmomatic v 0.39 [19]. Subsequently, host sequence reads were removed by mapping the reads of each of the samples to the Asian seabass reference genome (GCA\_001640805.1) using Bowtie2 v 2.5.1 [20].

2.3.2. Taxonomic Classification

The host-depleted reads were taxonomically classified according to the lowest common taxonomic ancestor (LCA) using Kraken2 v 2.1.3 with the NCBI viral and bacterial RefSeq complete genomes/proteins (downloaded August 2023) [21]. Virus and bacteria abundance estimates were reported at family level and normalized as a proportion of the sequencing reads in each of the samples using Bracken v 2.8 [22]. Moreover, the top 10 most abundant bacteria and virus families of



each sample were determined in reads per million (RPM) using formula: [the number of reads classified to a specific taxonomic group /total number of classified reads]  $\times 10^6$ . Relative taxa abundance bar plot and heatmap (based on RPM for each sample) were generated using ggplot2 [23] in R v 4.2.2 [24].

### 2.3.3. Metagenome Assembly and SDDV Genome Recovery

The host-depleted reads were assembled using MEGAHIT v 1.2.9 and contigs shorter than 200 bp were discarded [25]. Sequence similarity searches of the assembled contigs were conducted using NCBI BLASTx [26] and the full NCBI non-redundant protein (nr) database (downloaded on August 2024) with e-value cut off  $1 \times 10^{-5}$  to identify viral contigs, specifically SDDV. Subsequently, the SDDV contigs were subtracted from the samples that contained relative abundance of SDDV more than 70% (sample 4, 6, 7, 12, 21, and 23). Draft SDDV genomes were further subjected to genome assembly quality assessment by mapping all reads obtained for each sample to the SDDV TH2019 genome (MN562489) using Bowtie2 v 2.5.1 to determine genome size, depth, and breadth of read coverages [20]. Additionally, genome completeness was evaluated using Benchmarking Universal Single-Copy Orthologs (BUSCO v 5.7.1) analysis based on the conserved gene set of the iridoviridae\_odb10 database [27].

### 2.3.4. SDDV Genome Annotation

Open reading frames (ORFs) were predicted using Prodigal v 2.6.3 and their functions were predicted based on NCBI BLASTp [26] searches against the NCBI nr and conserved domain (CD) database (downloaded on January 2023) [28]. Genome segments were reoriented using progressive MAUVE using the SDDV TH2019 genome (MN562489) as reference before comparative genome analysis [29]. For comparative genome analysis, all SDDV genomes available in NCBI, C4575 (NC\_027778) and ZH-06/20 (OM037668), were included. Pairwise comparison of whole genome and multiple sequence alignment (MSA) was performed using MAFFT v 7 with the FFT-NS-I strategy [30]. Dot plots were generated using LAST local alignment implemented in MAFFT with SDDV TH2019 on the horizontal axis [30]. Nucleotide and amino acid sequence similarities of SDDV TH7\_2019 against other SDDV genomes were estimated using BLAST [26] and visualized using pyGenomeViz [31].

### 2.3.5. Comparative Genomics and Phylogenetic Analyses

Due to limitation of availability of gene sequences and completeness of the retrieved SDDV genomes in this study, maximum likelihood (ML) phylogenetic trees were constructed based on i) whole genome sequences (WGS), ii) single nucleotide polymorphisms (SNPs), iii) six concatenated iridoviruses core genes, and iv) individual iridoviruses core genes. Sequences of megalocytiviruses (MCVs) including ISKNV (NC\_003494), red seabream iridovirus (RSIV) (MK689686), European chub iridovirus (ECIV) (MK637631), and turbot reddish body iridovirus (TRBIV) (GQ273492) were included in the phylogenetic analysis while Singapore grouper iridovirus (SGIV) (AY521625) from family *Ranavirus* was used as an outgroup. A phylogenetic tree based on WGS was constructed using MAFFT-derived MSA and the IQ-TREE web interface with 1,000 bootstraps through the ultrafast bootstrap approximation approach [32]. The best fit nucleotide substitution model was selected according to the lowest Bayesian information criterion (BIC) score using ModelFinder [33]. SNP calling with SDDV TH2019 as a reference and subsequent SNP-based phylogenetic analysis were carried out using the Call SNPs & Infer Phylogeny (CSI) web service (<https://cge.food.dtu.dk/services/CSIPhylogeny/>) [34]. All available iridoviruses core gene sequences including DNA polymerase, DNA-dependent RNA polymerase II alpha subunit (RPO), DNA-dependent RNA polymerase II beta subunit (RPO2), D5 family NTPase ATPase, NTPase, flap endonuclease, ATPase, myristylated membrane protein, NIF-NLI interacting factor, Uvr/REP helicase, and major capsid protein (MCP) were retrieved from the SDDV genomes for phylogenetic analysis based on individual genes and concatenated core genes (Table S1). The six iridovirus core

gene sequences (D5 family NTPase ATPase, flap endonuclease, ATPase, myristylated membrane protein, NIF-NLI interacting factor, and Uvr/REP helicase) were concatenated and aligned using MEGA X [35] and the best fit nucleotide substitution models were selected according to the lowest BIC score. Maximum likelihood phylogenetic analyses were performed using MEGA X with 1,000 bootstraps [35].

Nucleotide and amino acid substitutions within the core genes were determined manually using the previous sequence alignments and SDDV TH2019 was used as reference for result interpretation. Tandem repeats were identified using Tandem Repeats Finder (TRF) with default parameters (<https://tandem.bu.edu/trf/trf.html>) [36].

Pan-genome analysis of the four complete SDDV genomes (TH7\_2019, TH2019, C4575, and ZH-06/20) was carried out using Roary v 3.13.0 [37] with identity threshold of 95%. Genes present in three or more genomes were defined as core genes whereas genes present in less than three genomes were defined as accessory genes.

3. Results

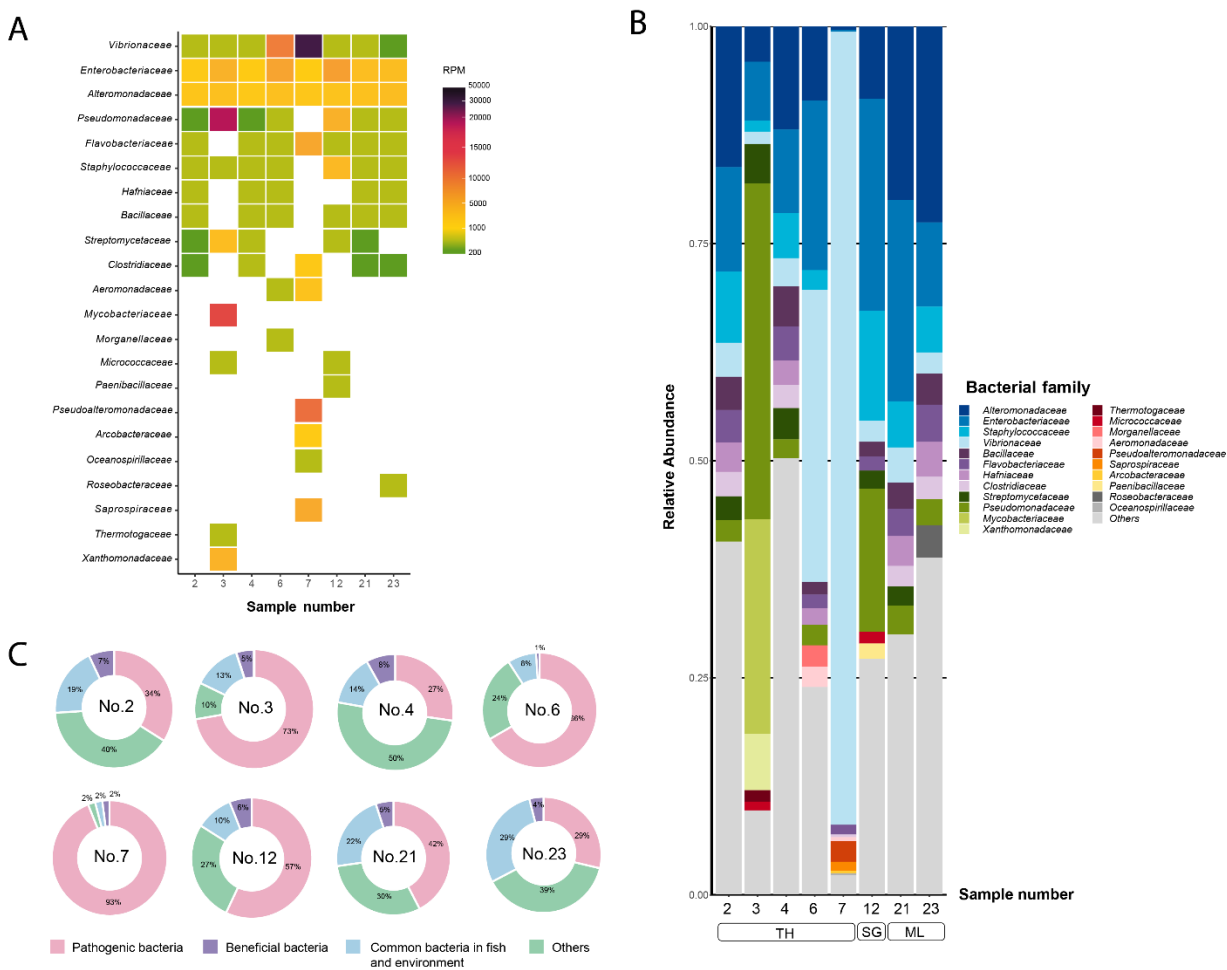
3.1. Taxonomic Profiles of Bacteria and Viruses

Metagenomic sequencing resulted in a total of 22 to 30 million reads per sample and contained approximately 87% to 94% of host reads. After read trimming and host removal, the number of sequences that remained for each sample ranged from 1.5 to 3.2 million. The host-depleted read pairs were subjected to taxonomic classification using Kraken2 and the percentage number of classified reads ranged from 1% to 6%, except sample 7 which exhibited remarkable percentage of microbial reads (42.81%) compared with other samples (Table 2). Among these microbial reads, the percentages of reads representing bacteria ranged from 84% to 99% with sample 23 possessing the lowest and sample 7 the highest number of bacterial reads, respectively. The taxonomic analysis revealed a total of 479 bacterial families across all samples.

