



# Article CRISPR-Cas-Based Adaptive Immunity Mediates Phage Resistance in Periodontal Red Complex Pathogens

Pradeep Kumar Yadalam <sup>1,\*</sup>, Deepavalli Arumuganainar <sup>2</sup>, Raghavendra Vamsi Anegundi <sup>1</sup>, Deepti Shrivastava <sup>3,\*</sup>, Sultan Abdulkareem Ali Alftaikhah <sup>4</sup>, Haifa Ali Almutairi <sup>4</sup>, Muhanad Ali Alobaida <sup>5</sup>, Abdullah Ahmed Alkaberi <sup>5</sup> and Kumar Chandan Srivastava <sup>6,7</sup>

- <sup>1</sup> Department of Periodontics, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha Dental College and Hospitals, Saveetha University, Chennai 600077, India; arvamsi2009@gmail.com
- <sup>2</sup> Department of Periodontics, Ragas Dental College and Hospital, 2/102, East Coast Road, Uthandi, Chennai 600119, India; deeps.271@gmail.com
- <sup>3</sup> Periodontics Division, Preventive Dentistry Department, College of Dentistry, Jouf University, Sakaka 72345, Saudi Arabia
- <sup>4</sup> College of Dentistry, Jouf University, Sakaka 72345, Saudi Arabia; sultan.abdulkarim.alfatakh@jodent.org (S.A.A.A.); dr.haifa.ali@jodent.org (H.A.A.)
- <sup>5</sup> General Dentist, Ministry of Health, Riyadh 12613, Saudi Arabia; maalobaida@moh.gov.sa (M.A.A.); aalkaberi@moh.gov.sa (A.A.A.)
- <sup>6</sup> Oral Medicine & Maxillofacial Radiology Division, Department of Oral & Maxillofacial Surgery & Diagnostic Sciences, College of Dentistry, Jouf University, Sakaka 72345, Saudi Arabia; drkcs.omr@gmail.com
- <sup>7</sup> Department of Oral Medicine and Radiology, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai 602105, India
- \* Correspondence: pradeepkumar.sdc@saveetha.com (P.K.Y.); sdeepti20@gmail.com (D.S.)

Abstract: Periodontal diseases are polymicrobial immune-inflammatory diseases that can severely destroy tooth-supporting structures. The critical bacteria responsible for this destruction include red complex bacteria such as Porphoromonas gingivalis, Tanerella forsythia and Treponema denticola. These organisms have developed adaptive immune mechanisms against bacteriophages/viruses, plasmids and transposons through clustered regularly interspaced short palindromic repeats (CRISPR) and their associated proteins (Cas). The CRISPR-Cas system contributes to adaptive immunity, and this acquired genetic immune system of bacteria may contribute to moderating the microbiome of chronic periodontitis. The current research examined the role of the CRISPR-Cas system of red complex bacteria in the dysbiosis of oral bacteriophages in periodontitis. Whole-genome sequences of red complex bacteria were obtained and investigated for CRISPR using the CRISPR identification tool. Repeated spacer sequences were analyzed for homologous sequences in the bacteriophage genome and viromes using BLAST algorithms. The results of the BLAST spacer analysis for T. denticola spacers had a 100% score (e value with a bacillus phage), and the results for T. forsthyia and P. gingivalis had a 56% score with a pectophage and cellulophage (e value: 0.21), respectively. The machine learning model of the identified red complex CRISPR sequences predicts with area an under the curve (AUC) accuracy of 100 percent, indicating phage inhibition. These results infer that red complex bacteria could significantly inhibit viruses and phages with CRISPR immune sequences. Therefore, the role of viruses and bacteriophages in modulating sub-gingival bacterial growth in periodontitis is limited or questionable.