Table 2. Classification of read counts of metagenomic samples using Kraken2.

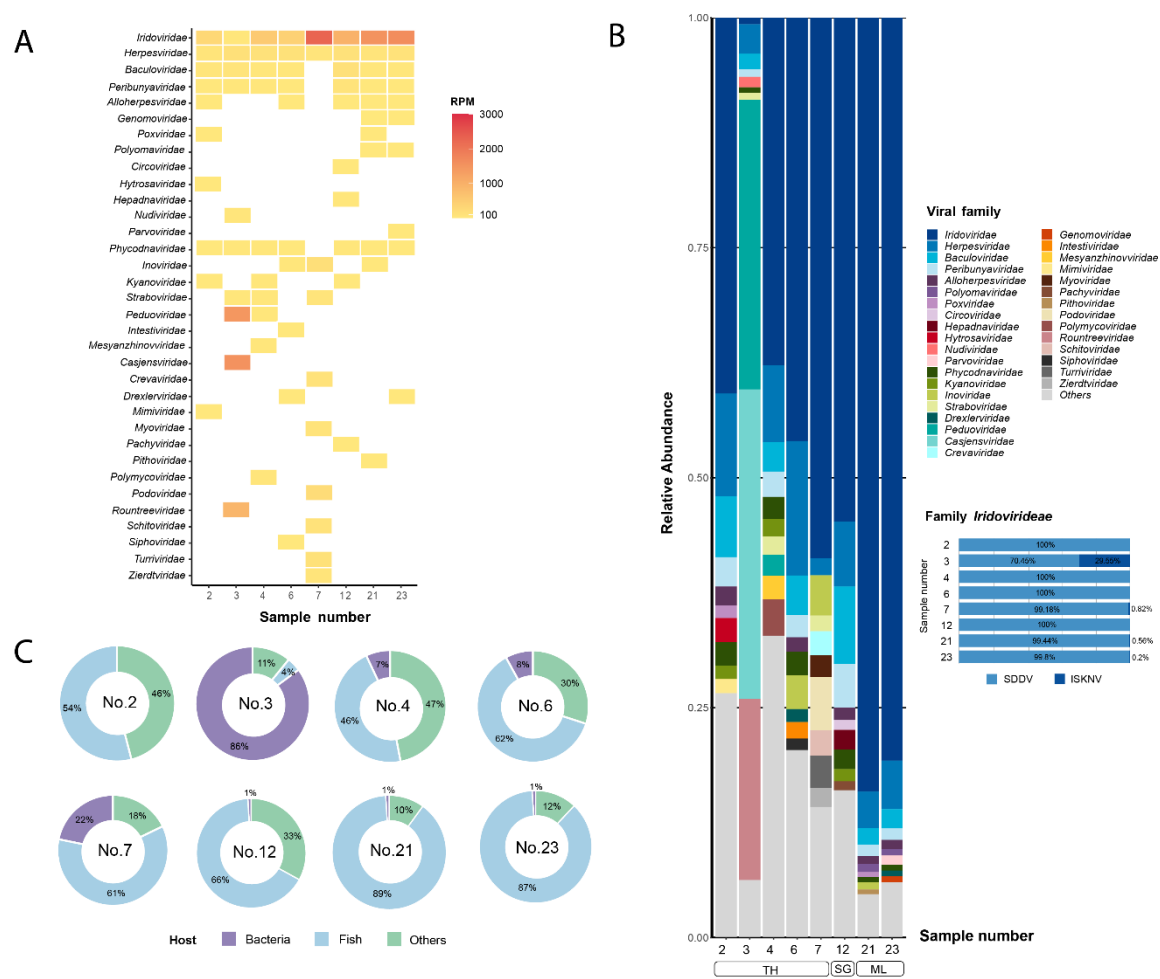
| Metagenomic sample no. | Number of raw reads | Unclassified reads (%) | Classified reads (%) | Bacterial reads (%) | Viral reads (%) |
|------------------------|---------------------|------------------------|----------------------|---------------------|-----------------|
| 2                      | 1,577,409           | 98.90                  | 1.10                 | 93.13               | 6.87            |
| 3                      | 3,012,060           | 94.16                  | 5.84                 | 92.04               | 7.96            |
| 4                      | 3,144,439           | 98.70                  | 1.30                 | 88.75               | 11.25           |
| 6                      | 1,422,958           | 97.30                  | 2.70                 | 96.62               | 3.38            |
| 7                      | 2,844,173           | 57.19                  | 42.81                | 99.09               | 0.91            |
| 12                     | 2,017,399           | 97.55                  | 2.45                 | 92.54               | 7.46            |
| 21                     | 1,515,485           | 98.77                  | 1.23                 | 85.10               | 14.90           |
| 23                     | 2,096,503           | 98.69                  | 1.31                 | 84.42               | 15.58           |

A heat map based on RPM was constructed to highlight the diversity of bacterial families across the samples (Figure 1A). The top 10 most abundant families present within each sample were selected for calculating family-level relative abundances, resulting in 22 dominant families (with the remaining families grouped as "others") (Figure 1B). Among these, nine families were identified as containing species that are known fish pathogens, including *Staphylococcaceae*, *Enterobacteriaceae*, *Vibrionaceae*, *Flavobacteriaceae*, *Pseudomonadaceae*, *Morganellaceae*, *Hafniaceae*, *Mycobacteriaceae*, and *Aeromonadaceae*. They were represented by high proportions of the total bacterial reads obtained for several samples (Figure 1C). The top 5 most prevalent bacterial families were *Enterobacteriaceae*, *Alteromonadaceae*, *Vibrionaceae*, *Staphylococcaceae*, and *Flavobacteriaceae*. Notably, *Vibrionaceae*, Gram negative bacteria known for causing SDS, exhibited a remarkably high abundance in samples 6 (33%) and 7 (91%). Meanwhile, sample 3 displayed a distinct bacterial profile, characterized by a high abundance of *Mycobacteriaceae* (24%), which was absent in other samples and high abundance of *Pseudomonadaceae* (38%).



**Figure 1.** Taxonomic analysis of bacterial reads at family level. The top 10 bacterial families are shown. (A) Heat map represents reads per million (RPM) values. The RPM values are represented according to color gradient legend in the right panel. (B) Bar plot represents bacterial diversity in terms of relative abundance, (C) Pie charts represent proportion of bacterial families based on their impact on fish health. TH: Thailand, SG: Singapore, ML: Malaysia.

Regarding virus taxonomic analysis, a total of 146 virus families were identified. A heat map showing the abundance (RPM) and the diversity of the virus families among samples is shown in Figure 2A. The top 10 most abundant virus families in each sample were chosen for family-level relative abundances calculation, resulting in 34 dominant families (with the remaining families grouped as "others") (Figure 2B). Among these, five families contain viruses that are known fish pathogens, while 22 families were representing bacteriophages. The top 5 most prevalent virus families were *Iridoviridae*, *Herpesviridae*, *Baculoviridae*, *Peribunyaviridae*, and *Phycodnaviridae*. The proportion of virus families known to be fish pathogens were dominated in all samples except sample 3 which exhibited a remarkably high proportion of bacteriophage families (85%) (Figure 2C). Notably, the family *Iridoviridae*, which includes SDDV, displayed the highest abundance in all samples except sample 3, which displayed its own unique viral community characteristics. Consistent with the bacterial taxonomic profile, sample 3 exhibited a unique virus profile with higher abundance of prokaryotic viruses, including *Casjensviridae* (33%), *Peduoviridae* (31%), and *Rountreeviridae* (19%). In all samples, *Iridoviridae* were predominantly represented by SDDV sequences, with the exception of sample 3 where approx. 30% of the *Iridoviridae* sequences were ISKNV sequences, while the remaining 70% were SDDV sequences (Figure 2B).



**Figure 2.** Taxonomic analysis of viral reads at family level. The top 10 viral families are shown. (A) Heat map represents reads per million (RPM) values. The RPM values are represented according to color gradient legend in the right panel. (B) Bar plot represents viral diversity in terms of relative abundance and proportion of species within the *Iridoviridae* family, (C) Pie charts represent proportion of viral families based on host. TH: Thailand, SG: Singapore, ML: Malaysia.

3.2. SDDV Genome Recovery

Following *de novo* assembly, the generated contigs were subjected to taxonomic classification. Notably, the largest contig (approximately 131 kb) was obtained from sample number 7. SDDV contigs were successfully retrieved from seven out of eight samples, with the exception of sample 3. The SDDV genomes retrieved from samples 2, 4, 6, 7, 12, 21, and 23 were named as SDDV strain TH2\_2016, TH4\_2017, TH6\_2018, TH7\_2019, SG12\_2019, ML21\_2019, and ML23\_2019, respectively. The assembled SDDV genomes exhibited varying degrees of coverage, ranging from 12% to 100% (sample 7; see Table 3). Furthermore, genome completeness was assessed using the BUSCO (Benchmarking Universal Single-Copy Orthologs) approach, evaluating a set of 10 single-copy iridovirus genes (Table S2). The result indicated the highest completeness within SDDV TH7\_2019 genome with following score: Complete (C): 10 (100%), Fragmented (F): 0, Missing (M): 0, followed by SDDV ML23\_2019 with following score: C: 8 (80%), F: 2 (20%), M: 0, and SDDV ML21\_2023 with following score: C: 7 (70%), F: 3 (30%), M: 0. Therefore, a complete genome was recovered only from sample 7 (SDDV TH7\_2019) and submitted to NCBI under the accession no. PP660347. Apart from the BUSCO set of iridovirus genes, complete major capsid protein (MCP) gene sequences were obtained from SDDV ML23\_2019 and submitted to NCBI (Table S2).