Keywords: CRISPR; periodontal disease; red complex bacteria; periodonitis; dysbiosis

## 1. Introduction

Periodontal diseases are polymicrobial immune–inflammatory diseases that can severely destroy the periodontal ligament and adjacent supportive alveolar bone [1,2]. They are prevalent worldwide, affecting large populations, and have become a public health concern. Dental biofilm is a forerunner in the development of periodontal disease. The sub-gingival



Citation: Yadalam, P.K.; Arumuganainar, D.; Anegundi, R.V.; Shrivastava, D.; Alftaikhah, S.A.A.; Almutairi, H.A.; Alobaida, M.A.; Alkaberi, A.A.; Srivastava, K.C. CRISPR-Cas-Based Adaptive Immunity Mediates Phage Resistance in Periodontal Red Complex Pathogens. *Microorganisms* **2023**, *11*, 2060. https://doi.org/10.3390/ microorganisms11082060

Academic Editors: Renato Fani, Eva Torres-Sangiao, Nelson C Soares and Monica Cartelle Gestal

Received: 13 June 2023 Revised: 23 July 2023 Accepted: 1 August 2023 Published: 11 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). microbiota contain more than 700 bacterial species [3,4]. However, the red complex includes *Porphyromonas gingivalis, Treponema denticola* and *Tannerella forsythia* (formerly *Bacteroides forsythus*), encompassing the most critical pathogens associated with human adult periodontal diseases [1,5]. Furthermore, the prevalence of potential periodontopathogens, including *Fusobacterium nucleatum, Prevotella species, Eikenella corrodens, Peptostreptococcus micros* and *Campylobacter rectus*, are enhanced in deep periodontal pockets [6], which leads to the spread of microorganisms to the distant site, causing cardiovascular disease, pulmonary infections, cancer initiation and promotion, pre-term low birth weight, Alzheimer's disease and Parkinson's disease [7].

Gingivitis and periodontitis destroy the surrounding structure of soft tissue and the hard tissues of teeth, leading to tooth mobility and loss. Diagnoses are based on clinical, radiographical and microbial investigations [8]. These investigations can identify and isolate specific microorganisms or whole metagenome profiles in infected oral cavities. However, isolating and identifying phages and viromes is not a regular clinical investigation. Additionally, the role of viruses and phages in periodontal disease is still debatable, as they have been proven clinically, but their exact role has not been discussed scientifically.

Virulent and temperate phages are two types of bacteriophages. They are viruses that infect bacteria and replicate using bacterial replication mechanisms. A virulent phage, or lytic phage, strictly follows a lytic cycle. In this cycle, the phage infects the bacterium, reproduces new phages using the bacterial machinery and ultimately causes the bacterial cell to lyse (break apart), releasing newly produced phages. This cycle eventually leads to the destruction of the bacterial cell. An example of a virulent phage is the T-even phage [9].

Temperate phages, conversely, can adopt either a lytic cycle similar to virulent phages or a lysogenic cycle. Instead of immediately killing the host cell in the lysogenic cycle, the phage integrates its DNA into the bacterial chromosome and becomes a prophage. The prophage DNA replicates passively along with the bacterial DNA during regular bacterial cell divisions. Under certain circumstances, such as when the bacterium is under stress, the prophage can be induced to switch to the lytic cycle [10].

Temperate phages can play a significant role in bacterial infection and in the ability of a bacterium to escape immunity and cause infection. One of the ways this occurs is through a process known as a lysogenic conversion, in which temperate phages integrate into the bacterial chromosome as prophages and can carry and express virulence factors. These virulence factors can enhance the bacterial host's ability to cause disease, increasing their pathogenic potential [2].

Another mechanism by which temperate phages can enhance bacterial virulence is through phage induction. Phage induction is a process in which a prophage is excised from the bacterial chromosome and enters the lytic cycle, facilitating cell lysis and the production and release of virulence molecules. This not only leads to the spread of the phage but also aids in the dispersion of virulence factors, posing a potential risk to human health [1].

Emerging technologies such as genome sequencing and transcriptomics have helped researchers dig deeper into the complex interplay between temperate phages and bacterial virulence. They have revealed more subtle ways that prophages can contribute to bacterial pathogenicity, such as influencing bacterial gene expression or metabolic processes. However, research in this field is ongoing, and many intricate layers remain to be explored about the relationships between temperate phages and bacterial virulence.