**Table 3.** Assembly statistics of SDDV genomes retrieved from metagenomic samples.



| Metagenomic sample no. | Number of SDDV contigs | SDDV strain name | Length (bp) | Percent covered <sup>1</sup> | BUSCO search |                |             |
|------------------------|------------------------|------------------|-------------|------------------------------|--------------|----------------|-------------|
|                        |                        |                  |             |                              | Complete (C) | Fragmented (F) | Missing (M) |
| 2                      | 36                     | TH2_2016         | 16,035      | 12.3                         | 1            | 0              | 9           |
| 4                      | 126                    | TH4_2017         | 90,045      | 80.0                         | 1            | 6              | 3           |
| 6                      | 71                     | TH6_2018         | 30,800      | 23.1                         | 1            | 2              | 7           |
| 7                      | 1                      | TH7_2019         | 131,759     | 100                          | 10           | 0              | 0           |
| 12                     | 113                    | SG12_2019        | 108,163     | 81.9                         | 3            | 7              | 0           |
| 21                     | 79                     | ML21_2019        | 127,278     | 92.9                         | 7            | 3              | 0           |
| 23                     | 28                     | ML23_2019        | 130,742     | 98.9                         | 8            | 2              | 0           |

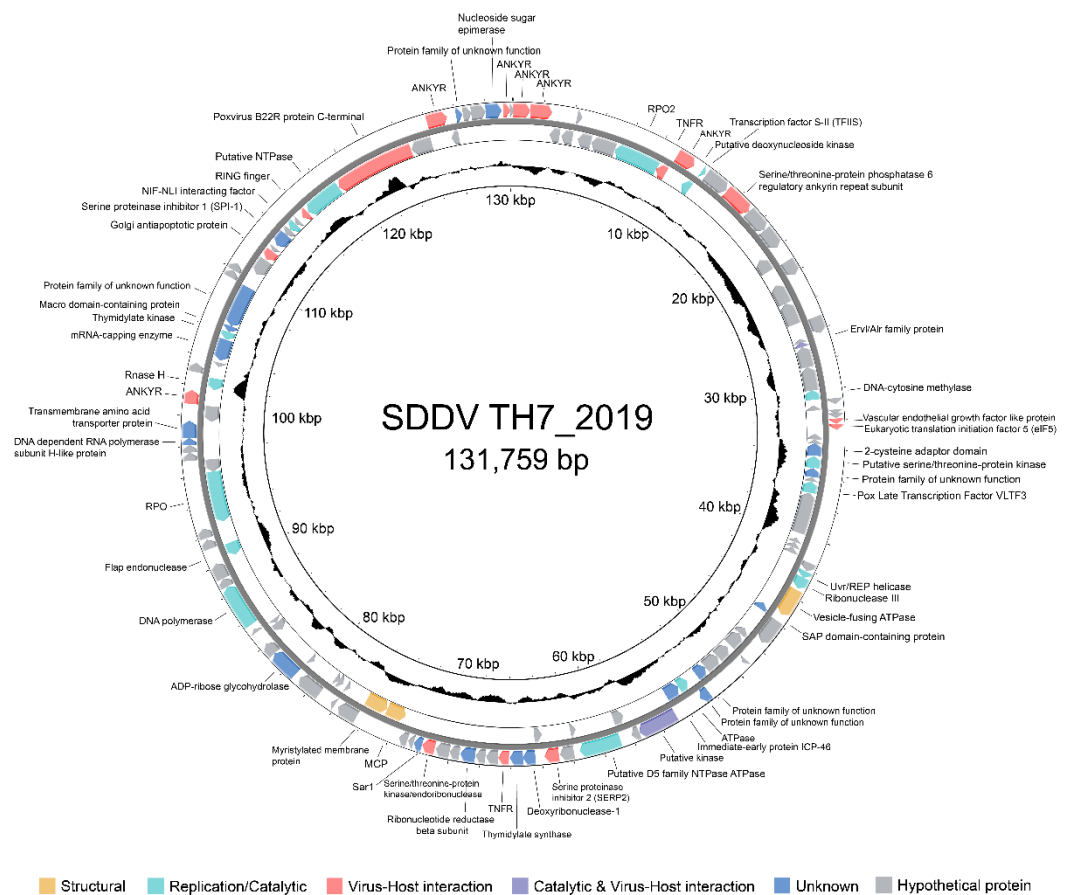
<sup>1</sup> Percent covered of query genomes compared with SDDV TH2019.

3.3. SDDV Genome Characterization

The genome of SDDV strain TH7\_2019 was subjected to genome annotation to predict ORFs and their potential function. The SDDV TH7\_2019 genome size was 131,759 bp with a G+C content of 36.6% (Table 4). A total of 134 ORFs encoding putative proteins were predicted based on BLASTp searches against nr and CD database. A circular genome map of the SDDV TH7\_2019 genome is shown in Figure 3. Detailed information on the predicted ORFs, including their lengths, homologous proteins based on BLASTp results, and predicted functions, is presented in Table S3. All predicted ORFs were homologous to SDDV and other iridovirus genes. Among these, 38 ORFs were functionally annotated, including 26 iridoviruses core genes. Functions of viral proteins were categorized into four types; i) Structural proteins: structural and virion-associated elements; ii) Replication/Catalytic proteins: essential proteins in DNA replication and transcription, or support virus replication and adaptation, or cell signaling; iii) Virus-Host interaction: virulence or immune evasion proteins, and iv) Unknown function. Several proteins that are crucial for the virus life cycle, including MCP (ORF079), DNA-dependent RNA polymerase II alpha (RPO) (ORF099) and beta (RPO2) (ORF008) subunit were identified. Moreover, viral proteins involved with virulence, including eukaryotic translational initiation factor 2 $\alpha$  (eIF2 $\alpha$ ) (ORF032), ribonuclease III (ORF046), Ery1/Alr family (ORF023), and tumor necrosis factor receptor (TNFR) homologs (ORF009 and 069) were present.

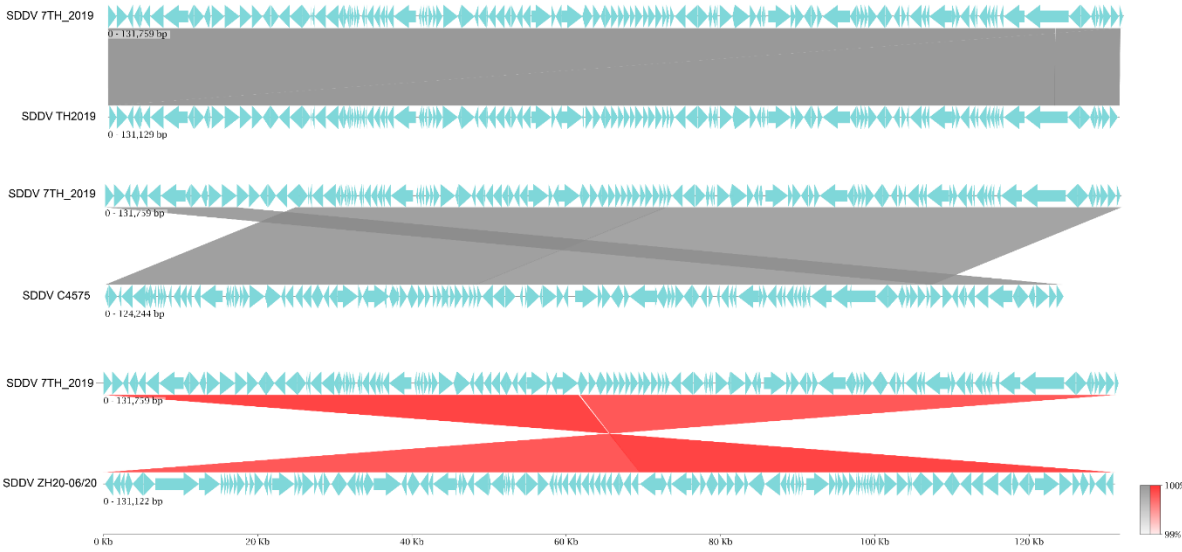
**Table 4.** General features of genomes used in this study.

| Species                                             | Isolate name | Host                                                 | Year | Geographical origin | GC (%) | Length (bp) | ORFs | GenBank accession no. | Reference  |
|-----------------------------------------------------|--------------|------------------------------------------------------|------|---------------------|--------|-------------|------|-----------------------|------------|
| Scale drop disease virus (SDDV)                     | TH7_2019     | Asian seabass ( <i>Lates calcarifer</i> )            | 2019 | Thailand            | 36.6   | 131,759     | 134  | PP660347              | This study |
| Scale drop disease virus (SDDV)                     | TH2019       | Asian seabass ( <i>Lates calcarifer</i> )            | 2018 | Thailand            | 36.6   | 131,192     | 135  | MN562489              | [12]       |
| Scale drop disease virus (SDDV)                     | C4575        | Asian seabass ( <i>Lates calcarifer</i> )            | 2015 | Singapore           | 37.0   | 124,244     | 129  | KR139659              | [4]        |
| Scale drop disease virus (SDDV)                     | ZH-06/20     | Yellow seabream ( <i>Acanthopagrus latus</i> )       | 2020 | China               | 36.56  | 131,122     | 135  | OM037668              | [9]        |
| Infectious spleen and kidney necrosis virus (ISKNV) |              | Mandarin fish ( <i>Siniperca chuatsi</i> )           | 2001 | China               | 54.78  | 111,362     | 124  | AF371960              | [52]       |
| European chub iridovirus (ECIV)                     | LEC15001     | European chub ( <i>Squalius cephalus</i> )           | 2005 | United Kingdom      | 38.5   | 128,216     | 108  | MK637631              | [49]       |
| Turbot reddish body iridovirus (TRBIV)              |              | Turbot ( <i>Scophthalmus maximus</i> )               | 2006 | China               | 55.0   | 110,104     | 114  | GQ273492              | [67]       |
| Red seabream iridovirus (RSIV)                      | KagYT-96     | Japanese amberjack ( <i>Seriola quinqueradiata</i> ) | 1996 | Japan               | 53.0   | 112,719     | 117  | MK689686              | [68]       |
| Singapore grouper iridovirus (SGIV)                 |              | Brown-spotted grouper ( <i>Epinephelus tauvina</i> ) | 2004 | Singapore           | 48.5   | 140,131     | 162  | AY521625              | [54]       |



**Figure 3.** Circular genome map of the SDDV TH7\_2019 genome. The outer and inner rings represent sense and antisense strands, respectively. Arrows indicate open reading frames (ORFs) and the direction of their transcripts. The ORFs are colored based on their functions.