Recent studies have reported the diversity of viruses in the oral cavity; most viromes contain bacteriophages [11]. Bacteriophage-based therapeutics are currently under investigation for different diseases, as they bypass the problem of antibiotic resistance. In periodontitis, bacteriophage-based therapy can also be applied, as it can overcome the bacteria's multi-drug resistance and the disease's recurrence [9]. However, acquired resistance against the phage limits the therapeutic potential of bacteriophages [12]. In most bacteria and archaea CRISPR-Cas (clustered regularly interspaced short palindromic repeats), the system contributes to adaptive immunity in most bacteria and archaea via a DNA-encoded, RNA-mediated and nucleic acid targeting mechanism [13]. Different types of CRISPR-Cas

systems have been identified, each with unique characteristics. For instance, the CRISPR-Cas13a system, a type VI-A system, targets messenger RNA rather than DNA. This system effectively inhibits certain phages, protecting the bacterial cell from infection.

Similarly, the type I-C CRISPR-Cas system has shown activity in inhibiting phage antagonists, providing a certain level of immunity to the bacteria. However, this immunity was found to be limited in a study conducted on *E. lenta*, a species of bacteria, indicating that the efficacy of the CRISPR-Cas system can vary depending on specific factors, such as the type of bacteria and phage involved [14].

Although phages are the primary targets of CRISPR-Cas systems, these systems can also target other genetic elements, such as integrative conjugative elements (ICEs) [15]. It was found that more than 80% of isolates with an active CRISPR-Cas system have spacers (segments of foreign DNA stored in the bacterial genome) that target ICEs or similar elements.

Clustered regularly interspaced short palindromic repeats (CRISPR) belong to a family of DNA sequences derived from bacteriophages and are characterized by short, direct repeats separated by spacers. The CRISPR-Cas adaptive immune system, which provides immunological memory by introducing short DNA sequences from phage and other parasite DNA elements into CRISPR loci on the host genome, is present in about half of all bacteria. In contrast to the fast evolution of CRISPR loci in their natural environments, bacterial species normally develop phage resistance through phage receptor mutations or deletions [16]. CRISPR and CRISPR-associated (Cas) genes confer resistance to exogenous sequences of bacteriophages/viruses. Their recognition depends on the similarity between sequences of targeted phage DNA segments and the spacers [17]. CRISPR-Cas systems are present in about 45% of bacterial species and in 80% of archaea. Structurally, the CRISPR-Cas system consists of a group of repeats interspersed by spacers, which are short DNA stretches along with a set of Cas genes in proximity [18]. Immunity is built by gaining short stretches of interfering nucleic acids into CRISPR loci as 'spacers' [19]. These immune markers are transcribed and processed into small non-coding interfering CRISPR RNAs (crRNAs) that guide Cas proteins toward target nucleic acids for the specific cleavage of homologous sequences. A new spacer is always added to the AT-rich leader site of CRISPR, which is thought to include unique sequence features for direct spacer DNA insertion [20]. Although the search results provided do not explicitly mention the interaction between CRISPR systems and phages in the context of periodontal bacteria, it can be inferred that the CRISPR system functions as a defense mechanism against phages in these bacteria, similar to its role in other bacterial species. Moreover, developing CRISPRbased therapeutics against periodontal bacteria may also involve strategies that leverage this system's anti-phage properties.

Machine learning has been used in multiple ways to improve the efficacy and accuracy of the CRISPR-Cas system. One such application is being developed with a deep learning model known as CRISPRon. A study utilized a machine learning model to classify CRISPR arrays. This step is part of a broader CRISPR identification pipeline, which is used to identify potential targets for CRISPR-Cas-mediated gene editing. The Extra Trees classifier from the Python Scikit-learn package was integrated into this pipeline to classify CRISPR candidates.

The current study aimed to investigate red complex bacteria's acquired phage resistance marker profile via the genome analysis of patient samples. Further, this study attempted to identify spacer sequences, and spacers were BLASTED against the bacteriophage database to identify homologous sequences in phages with machine learning. Additionally, the present study aimed to identify the role of CRISPR-Cas in the phage resistance of periodontal red complex bacteria with a machine learning model.

#### 2. Materials and Methods

The study protocol was approved by the institutional ethical committee (IHEC/SDC/ FACULTY/PERIO/020), Saveetha Dental College. Five plaque samples of periodontitis patients were sent to a lab for identification. Later, whole-genome sequences were obtained from the NCBI NR sequence database of *P. gingivalis*, *T. denticola* and *Tanerella* [21].

The genomic query sequence, in FASTA format, was the input for the crispr.i2bc.parissaclay.fr CRISPR tool. Potential locations of CRISPRs, including at least one motif, were identified by finding the maximal direct repeats. The CRISPR pattern of two direct repeats and one spacer was considered a maximal repeat, and repeated sequences were separated by a sequence of about the same length. Whole-genome sequences of red complex bacteria were crosschecked for similarities in the NCBI genomic database. Once matched, the sequences were analyzed for CRISPR using the CRISPR identification tool (University of Paris; CRISPR.i2bc.paris-saclay.fr, accessed on 2 April 2023) [20]. Homologous sequences were obtained and confirmed using the NCBI BLAST algorithm according to standards [21]. CRISPR sequences were identified using CRISPR identification tools [21]. After obtaining the results, spacer sequences were downloaded for each bacterium, analyzed and crosschecked for homologous sequences in phage and virus genomic databases using the BLAST algorithm.

Predicting CRISPR sequences from spacers using AI can improve the efficiency of identifying and characterizing CRISPR systems. The BLAST results were added to each sequence as a separate class if necessary [22,23]. This class prediction was made by using the Orange machine learning tool. Orange supports file loading, transformation and explorative analysis and addresses all the essential phases of necessary frameworks. The pre-processing process includes cleaning the data and preparing the data. In this stage or step, we cleaned and arranged the data that we obtained. We identified a set of missing data in these data, for which the missing features were removed and outliers were removed and normalized, and the data were split into training and test data with 80/20 percent and cross-validation of 20. Machine and deep learning algorithms, such as SVM, Random Forest and Neural Networks, were applied to the CRISPR sequence dataset.

#### 2.1. Neural Network

Artificial neural networks (ANNs), modeled according to how biological nervous systems process information, comprise interconnected components called neurons that process and work together to find answers to particular problems. Similar to humans, ANNs base their learning on examples. Instead of a list of guidelines for carrying out a specific task, they are given examples to analyze and devise a solution.

### 2.2. SVM

Support Vector Machines (SVMs) are supervised learning models used in machine learning for classification and regression analysis. They are associated with learning algorithms that analyze data to find patterns and predict outcomes.

SVMs are particularly effective in high-dimensional spaces, where the volume of features (or variables) in the data is high. Even in cases in which the number of dimensions (features) exceeds the volume of samples (individual data points), SVMs can still provide an effective analysis.

#### 2.3. Random Forest

Random Forest is a supervised learning method used for classification and regression tasks. It works by generating plenty of decision trees during training. The term "forest" in the name represents an ensemble of decision trees. The main principle behind Random Forest is that a combination of learning models (in this case, decision trees) increases the overall result. Hence, for a more precise and reliable forecast, Random Forest constructs and combines many decision trees [11]. An uncorrelated forest of decision trees is produced using the Random Forest algorithm, an extension of the bagging method that uses feature randomness and bagging. A random subset of features is produced with feature randomness, which adds to the diversity and robustness of the model.

#### 2.4. AUC-ROC Curve

The classification model's performance metric is AUC-ROC. The AUC-ROC metric shows a model's class-distinguishing ability. As the AUC becomes higher, the model becomes better. AUC-ROC curves graphically show the trade-off between sensitivity and specificity for every possible cut-off for a test or combination of tests.