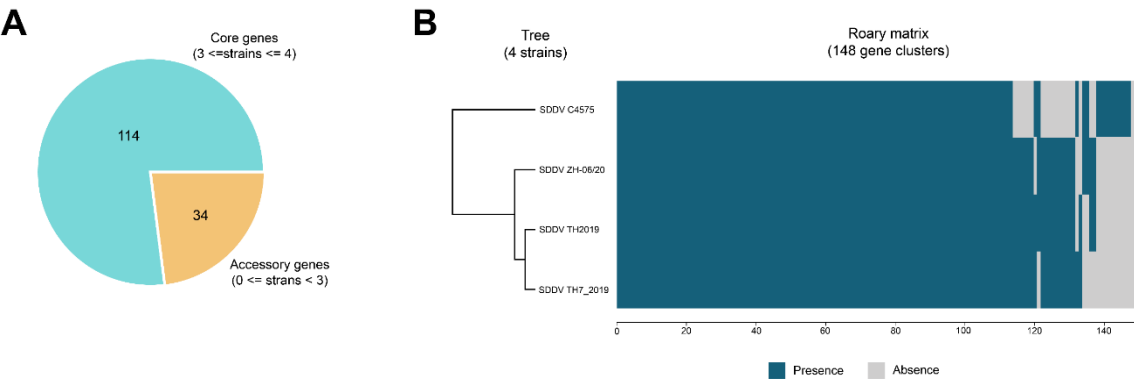
Pairwise comparison demonstrated high similarity between the SDDV TH7\_2019 genome and other previously published SDDV genomes. The SDDV TH7\_2019 genome exhibited 99.80%, 99.71%, and 99.69% nucleotide identity with SDDV strains TH2019, ZH-06/20, and C4575, respectively (Figure 4). Dot plot analysis indicated substantial collinearity across the genomes, with the SDDV ZH-06/20 strain exhibiting an inverted direction (Figure S1). Gene by gene comparisons of predicted ORFs from the SDDV TH7\_2019 genome indicated high amino acid similarity (<95% to 100%) for 110 to 122 ORFs of 134 ORFs when compared to other SDDV strains. Conversely, the number of ORFs exhibiting lower sequence similarity (<90%) were 6 to 14 (Table S4). Some of these variations can be attributed to the presence of varying numbers of tandem repeat sequences in ORFs. The most variable number of tandem repeats was found in ORF076 and ORF077 (SDDV TH7\_2019), resulting in relatively low sequence similarities (78% to 79%) between Thai SDDV strains and other SDDV strains (Table S5). Variations in the number of tandem repeats were also observed in ORF051 and ORF088, resulting in sequence similarities of 97% and 98%, respectively.



**Figure 4.** Linear map of whole genome sequence alignment of SDDV TH7\_2019 against SDDV TH 2019, C4575, and ZH-06/20. Grey linkages indicate nucleotide similarity percentages and red linkages indicate nucleotide similarity percentages of inverted sequences.

3.4. Comparison Genomics and Phylogenetic Analysis

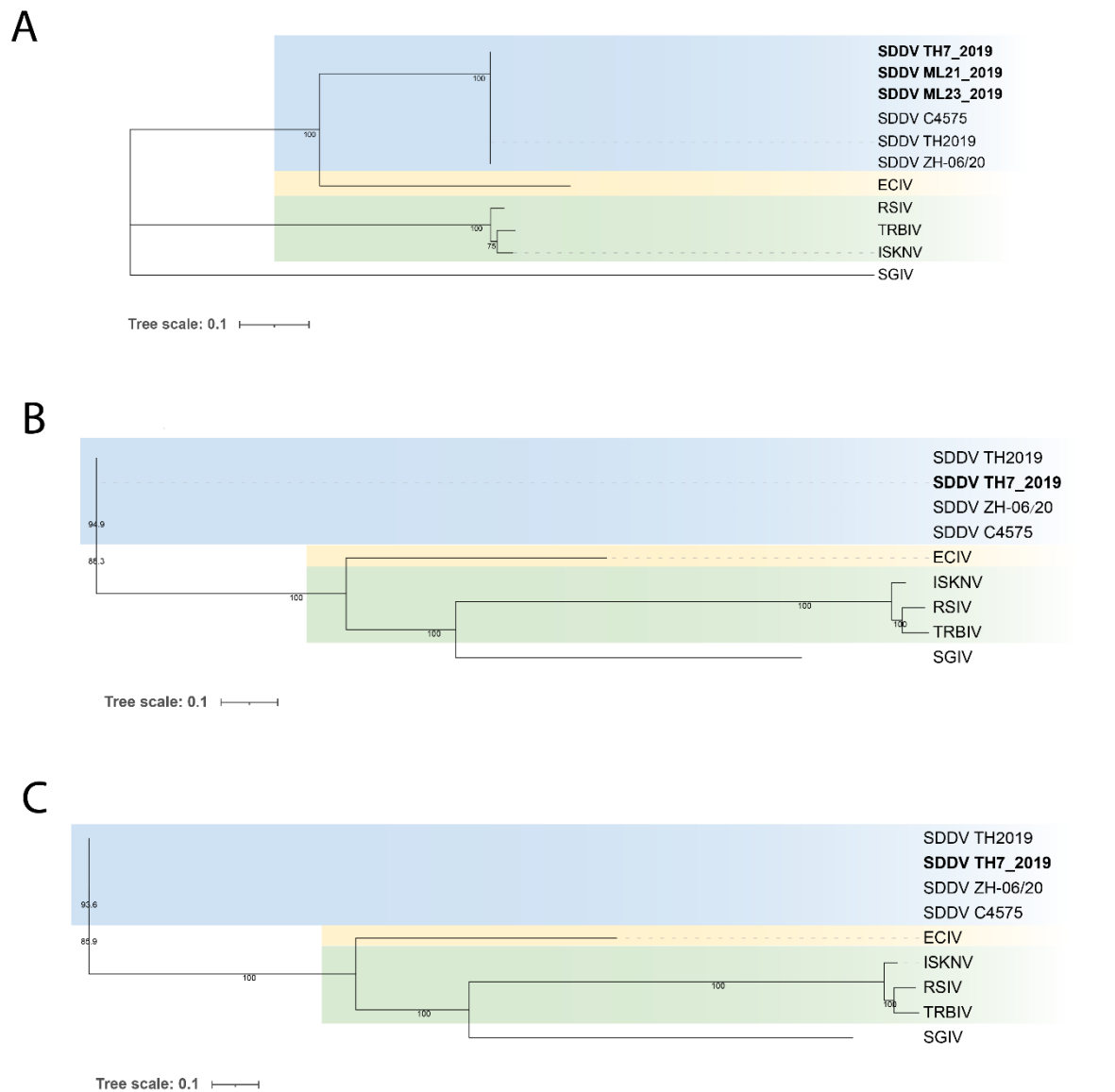
Pan-genome analysis identified a total of 148 putative protein-coding sequences among SDDV genomes including 114 core genes (present in at least 3 of 4 genomes) and 34 accessory genes (present in less than 3 genomes) (Figure 5). Predicted functions of the core genes are primarily involved with essential biological functions such as DNA replication, transcription, translation, structural and cell metabolisms. Based on presence and absence of genes, SDDV Thai strains (SDDV TH2019 and TH7\_2019) were grouped into the same cluster whereas SDDV strains from Singapore (SDDV C4575) and China (SDDV ZH-06/20) were clustered separately. Compared to the other strains, SDDV C4575 possessed a unique pattern with ten and 16 unique presence and absence genes, respectively. The function of the unique accessory genes of SDDV C4575 are largely unknown, although some were predicted to encode for mRNA capping enzyme and signaling peptide. Similarly, the majority of accessory genes that were absent in SDDV C4575 were predicted to have unknown functions, however some were predicted to contain signaling peptide and/or zinc ring finger motifs (see Supplemental Table S6).



**Figure 5.** Pan-genome analysis of SDDV genomes. (A) Pie chart showing the number of core and accessory genes. (B) Matrix showing the distribution of core genes and accessory genes. Blue and pink indicate gene presence and absence, respectively. The tree represents the relationships based on gene presence/absence content. The x-axis shows the number of genes.