Classification models are evaluated using AUC-ROC. The AUC-ROC metric can determine a model's ability to distinguish classes. Models with higher AUCs are better. AUC-ROC curves are widely used to visually depict the relationship between sensitivity and specificity for each conceivable cut-off for a test or collection of tests. One way to assess a model's accuracy is the area under the curve. A good model has an AUC near 1. A model with a low AUC has the worst separability.

Precision should ideally be 1 (high) for a good classifier. Precision becomes 1 only when the numerator and denominator are equal, i.e., TP = TP + FP. This also means that FP is 0.

Precision = 
$$TP \div TP + FP$$

The recall should ideally be 1 (high) for a good classifier. The recall becomes 1 only when the numerator and denominator are equal, i.e., when TP = TP + FN, which means that FN is 0. As FN increases, the value of the denominator becomes more significant than the numerator, and the recall value decreases.

$$\text{Recall} = \text{TP} \div \text{TP} + \text{FN}$$

Therefore, the ideal precision and recall for a competent classifier are 1, implying that FP and FN are 0. Therefore, we need a statistic that considers both recall and precision. The F1-score is a statistic that considers both precision and recall: [(Precision × Recall)/(Precision + Recall)] × 2.

#### 3. Results

#### 3.1. Identification of CRISPR in P. gingivalis

*P. gingivalis* sequences were analyzed for CRISPR sequences with the identification tool, and the results show *P. gingivalis* TDC60 DNA and the complete-genome CRISPR ranking for the following sequence: 6 Crispr\_begin\_position:  $218514 \rightarrow \text{Crispr}_end_position$ : 2189300.

#### 3.2. Treponema Denticola CRISPR

*Treponema denticola* chromosomes and the complete genome were analyzed for CRISPR sequences with the CRISPR identification tool (CRISPR ranking for the following sequence: 6 Crispr\_begin\_position:  $367189 \rightarrow \text{Crispr}_end_position: 370884$ ).

BLAST—RESULTS Bacillus phage 34.2.

100% query cover; e value: 018, 100%.

The results of the above red complex organisms with the identified CRISPR query cover, spacers and BLAST were determined for sequence similarity and for the identification of microbes.

#### 4. Discussion

The immune–inflammatory disease periodontitis can destroy periodontal ligaments and adjacent supportive alveolar bone. Increased oral biofilm buildup, oral inflammation, the recession of gingival tissues and the destruction of the periodontium are symptoms of periodontitis [2]. It is primarily caused by red complex bacterial infections. Red complex bacteria include *Porphyromonas gingivalis, Treponema denticola* and *Tannerella forsythia*, which are highly invasive and secrete huge amounts of proteases and proteinases that degrade the host's collagen and destroy host immune cells, such as neutrophils [24]. These bacteria are located in periodontal pockets and lead to the destruction of periodontal tissues. Several treatment modalities have been employed to treat periodontitis, including antibiotics, pre and probiotics, lasers and ozone therapy, which have shown satisfying results [25,26]. Periodontal disease is also associated with many systemic diseases, including the risk of cardiovascular disease, rheumatoid arthritis and cancer [27–29].

Apart from the bacteria, the oral microbiome comprises several archaea, protozoa and viruses and is one of the most dynamic microbial communities in the human body. Dysbiosis of the oral microbiota can affect the host's immune system and potentially increase periodontitis incidence. Phages are the most common virus in the oral cavity, as is well known. Even though some phages have the virulence to infect and eliminate the periodontitis pathogen, they can survive in the phage-rich environment. Understanding the molecular mechanism underlying phage resistance in periodontopathogens can enable the clinical management of periodontitis more effectively [30]. In the periodontal pocket, bacteriophages are prominent viruses, and the CRISPR-Cas system in the bacterial system might protect the red complex bacteria from these bacteriophages [31]. Clustered regularly interspaced short palindromic repeats (CRISPRs) and their associated proteins (Cas) confer adaptive immunity systems in bacteria and archaea against foreign elements such as bacteriophages/viruses, plasmids and transposons. A CRISPR-Cas genetic structure comprises a series of repeats separated by spacers and a set of Cas genes nearby. In addition to defending against bacteriophages and mobile genetic elements, CRISPR-Cas appears to affect bacterial dormancy, stress, pathogenicity and immune system evasion [32,33]. The CSRISPR-Cas analysis revealed that P. gingivalis selectively acquire DNA sequences for their survival and provide protection against foreign RNA and DNA [34]. It has been shown that the virome from the sub-gingival biofilm is distinct from the healthy and periodontal disease state, which implies that the bacterial population might influence phage survival in the oral cavity [35]. We aimed to understand the role of CRISPR-Cas in mediating the phage resistance of red complex bacteria. We hypothesize that CRISPR-Cas systems in red complex bacteria could help protect themselves in the periodontal pocket environment, where bacteriophages are abundant. In addition, the system also helps the bacteria to inhibit the growth of bacteriophages/viruses implicated in the biofilm community.

Regarding CRISPR-Cas systems in the genomes of red complex bacteria from periodontally affected subjects, our study, which utilized a query sequence of all three bacteria with more than 50 percent, is sufficient to prove repeated attacks of bacteriophages in bacteria. CRISPR gained red complex bacteria sequences that constantly fought phage communities in the sub-gingival microenvironment [31]. More than 2000 oral phages have been reported or are expected to infect species of the phylas Actinobacteria, Bacteroidetes, Firmicutes, Fusobacteria, Proteobacteria and three more (few phages only). The role of phages in periodontal disease has been proven by various studies [36,37], especially cellulobacteria phages Figure 1 and Table 1, bacillus phages Figure 2 and Table 2 and pectobacteria phages Figure 3 and Table 3. Because the CRISPR space is complementary to these phages, the presence of cellulophages (e value: 0.2), (P. gingivalis), bacillus phages (e value: 0.18) (T. denticola) and pectophages (e value: 0.001) (Tannerella) indicate that these bacteria are in a constant fight with phage communities [19]. Zhou et al. found larger Shannon–Wiener diversities of DRs in periodontal disease than those in healthy periodontia, but not for spacer composition [32]. This may imply that healthy persons have a robust bacterial population resistant to phage invasion. The concept of viral modification in bacterial and host environments in the periodontal sub-gingival microenvironment must be revisited, as these bacteria have developed adaptive immune CRISPR spacer responses [38,39].



**Figure 1.** Given homologous cellulophagia query cover of 56.5%; e value: 0.2 s, 100% (BLAST RESULTS).

**Table 1.** CRISPR arrays with spacers of *immunological memories that resist future infections using spacers, or viral DNA fragments for P. gingivalis.* Homologous sequence. Identification using the BLAST algorithm with the bacteriophage database.