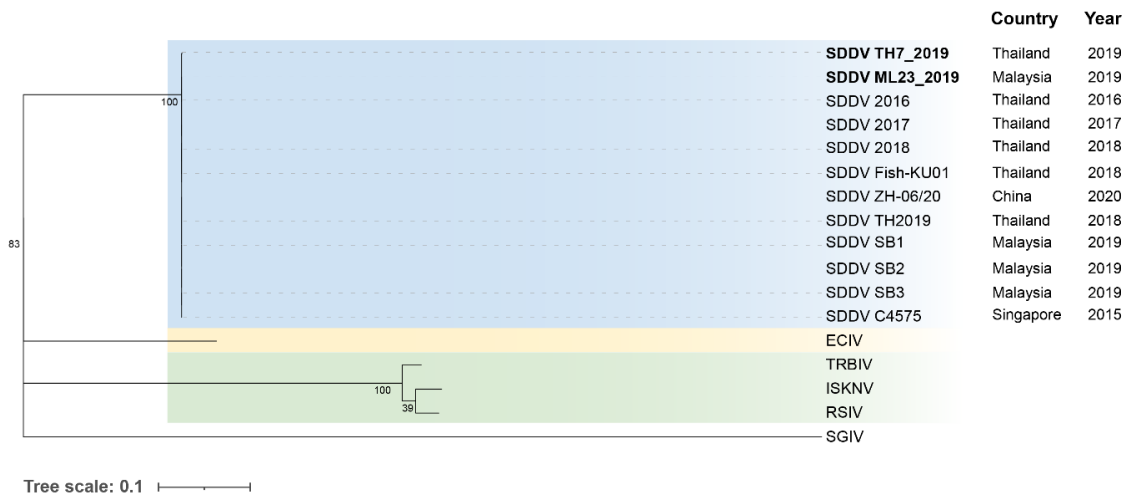
Using SDDV TH2019 as reference, SNPs were identified across seven core genes (D5 family NTPase ATPase, flap endonuclease, ATPase, myristylated membrane protein, NIF-NLI interacting factor, Uvr/REP helicase, and MCP) (Table S7). A total of 13 SNPs were identified and characterized, including five synonymous (not changing the amino acid sequence) and eight nonsynonymous (changing the amino acid sequence) variants. All nonsynonymous variants were missense variants. The SDDV ZH-06/20 exhibited the highest number of SNPs (7 SNPs), while none were detected in the SDDV TH7\_2019. SDDV strains from Malaysia possessed unique SNPs within the NTPase, flap endonuclease, and NIF-NLI interacting factor genes. Additionally, variations of SNPs were found within the Uvr/REP helicase gene among the Thai strains.

Phylogenetic analysis based on WGS, SNPs, concatenated and individual core genes indicated that all SDDV strains clustered within a single clade, distinct from other iridoviruses (Figures 6 and 7 and S2-5). Notably, SDDV exhibited the closest relationships with ECIV followed by the ISKNV clade (ISKNV, RSIV, and TRBIV).



**Figure 6.** Maximum-likelihood tree based on (A) concatenated 6 core genes. The tree was constructed using MEGA X with General Time Reversible (GTR) + G nucleotide substitution model, (B) whole genome sequences. The tree was constructed using IQ TREE with GTR + F + G4 nucleotide substitution model, (C) single nucleotide polymorphisms (SNPs). The tree was constructed using IQ TREE with Transversion model (TVM) + F +ASC + G4 nucleotide substitution model. All trees were constructed with 1,000 replications and bootstrap support values are shown at the nodes. Names in bold represent SDDV strains from this study. Scale bar represents nucleotide substitution per site.

Singapore grouper iridovirus (SGIV), belonging to the genus *Ranavirus*, was used as an outgroup. SDDV: scale drop disease virus, ECIV: European chub iridovirus, ISKNV: infectious spleen and kidney necrosis virus, RSIV: red sea bream iridovirus, TRBIV: turbot reddish body iridovirus.



**Figure 7.** Maximum-likelihood tree based on major capsid protein (MCP) gene constructed using MEGA X software with Kimura 2-parameter (K2) + I nucleotide substitution model and 1,000 replications. Scale bar represents nucleotide substitution per site. Names in bold represent SDDV strains from this study. SGIV, belonging to the genus *Ranavirus*, was used as an outgroup. Bootstrap support values in percentage are shown at the tree node. SDDV: scale drop disease virus, ECIV: European chub iridovirus, ISKNV: infectious spleen and kidney necrosis virus, RSIV: red sea bream iridovirus, TRBIV: turbot reddish body iridovirus.

4. Discussion

Scale drop syndrome (SDS) has been a significant threat to Asian seabass aquaculture by causing severe economic losses across Southeast Asia [4,7,8]. Affected fish showed lethargy, darkened bodies, gill pallor, hemorrhage, fin and tail rot, scale loss, and spleen and kidney enlargement. Other pathogens, such as *V. harveyi* [5], *T. maritimum* [6], ISKNV [38], *L. calcarifer* herpes virus (LCHV) [39,40], and RSIV [41,42], can also cause similar symptoms or occasionally co-infected with SDDV. In this study, metagenomics analysis was used to study viral and bacterial communities within the tissues of diseased Asian seabass exhibiting pathological characteristics of SDS. Taxonomic analysis of bacteria identified members of the pathogenic bacterial families *Staphylococcaceae*, *Enterobacteriaceae*, *Vibrionaceae*, *Flavobacteriaceae*, *Pseudomonadaceae*, *Morganellaceae*, *Hafniaceae*, *Mycobacteriaceae*, and *Aeromonadaceae*. Among these, members of the *Vibrionaceae* family were suspected to be the cause of SDS observed in our samples, as its relative abundance was notably high. However, it is worth noting that only two samples exhibited high abundance of the *Vibrionaceae* family. Most of the viral sequences in the samples represented members of the *Iridoviridae* Family, predominantly SDDV. Based on the abundance of viral and bacterial reads, our findings suggest that SDDV and not *Vibrio* spp. was the primary pathogen responsible for causing SDS, while *Vibrio* spp. may also act opportunistically to increase disease severity.

In addition to these pathogens, sequences representing other potential pathogenic viruses and bacteria were discovered, raising concerns about opportunistic infections and disease transmission. For example, the *Herpesviridae* family was frequently identified as the second-ranked pathogen in many samples. This virus was reported to co-infect subclinically with SDDV in Asian seabass and may currently be endemic in this species [43]. For sample 3, which displayed the most distinctive bacterial and viral profiles, we observed a high abundance of the family *Pseudomonadaceae* and *Mycobacteriaceae*, corresponding with a high abundance of bacteriophage family *Peduoviridae*. There is an increasing number of studies that demonstrated the role of the pathobiome in disease [44] and



further work is needed to investigate the role of SDDV, *Vibrio* spp. and potentially other species in SDS.

Our study successfully employed metagenomics to recover a complete SDDV genome directly from the tissue samples, but this approach does have limitations. The detection of viral sequences can be challenging due to the overwhelming presence of host sequences and other microbial species. In our study, sample 7, which had a relatively high number of iridovirus reads, enabled de novo assembly of a single SDDV-like contig of 131 kb in length. This contig was confirmed as SDDV through reference genome mapping, comparative genomics, and phylogenetic analyses. However, samples with fewer iridovirus reads yielded only partial SDDV genomes. The completeness of viral genomes assembled via metagenomics depends on the abundance of viral reads in the samples, which could be influenced by the severity or stage of the viral infection. Therefore, a high sequencing depth is recommended for metagenomics analysis. Interestingly, sample 7 was obtained from fin tissue, which is generally considered to be a lower priority target organ for SDDV during the early stages of infection [43]. This aligns with a recent study by Charoenwai et al. [17] which also detected SDDV in non-destructive samples like mucus and fin clips.

The SDDV TH7\_2019 genome was found to be 131,759 bp in length, with a GC content of 36.6%. This genome length is characteristic of MCV genomes, which are typically larger compared to other viruses in the *Iridoviridae* family. Recent research has categorized the *Megalocytyivirus* genus into two main clusters, with one cluster, containing viruses like ISKNV, RSIV, and TRBIV, known as the ISKNV-like cluster. The other cluster is more distinct and includes SDDV and ECIV. This distinct grouping has led to the proposal of a new cluster named the SDDV-like cluster [9,45,46]. Our genome characterization and phylogenetic analysis support the proposal of an SDDV-like cluster, based on both GC content and genome length. The SDDV genome sequenced in this study, along with previously characterized SDDV and ECIV genomes, share a relatively low GC content (< 40%, ranging from 36.5% to 37%) and a relatively larger size (> 128 kb, ranging from 128 kbp to 131 kbp). These distinct features suggest that these viruses form a novel clade within the genus *Megalocytyiviruses*, separate from other known MCVs [4,12,47,48]. Analysis of nucleotide and amino acid sequence similarities between SDDV genomes revealed a high degree of similarity (> 99%) within the strains, even across different locations and years.

Interestingly, the first characterized SDDV genome (SDDV C4575) from Singapore has the shortest genome size (124 kb) compared to other SDDV genomes, which typically range around 131 kb [4,9,12]. A recent comparative genomics study comparing Thai SDDV strains TH2019 and SDDV C4575, revealed a 7.6-kb-long unique region encoding for ORFs 15-20 with unknown functions in SDDV TH2019 [12]. However, our study identified the presence of this region in SDDV SG12\_2019, which originated from Singapore. The missing region in SDDV C4575 could potentially be either the variations within Singapore strains or a consequence of the limitations of the sequencing technique, VIDISCA-454 (virus discovery cDNA-AFLP combined with Roche 454). This method is often challenged by high interference from background sequences and limited availability of reference genomes, which may have resulted in incomplete sequencing of certain genomic regions [49,50]. Pan-genome analysis revealed a highly conserved genome structure shared by all SDDV strains with the majority of genes (70%) defined as SDDV core genes. It is a common occurrence for core genes to be primarily associated with genome replication, transcription, and modification, as they are widely recognized as essential genes across many viral species. Meanwhile, variations were observed among the accessory genes. SDDV C4575 exhibited the most distinct gene presence/absence pattern. Even though the functions of the majority of these genes are unknown, some are potentially linked to enhancing mRNA stability during the translation process. The absence of the 7.6 kb region, which contains 6 ORFs, in SDDV C4575 would have influenced the results of the pan-genome analysis. Further functional characterization of the accessory genes could elucidate potential differences contributing to virulence or adaptation ability among SDDV strains.