Spacer_Begin_Position	Spacer_Length	Spacer_Sequence		
2181514	36	AAACGAAATGAAAAAGACAACAAACAGAAGACCCTC		
2181580	36	GTGCCAGCTGCAGGGGGGATGACATAGCCATTGACGA		
2181646	36	GTGCCAGCTGCAGGGGGGATGACATAGCCATTGACGA		
2181712	36	TCCGCGCCGCGAGGTGGAGACCCTGCCGGAGGCGAA		
2181778	37	CCTGAGAAAGAGGGGAGGAGGAGCGATAGACGAAGT		
2181845	36	CGAGAATCTATTGAGTAGCGAAGTCGTCACAAAGAT		
2181911	36	AGCCATAGCTCTTCAATTTCAATTTCTTCTTTTAAT		
2181977	36	CGAAAATAACAAAAATAGATATATTTATAAAAAAGA		
2182043	34	ACTCTTATCATCTACTATCTCAAAAGCTCTTTTT		
2182107	38	TCGCTATAACCCTATGTGATTCAGGAATCGGCTTGCTA		
2182175	36	GAAACATTCGAGCCGTATTCAATTACGCCATCAATG		
2182241	38	CACCCATTGTGCCGCCGTCCTGACCGAAAACTTCTTTA		
2182309	38	TCGCTATAACCCTATGTGATTCAGGAATCGGCTTGCTA		
2182377	36	GAAACATTCGAGCCGTATTCAATTACGCCATCAATG		
2182443	38	CACCCATTGTGCCGCCGTCCTGACCGAAAACTTCTTTA		
2182511	38	TCGCTATAACCCTATGTGATTCAGGAATCGGCTTGCTA		
2182579	36	GAAACATTCGAGCCGTATTCAATTACGCCATCAATG		
2182645	38	CACCCATTGTGCCGCCGTCCTGACCGAAAACTTCTTTA		
2182713	36	ATTTCTTCATCTCGCGCTTGCTCAAAAGCGCGTTCA		
2182779	34	TGCATCAAGTCACGAACTTTCTGCGAGATGGAAA		
2182843	36	CCGTCTCGAAAAAATATCGGGACGTTTTTGTTTTCT		
2182909	35	CTCCTTCACTTTGTCGACAATGTGCACTGTATTTG		
2182974	36	AGGCGGAGTATCTCTTTGCCACCCAGTCCGCGCGCA		
2183040	36	CGTAACCGCCTCGGTAGACCGCTCCGCACGGTCGTT		
2183106	37	GTGCATTCCGGACAGCTTTCGCTTAAAAAGTTAGCGG		
2183173	37	AAGAACGAACGCCTGCGCGATAAGCACCGCGAGCGTA		
2183240	36	ATTTACCTGCAGACTTGTGCCCACCCACTTGATAGA		
2183306	36	TGGTCACGGAGCGATACCATGAGTGTTTAGTAGATG		
2183372	36	CGGAGGAGATCAGCTATGCGGATGATACCACCCGTG		
2183438	36	TCAGAAAACTCGCGTCCATCTGATAGATGTACACGA		
2183504	36	TAAACGATCGAGGCGCGGAGACCCTCCTTGCCAGTA		
2183570	37	CCGTTCAGGAAAAAGTAACCGAGCTGAAGACCATGCT		
2183637	35	AGGTGCTATCGCAGGACTGCAGGACATCCTCAATC		