Gene prediction revealed 134 putative ORFs in the SDDV TH7\_2019 genome, with known functions attributed to some genes. Out of these, 26 genes were identified as iridovirus core genes, showing high homology to SDDV and other viruses within the family. Notably, among these core

genes, MCP (ORF079) is a structural component of the virus particles, constituting 40-50% of the viral particle [51]. The other iridovirus core genes are mostly associated with DNA replication, transcription, cell metabolism. The genes described as essential for the viral life cycle are DNA polymerase (ORF008 and ORF099), DNA repair protein (ORF096), D5 family NTPase (ORF062), DNA binding/packing protein (ORF039), and helicase (ORF124) [51–53]. In addition to the essential genes, several genes associated with virulence and host immune interaction of SDDV were also identified. Some of these genes are related to host apoptosis manipulation including tumor necrosis receptor (TNFR) homologs (ORF009 and ORF069), and Golgi antiapoptotic protein (ORF117). *Iridovirus*, like many other viral families including *Poxviridae*, possess genes that encode proteins capable of suppressing host apoptosis. Apoptosis is a natural cellular self-destruct mechanism that eliminates virus-infected cells. By inhibiting apoptosis, these viral genes prolong the survival of infected cells, creating a more favorable environment for viral replication [54]. ORF072 was identified as a gene encoding the small subunit of ribonucleotide reductase. This protein, previously studied in poxviruses, functions by binding to the host ribonucleotide reductase large subunit. This interaction induces host ribonucleotide reduction, thereby facilitating the viral replication process [55]. TNFR homologs or TNFR-associated protein genes are commonly found in other fish iridoviruses, whereas gene loss events have been reported in many non-fish iridoviruses, such as those affecting amphibians and reptiles [56]. These genes may have significantly contributed to the adaptation to different natural host species during iridovirus-host co-evolution [56]. Furthermore, the SDDV genome contains six ORFs encoding ankyrin repeat-containing proteins, which may encode repressors of the host immune response [57]. Collectively, the SDDV TH7\_2019 genome consists of genes encoding host immune evasion functions. These genes potentially contribute to prolonged SDDV infection by inhibiting apoptosis and triggering inflammatory responses during infection. This aligns with observations of the delayed appearance of clinical signs of SDDV infection and host responses characterized by the release of chemokines, interleukins, and tumor necrosis factors [43].

Phylogenetic analyses utilizing various approaches, including WGS, SNPs, and core genes, consistently demonstrated a close relationship between SDDV strains, grouping them within a single clade. The high degree of conservation observed across the SDDV strains provides flexibility in the use of available resources for SDDV identification and classification. This observation supports the potential of utilizing the MCP and core genes, such as ATPase, which have been identified in several studies, for the development of diagnostic tools and vaccines [11,57–59].

Variations within SDDV strains were highlighted by determination of tandem repeats and SNPs. Tandem repeats are short lengths of DNA that are repeated multiple times contributes by DNA slippage during replication process. Variation in numbers of tandem repeats referred as a variable number tandem repeats (VNTRs) facilitates studies of genetic diversity and evolution. Several previous studies have demonstrated the potential of using repeat regions in epidemiological studies of viruses, such as white spot syndrome virus (WSSV) and African swine fever virus (ASFV), to trace the origin and virus distribution [60–62]. In these viruses, variations in repeat sequences among populations enable the development of genetic markers for strain discrimination [61,63]. Repeat sequences were previously reported in genes encoded for myristoylated membrane, hypothetical protein, ADP-ribose glycohydrolase and putative ankyrin repeat protein through genome comparison between SDDV C4575 and TH2019 [12]. Similarly, our study identified the repeat sequences across these genes with additional SDDV genomes. A significantly different number of repeats were found in gene encoding putative membrane and hypothetical protein (ORF077 and ORF055) between Thai and other SDDV strains from Singapore and China. These genes could potentially serve as genetic markers for further research on genetic diversity in SDDV.

Additionally, determination of SNPs within the core genes indicated variations in SNPs patterns across SDDV strains from different geographical origins. Notably, the SDDV strains from Malaysia displayed unique missense mutations in genes encoding the NTPase, flap endonuclease, and NIF-NLI interacting factor. Missense mutations can alter the amino acid sequence, potentially affecting protein function. While the SDDV strains from China harbored the highest number of SNPs, most were silent mutations, potentially having a less significant impact compared to the missense

mutations observed in Malaysia. Variations in SNPs were detected within the Uvr/REP helicase gene of Thai SDDV isolates collected across different years. Previous reports have also highlighted differences in amino acid substitutions within genes encoding ATPase and myristylated membrane proteins between SDDV isolates from 2019 and 2016 [12]. This suggests a potential for temporal accumulation of mutations, leading to increased strain divergence over time. Viral mutations, particularly missense mutations, can be a mechanism for adaptation and survival under selective pressure [64]. Future studies should investigate whether these observed amino acid substitutions impact the structural integrity or functionality of SDDV proteins, potentially influencing viral fitness or virulence, as observed in other viruses [65]. Nevertheless, the relatively small sample size and limited availability of SDDV genomes from diverse geographical regions of this study restrict the generalizability of our findings. Future research should prioritize expanding the SDDV strain collection by including more samples from a wider range of countries. This will provide a more comprehensive understanding of SDDV genetic diversity and potential geographical trends in mutation patterns.

Altogether, SDDV isolated from diseased Asian seabass from across Southeast Asian countries exhibited high similarity in their genomes and several genes still hold promise as targets for diagnostic approaches or vaccine development. However, differences between SDDV strains were also observed and these can likely be attributed to differences in both host and geographic location. This raises a concern related to the translocation of seabass stocks across countries, specifically the movement of fingerlings from hatcheries in one country to grow-out facilities in another. Such translocations could potentially introduce different viral strains into new environments, each with its own level of virulence and adaptability. As a result, the efficacy of existing diagnostic tools and vaccines in the affected countries could decrease over time. This highlights the need for ongoing updates in disease surveillance, as well as the implementation of effective control measures and virus prevention strategies.

## 5. Conclusions

This study employed metagenomic techniques to investigate the microbial diversity associated with scale drop disease (SDS) and successfully recovered a complete genome of the causative agent, scale drop disease virus (SDDV). We further characterized the SDDV genome through phylogenetic analysis and functional annotation, revealing insights into genome structure, evolutionary relationships, and potential virulence factors. These findings provide a foundation for developing targeted diagnostic tools and effective disease control strategies for SDDV.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: Dot plot of whole genome sequence alignment between the reference genome (SDDV TH 2019) and (A) SDDV TH2019, (B) C4575, and (C) ZH-06/20; Figure S2: Maximum-likelihood tree based on (A) DNA-dependent RNA polymerase II alpha subunit (RPO), (B) DNA-dependent RNA polymerase II beta subunit (RPO2), and (C) NTPase gene; Figure S3: Maximum-likelihood tree based on (A) ATPase, (B) myristylated membrane protein, and (C) NIF-NLI interacting factor gene; Figure S4: Maximum-likelihood tree based on (A) D5 family NTPase ATPase and (B) DNA polymerase gene; Figure S5: Maximum-likelihood tree based on (A) Uvr/REP helicase and (B) flap endonuclease gene; Table S1: Information of gene sequences of SDDV used in this study; Table S2: BUSCO assessment results of recovered SDDV genomes in this study; Table S3: Genome annotation of the SDDV TH7\_2019; Table S4: Gene-to-gene comparison of ORFs of the SDDV TH7\_2019 against previously published SDDV genomes; Table S5: Tandem repeats in SDDV genomes; Table S6: Gene presence and absence of SDDV genomes; Table S7: Summary of single nucleotide polymorphisms (SNPs) in SDDV core genes.

**Author Contributions:** Conceptualization, S.S. and C.R.; methodology, P.C.; validation, D.B. and R.v.A.; formal analysis, P.C.; investigation, S.S.; resources, J.A.D., M.N.A.A., P.P.; data curation, P.C.; writing—original draft preparation, P.C.; visualization, P.C.; project administration, S.S. and C.R.; funding acquisition, S.S. and C.R.; writing—review and editing, J.A.D., M.N.A.A., D.B., R.v.A., S.S., and C.R. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study due to our study only based on metagenomic data retrieved from the previous studies and does not involve any animal

experimentation, biological sampling, or direct involvement of human subjects. As such, it falls outside the scope of requirements for Animal Ethics or Institutional Review Board approval.

**Funding:** Putita Chokmangmeepisarn was supported by Chulalongkorn University – NSTDA for Doctoral Scholarship. This project was funded by the National Science, Research and Innovation Fund, Thailand Science Research and innovation (TSRI), and the 90<sup>th</sup> anniversary of Chulalongkorn University scholarship, the budget for the International Researcher Conference by Ratchadapisek Sompoch Fund 2024, the National Research Council of Thailand (NRCT) under Mid-Career Talented Researchers Grant no. NRCT5-RSA63001-01, NRCT: R. Thanawongnuwech NRCT Senior scholar 2022 #N42A650553 and Thailand Science Research and Innovation Fund Chulalongkorn University (FF67\_4709668).

**Data Availability Statement:** All sequence data presented in this study were submitted to the National Center for Biotechnology Information (NCBI) database. The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

**Acknowledgments:** We would like to thanks the Aquatic Animal Health and Therapeutics Laboratory, Institute of Bioscience, Universiti Putra Malaysia and Tropical Futures Institute, James Cook University, Singapore for providing the tissue samples. Metagenomics data was provided by the Fish Health Platform, Center of Excellence for Shrimp Molecular Biology and Biotechnology (Centex Shrimp), Mahidol University, Thailand.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Yue, G.H.; Zhu, Z.Y.; Lo, L.C.; Wang, C.M.; Lin, G.; Feng, F.; Pang, H.Y.; Li, J.; Gong, P.; Liu, H.M.; Tan, J.; Chou, R.; Lim, H.; Orban, L. Genetic variation and population structure of Asian seabass (*Lates calcarifer*) in the Asia-Pacific region, *Aquaculture*. **2009**, *1*, 22-28. <https://doi.org/10.1016/j.aquaculture.2009.03.053>.
2. Gibson-Kueh, S.; Chee, D.; Chen, J.; Wang, Y.H.; Tay, S.; Leong, L.N.; Ng, M.L.; Jones, J.B.; Nicholls, P.K.; Ferguson, H.W. The pathology of 'scale drop syndrome' in Asian seabass, *Lates calcarifer* Bloch, a first description. *J. Fish. Dis.* **2012**, *35*, 19-27. <https://doi.org/10.1111/j.1365-2761.2011.01319.x>.
3. Gibson-Kueh, S. Diseases of Asian seabass (or barramundi), *Lates calcarifer* Bloch, Murdoch University, 2012, pp. 162.
4. de Groof, A.; Guelen, L.; Deijs, M.; van der Wal, Y.; Miyata, M.; Ng, K.S.; van Grinsven, L.; Simmelink, B.; Biermann, Y.; Grisez, L.; van Lent, J.; de Ronde, A.; Chang, S.F.; Schrier, C.; van der Hoek, L. A Novel Virus Causes Scale Drop Disease in *Lates calcarifer*. *PLoS Pathog.* **2015**, *11*, e1005074. <https://doi.org/10.1371/journal.ppat.1005074>.
5. Kayansamruaj, P.; Dong, H.T.; Hirono, I.; Kondo, H.; Senapin, S.; Rodkhum, C. Genome characterization of piscine 'Scale drop and Muscle Necrosis syndrome'-associated strain of *Vibrio harveyi* focusing on bacterial virulence determinants. *J. Appl. Microbiol.* **2018**, *124*, 652-666. <https://doi.org/10.1111/jam.13676>.
6. Mabrok, M.; Algammal, A.M.; Sivaramasamy, E.; Hetta, H.F.; Atwah, B.; Alghamdi, S.; Fawzy, A.; Avendano-Herrera, R.; Rodkhum, C. Tenacibaculosis caused by *Tenacibaculum maritimum*: Updated knowledge of this marine bacterial fish pathogen. *Front. Cell Infect. Microbiol.* **2022**, *12*, 1068000. <https://doi.org/10.3389/fcimb.2022.1068000>.
7. Nurliyana, M.; Lukman, B.; Ina-Salwany, M.Y.; Zamri-Saad, M.; Annas, S.; Dong, H.T.; Rodkhum, C.; Amal, M.N.A. First evidence of scale drop disease virus in farmed Asian seabass (*Lates calcarifer*) in Malaysia. *Aquaculture*. **2020**, *528*, 735600. <https://doi.org/10.1016/j.aquaculture.2020.735600>.
8. Senapin, S.; Dong, H.; Meemetta, W.; Gangnonngiw, W.; Sangsuriya, P.; Vanichviriyakit, R.; Sonthi, M.; Nuangsaeng, B. Mortality from scale drop disease in farmed *Lates calcarifer* in Southeast Asia. *J. Fish. Dis.* **2019**, *42*, 119-127. <https://doi.org/10.1111/jfd.12915>.
9. Fu, Y.; Li, Y.; Fu, W.; Su, H.; Zhang, L.; Huang, C.; Weng, S.; Yu, F.; He, J.; Dong, C. Scale Drop Disease Virus Associated Yellowfin Seabream (*Acanthopagrus latus*) Ascites Diseases, Zhuhai, Guangdong, Southern China: The First Description. *Viruses*. **2021**, *13*, 8. <https://doi.org/10.3390/v13081617>.
10. Fu, X.; Li, N.; Liu, L.; Lin, Q.; Wang, F.; Lai, Y.; Jiang, H.; Pan, H.; Shi, C.; Wu, S. Genotype and host range analysis of infectious spleen and kidney necrosis virus (ISKNV). *Virus Genes*. **2011**, *42*, 97-109. <https://doi.org/10.1007/s11262-010-0552-x>.
11. Fusianto, C.K.; Becker, J.A.; Subramaniam, K.; Whittington, R.J.; Koda, S.A.; Waltzek, T.B.; Murwantoko; Hick, P.M.; Zheng, C. Genotypic Characterization of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Southeast Asian Aquaculture. *Transb. Emerg. Dis.* **2023**, *2023*, 1-16. <https://doi.org/10.1155/2023/6643006>.



12. Kayansamruaj, P.; Soontara, C.; Dong, H.T.; Phiwsaiya, K.; Senapin, S. Draft genome sequence of scale drop disease virus (SDDV) retrieved from metagenomic investigation of infected barramundi, *Lates calcarifer* (Bloch, 1790). *J. Fish. Dis.* **2020**, *43*, 1287-1298. <https://doi.org/10.1111/jfd.13240>.
13. Mokili, J.L.; Rohwer, F.; Dutilh, B.E. Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol.* **2012**, *2*, 63-77. <https://doi.org/10.1016/j.coviro.2011.12.004>.
14. Munang'andu, H.M. Environmental Viral Metagenomics Analyses in Aquaculture: Applications in Epidemiology and Disease Control. *Front. Microbiol.* **2016**, *7*, 1986. <https://doi.org/10.3389/fmicb.2016.01986>.
15. Rigou, S.; Santini, S.; Abergel, C.; Claverie, J.M.; Legendre, M. Past and present giant viruses diversity explored through permafrost metagenomics. *Nat. Commun.* **2022**, *13*, 5853. <https://doi.org/10.1038/s41467-022-33633-x>.
16. Yang, J.; Yang, F.; Ren, L.; Xiong, Z.; Wu, Z.; Dong, J.; Sun, L.; Zhang, T.; Hu, Y.; Du, J.; Wang, J.; Jin, Q. Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach. *J. Clin. Microbiol.* **2011**, *49*, 3463-3469. <https://doi.org/10.1128/JCM.00273-11>.
17. Charoenwai, O.; Senapin, S.; Dong, H.T.; Sonthi, M. Detection of scale drop disease virus from non-destructive samples and ectoparasites of Asian sea bass, *Lates calcarifer*. *J. Fish. Dis.* **2020**. <https://doi.org/10.1111/jfd.13290>.
18. Charoenwai, O.; Meemetta, W.; Sonthi, M.; Dong, H.T.; Senapin, S. A validated semi-nested PCR for rapid detection of scale drop disease virus (SDDV) in Asian sea bass (*Lates calcarifer*). *J. Virol. Methods.* **2019**, *268*, 37-41. <https://doi.org/10.1016/j.jviromet.2019.03.007>.
19. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* **2014**, *30*, 15, 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>.
20. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **2012**, *9*, 357-359. <https://doi.org/10.1038/nmeth.1923>.
21. Wood, D.E.; Lu, J.; Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome. Biol.* **2019**, *20*, 257. <https://doi.org/10.1186/s13059-019-1891-0>.
22. Lu, J.; Breitwieser, F.P.; Thielen, P.; Salzberg, S.L. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **2017**, *3*, e104. <https://doi.org/10.7717/peerj-cs.104>.
23. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2009.
24. Rstudio Team. RStudio: Integrated Development Environment for R. 2015. <http://www.rstudio.com/>.
25. Li, D.; Luo, R.; Liu, C.M.; Leung, C.M.; Ting, H.F.; Sadakane, K.; Yamashita, H.; Lam, T.W. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods.* **2016**, *102*, 3-11. <https://doi.org/10.1016/j.ymeth.2016.02.020>.
26. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: architecture and applications. *BMC Bioinformatics.* **2009**, *10*, 421. <https://doi.org/10.1186/1471-2105-10-421>.
27. Manni, M.; Berkeley, M.R.; Seppey, M.; Simão, F.A.; Zdobnov, E.M. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* **2021**, *38*, 4647-4654. <https://doi.org/10.1093/molbev/msab199>.
28. Hyatt, D.; Chen, G.L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics.* **2010**, *11*, 119. <https://doi.org/10.1186/1471-2105-11-119>.
29. Darling, A.C.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome. Res.* **2004**, *14*, 1394-1403. <https://doi.org/10.1101/gr.2289704>.
30. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* **2019**, *20*, 1160-1166. <https://doi.org/10.1093/bib/bbx108>.
31. Shimoyama, Y. pyGenomeViz: A genome visualization python package for comparative genomics. 2022. <https://github.com/moshi4/pyGenomeViz>.
32. Trifinopoulos, J.; Nguyen, L.T.; von Haeseler, A.; Minh, B.Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **2016**, *44*, W232-235. <https://doi.org/10.1093/nar/gkw256>.
33. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermini, L.S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods.* **2017**, *14*, 587-589. <https://doi.org/10.1038/nmeth.4285>.



34. Kaas, R.S.; Leekitcharoenphon, P.; Aarestrup, F.M.; Lund, O. Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms. *PLOS ONE*. **2014**, *9*, e104984. <https://doi.org/10.1371/journal.pone.0104984>.
35. Kumar, S.; Stecher, G.; Li, M.; Nnyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547-1549. <https://doi.org/10.1093/molbev/msy096>.
36. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573-580. <https://doi.org/10.1093/nar/27.2.573>.
37. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.G.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. **2015**, *31*, 3691-3693. <https://doi.org/10.1093/bioinformatics/btv421>.
38. Kerddee, P.; Dong, H.T.; Chokmangmeepisarn, P.; Rodkhum, C.; Srisapoome, P.; Areechon, N.; Del-Pozo, J.; Kayansamruaj, P. Simultaneous detection of scale drop disease virus and *Flavobacterium columnare* from diseased freshwater-reared barramundi *Lates calcarifer*. *Dis. Aquat. Organ.* **2020**, *140*, 119-128. <https://doi.org/10.3354/dao03500>.
39. Chang, S.F.; Ng, K.S.; Grisez, L.; De Groof, A.; Vogels, W.; Van der Hoek, L.; Deijs, M. Novel Fish Pathogenic Virus, in: W.I.P.O.I. Bureau (Ed.) 2018.
40. Dang, M.; Dien, T.D.; Van, U.P.; Ha, V.T.; Dung, V.V.; Hieu, N.T.D.; Hua, V.C.; Hue, N.T.K.; Giang, N.T.T.; Truong, V.H.; Dong, H.T. The first description of histopathology of *Lates calcarifer* herpesvirus (LCHV) infection in barramundi (*Lates calcarifer*). *Aquaculture*. **2023**, *565*, 739091. <https://doi.org/10.1016/j.aquaculture.2022.739091>.
41. Girisha, S.K.; Puneeth, T.G.; Nithin, M.S.; Naveen Kumar, B.T.; Ajay, S.K.; Vinay, T.N.; Suresh, T.; Venugopal, M.N.; Ramesh, K.S. Red sea bream iridovirus disease (RSIVD) outbreak in Asian seabass (*Lates calcarifer*) cultured in open estuarine cages along the west coast of India: First report. *Aquaculture*. **2020**, *520*, 734712. <https://doi.org/10.1016/j.aquaculture.2019.734712>.
42. Sumithra, T.G.; Krupesha Sharma, S.R.; Neelima, L.; Dhanutha, N.R.; Joshy, A.; Anusree, V.N.; Gayathri, S.; Raghu, R.K.; Praveen, N.D.; Thomas, S.; Rajesh, K.M. Red sea bream iridovirus infection in cage farmed Asian sea bass (*Lates calcarifer*): Insights into the pathology, epizootiology, and genetic diversity. *Aquaculture*. **2022**, *548*, 737571. <https://doi.org/10.1016/j.aquaculture.2021.737571>.
43. Domingos, J.A.; Shen, X.; Terence, C.; Senapin, S.; Dong, H.T.; Tan, M.R.; Gibson-Kueh, S.; Jerry, D.R. Scale Drop Disease Virus (SDDV) and *Lates calcarifer* Herpes Virus (LCHV) Coinfection Downregulate Immune-Relevant Pathways and Cause Splenic and Kidney Necrosis in Barramundi Under Commercial Farming Conditions. *Front. Genet.* **2021**, *12*, 666897. <https://doi.org/10.3389/fgene.2021.666897>.
44. Bass, D.; Stentiford, G.D.; Wang, H.C.; Koskella, B.; Tyler, C.R. The Pathobiome in Animal and Plant Diseases. *Trends Ecol. Evol.* **2019**, *34*, 996-1008. <https://doi.org/10.1016/j.tree.2019.07.012>.
45. Kim, W.S.; Oh, M.J.; Kim, J.O.; Kim, D.; Jeon, C.H.; Kim, J.H. Detection of megalocytivirus from imported tropical ornamental fish, paradise fish *Macropodus opercularis*. *Dis. Aquat. Organ.* **2010**, *90*, 235-239. <https://doi.org/10.3354/dao02232>.
46. Kurita, J.; Nakajima, K. *Megalocytiviruses*. *Viruses*. **2012**, *4*, 521-538. <https://doi.org/10.3390/v4040521>.
47. Fu, Y.; Li, Y.; Zhang, W.; Fu, W.; Li, W.; Zhu, Z.; Weng, S.; He, J.; Dong, C. Effectively protecting Asian seabass *Lates calcarifer* from ISKNV-I, ISKNV-II, RSIV-II and SDDV by an inactivated ISKNV-I and SDDV bivalent vaccine. *Aquaculture*. **2023**, *566*, 739218. <https://doi.org/10.1016/j.aquaculture.2022.739218>.
48. Halaly, M.A.; Subramaniam, K.; Koda, S.A.; Popov, V.L.; Stone, D.; Way, K.; Waltzek, T.B. Characterization of a Novel Megalocytivirus Isolated from European Chub (*Squalius cephalus*). *Viruses*. **2019**, *11*, 5. <https://doi.org/10.3390/v11050440>.
49. de Vries, M.; Deijs, M.; Canuti, M.; van Schaik, B.D.; Faria, N.R.; van de Garde, M.D.; Jachimowski, L.C.; Jebbink, M.F.; Jakobs, M.; Luyf, A.C.; Coenjaerts, F.E.; Claas, E.C.; Molenkamp, R.; Koekkoek, S.M.; Lammens, C.; Leus, F.; Goossens, H.; Ieven, M.; Baas, F.; van der Hoek, L. A sensitive assay for virus discovery in respiratory clinical samples. *PLoS One*. **2011**, *6*, e16118. <https://doi.org/10.1371/journal.pone.0016118>.
50. De Vries, M.; Oude Munnink, B.B.; Deijs, M.; Canuti, M.; Koekkoek, S.M.; Molenkamp, R.; Bakker, M.; Jurriaans, S.; Van Schaik, B.D.C.; Luyf, A.C.; Olabarriaga, S.D.; Van Kampen, A.H.C.; Van der Hoek, L. Performance of VIDISCA-454 in feces-suspensions and serum. *Viruses*. **2012**, *4*, 1328-1334. <https://doi.org/10.3390/v4081328>.

51. He, J.G.; Deng, M.; Weng, S.P.; Li, Z.; Zhou, S.Y.; Long, Q.X.; Wang, X.Z.; Chan, S.-M. Complete Genome Analysis of the Mandarin Fish Infectious Spleen and Kidney Necrosis Iridovirus. *Virology*. **2001**, *291*, 126-139. <https://doi.org/10.1006/viro.2001.1208>.
52. Gray, M.J.; Chinchar, V.G. *Ranaviruses*. Springer Cham. 2015.
53. Song, W.J.; Qin, Q.W.; Qiu, J.; Huang, C.H.; Wang, F.; Hew, C.L. Functional genomics analysis of Singapore grouper iridovirus: complete sequence determination and proteomic analysis. *J. Virol.* **2004**, *78*, 12576-90. <https://doi.org/10.1128/JVI.78.22.12576-12590.2004>.
54. Huang, X.; Huang, Y.; Cai, J.; Wei, S.; Gao, R.; Qin, Q. Identification and characterization of a tumor necrosis factor receptor-like protein encoded by Singapore grouper iridovirus. *Virus Res.* **2013**, *178*, 340-348. <https://doi.org/10.1016/j.virusres.2013.09.023>.
55. Gammon, D.B.; Gowrishankar, B.; Duraffour, S.; Andrei, G.; Upton, C.; Evans, D.H. Vaccinia virus-encoded ribonucleotide reductase subunits are differentially required for replication and pathogenesis. *PLoS Pathog.* **2010**, *6*, e1000984. <https://doi.org/10.1371/journal.ppat.1000984>.
56. Huang, Y.; Huang, X.; Liu, H.; Gong, J.; Ouyang, Z.; Cui, H.; Cao, J.; Zhao, Y.; Wang, X.; Jiang, Y.; Qin, Q. Complete sequence determination of a novel reptile iridovirus isolated from soft-shelled turtle and evolutionary analysis of *Iridoviridae*. *BMC Genomics*. **2009**, *10*, 224. <https://doi.org/10.1186/1471-2164-10-224>.
57. Kurita, J.; Nakajima, K. *Megalocytiviruses*. *Viruses*. **2012**, *4*, 521-538. <https://doi.org/10.3390/v4040521>.
58. Eaton, H.E.; Ring, B.A.; Brunetti, C.R. The genomic diversity and phylogenetic relationship in the family *Iridoviridae*. *Viruses*. **2010**, *2*, 1458-75. <https://doi.org/10.3390/v2071458>.
59. Song, J.Y.; Kitamura, S.; Jung, S.J.; Miyadai, T.; Tanaka, S.; Fukuda, Y.; Kim, S.R.; Oh, M.J. Genetic variation and geographic distribution of megalocytiviruses. *J. Microbiol.* **2008**, *46*, 29-33. <https://doi.org/10.1007/s12275-007-0184-6>.
60. Syed Musthaq, S.; Sudhakaran, R.; Ishaq Ahmed, V.P.; Balasubramanian, G.; Sahul Hameed, A.S. Variability in the tandem repetitive DNA sequences of white spot syndrome virus (WSSV) genome and suitability of VP28 gene to detect different isolates of WSSV from India. *Aquaculture*. **2006**, *256*, 34-41. <https://doi.org/10.1016/j.aquaculture.2006.01.036>.
61. Wongteerasupaya, C.; Pungchai, P.; Withyachumnarnkul, B.; Boonsaeng, V.; Panyim, S.; Flegel, T.W.; Walker, P.J. High variation in repetitive DNA fragment length for white spot syndrome virus (WSSV) isolates in Thailand. *Dis. Aquat. Organ.* **2003**, *54*, 253-7. <https://doi.org/10.3354/dao054253>.
62. Zhu, Z.; Xiao, C.T.; Fan, Y.; Cai, Z.; Lu, C.; Zhang, G.; Jiang, T.; Tan, Y.; Peng, Y. Homologous recombination shapes the genetic diversity of African swine fever viruses. *Vet. Microbiol.* **2019**, *236*, 108380. <https://doi.org/10.1016/j.vetmic.2019.08.003>.
63. Goller, K.V.; Malogolovkin, A.S.; Katorkin, S.; Kolbasov, D.; Titov, I.; Höper, D.; Beer, M.; Keil, G.M.; Portugal, R.; Blome, S. Tandem repeat insertion in African swine fever virus, Russia, 2012. *Emerg. Infect. Dis.* **2015**, *21*, 731-2. <https://doi.org/10.3201/eid2104.141792>.
64. Zhou, X.; Iversen, E.S. Jr.; Parmigiani, G. Classification of Missense Mutations of Disease Genes. *J. Am. Stat. Assoc.* **2005**, *100*, 51-60. <https://doi.org/10.1198/016214504000001817>.
65. Balasco, N.; Damaggio, G.; Esposito, L.; Colonna, V.; Vitagliano, L. A comprehensive analysis of SARS-CoV-2 missense mutations indicates that all possible amino acid replacements in the viral proteins occurred within the first two-and-a-half years of the pandemic. *Int. J. Biol. Macromol.* **2024**, *266*, 131054. <https://doi.org/10.1016/j.ijbiomac.2024.131054>.

66. Shi, C.Y.; Jia, K.T.; Yang, B.; Huang, J. Complete genome sequence of a *Megalocyttivirus* (family *Iridoviridae*) associated with turbot mortality in China. *Virol J.* **2010**, *7*, 159. <https://doi.org/10.1186/1743-422X-7-159>.
67. Ito, T.; Yoshiura, Y.; Kamaishi, T.; Yoshida, K.; Nakajima, K. Prevalence of red sea bream iridovirus among organs of Japanese amberjack (*Seriola quinqueradiata*) exposed to cultured red sea bream iridovirus. *J. Gen. Virol.* **2013**, *94*, 2094-2101. <https://doi.org/10.1099/vir.0.052902-0>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.