Table 1. Cont.	Tal	ole	1.	Cont.
----------------	-----	-----	----	-------

Spacer_Begin_Position	Spacer_Length	Spacer_Sequence		
2183702	36	AGCCGCTCGACTTGACGCCATGCAAAAACAGATAGA		
2183768	36	TCTGCGAGTTGTGAGAGGCAATAAACTGCTGGGAGC		
2183834	35	ATTGCTGTTAATTCTGTCATCTCTTATTTCTTTTA		
2183899	35	TCTATCTGCGTATAGCCCTCGCCGTCCACGCCCTG		
2183964	37	GGGAAGCTGCTTTGCTCGCTGAGATAGCAGCACTTCA		
2184031	36	TTTATTGACGCCACCCGCCGACGAAAAAAAATCAT		
2184097	36	TTTCGGTCTTTACGTTTGTCGCCACGGATACATGCT		
2184163	37	ATTGTAAATAAATTACATGGCTATTGAAAAACAAATT		
2184230	36	GATTGTACGACTTTGCTATAAAGTCTGAGTTATTAA		
2184296	36	TATCAACAATCACCGTCATATGTGTAATATACTTGA		
2184362	36	GGGCAGGCGTATTGCCCCACTTCTCCCCGAATGCAT		
2184428	36	TGAGGAATCATATCAGTGTTTATTTTTTCATCGATT		
2184494	36	TGGAAGTGGGTAGAAGAAAGCCCCAACGTGTTTAAG		
2184560	36	CATGACAAAGAGACGGTTATCGGTAGAGGACAGGCA		
2184626	35	CGCGTGGAAGGGGCATGTACACTTGTAGTTCGCCC		
2184691	34	TGACAGGCCCTGCAGCGTGTGAGAGCGGAAATGC		
2184755	37	ATTTTCAATCATGATATTTTATTTTTCCTGCAAACGC		
2184822	36	TCGCATGTGGGAGCGCGGCGGTCTCTGCTTTACGAA		
2184888	36	TCAGCGTGATGAGCGCTTGAGGCTCCTCCATCGAGT		
2184954	36	AAGGATGATTTGGAAAATTTAGTAAGATAGTTGATA		
2185020	37	TTCTTGGAGAAAGCGAAGACCATGAACCTAAGCGTCG		
2185087	36	GGAAATATAGTTATTGTATCTACTAAAAGACATAGT		
2185153	34	ACCCCATCTTGCAGAGTATATGCGAGCAAAATTT		
2185217	34	TAGTTAGCACAGTTGCTACTATCGTAGTAGCTGT		
2185281	35	AGATAAACTTTCTTCTCGAATTAAGAAAATCGAGA		
2185346	36	CCGCGGCCATCGAGGCCACCGCCTCCGTCCTCCGCG		
2185412	37	TATGAAAACAGAAAAGAACTTCTCAGCCCTGAGTTTG		
2185479	37	CAAACTCAATGATTATCTGTCAAGAAGCAAGAAACGA		
2185546	37	AGTAAAAATTACCCTAGATGCCGAAACGGACGCCTT		
2185613	36	CAGAATTTGCACGAACAGTATGATGTTCGTGTTCTT		
2185679	36	AAGCGCGAGACAGGCCGAGCCGGCACAGCTTAGTGC		
2185745	36	AAAACGCCGATAAAATAGCGTTCGAGATTTTCCGCA		
2185811	37	TAGTTGTAGCGATTGTCTCAGTTGCATTACTCCTTAC		
2185878	37	AAATAACGAGAAAAAGAATGCTTAAATTGTTCTTCGT		
2185945	36	AGAAGAGGGTAAACTATTTGCTAATCTTGAATGCTT		
2186011	36	TCTCAATATCTTTCATAGCTACTAAAAATTTACGAA		
2186077	35	ATTGCTGTTAATTCTGTCATCTCTTATTTCTTTTA		
2186142	35	TCTATCTGCGTATAGCCCTCGCCGTCCACGCCCTG		
2186207	36	TGCCCCTCACCCATCTAACCTCGAGCCGTTGAGCC		
2186273	36	GATTACATAATGATAGACGACAGAGATTGTGCAGAA		
2186339	35	GTACTGATAATTACGCTGCAAGGTCAGACGGTGAT		
2186404	36	TTGCCAGGGCTTGCTGATGCGCGCGCGCTCCAGCTGCT		
2186470	36	ACAGAACCAGCTCCGTCAAATCTCCCGCTTTTTGTC		
2186536	36	TGCCCCTCACCCATCTAATCTCGAGCCGCTGAGCC		
2186602	36	TATTCATTCGCTCAAGCGAGGGCATCCTGCTGCAGC		
2186668	36	GCTCGCGGAGGGAGAGGGTGCCGGTATCCTGCCTCC		
2186734	36	AGAGTCTGTATGTGAAAGTGTAGTTCGAAACATTAT		
2186800	34	AGAATACTTTGAAGTTGTATTCAATTACTTTGAA		
2186864	36	CCGTTGTGACAGAGCTGCGCCGTCGTGGCTATGATG		
2186930	36	TGAAAGCCCGAAGATAATCTACACGCAAGATTGTTA		
2186996	35	CGTCATCAGGTGGATATTCTTACTGCTATCCACGA		
2187061	36	TCGTCTGCGACGTATCGCAGCTCTGCCAGCTCCGCG		
2187127	36	GCATCTTCGCAGGAAAGAAGAAGGCTCCGTCCTCGA		
2187193	36	AACAGTCAAGGGGGGGGGCCGACCTCTCCGGCGGATTT		
2187259	36	GCAGTGCGGCCGACAAGGCTAACCTCGCTCAGCTGA		
2187325	36	AATCCATGTTCATGAGGAGAGGAGAGAGGATACGAGTTCTATT		
	- *			

Spacer_Begin_Position	Spacer_Length	Spacer_Sequence			
2187391	35	CCTTTTGCTATTGCAGCAAACATATACAATAATAA			
2187456	36	ACTTCATGGATTTAGCGAGATACTCATTATAATTAA			
2187522	36	TCTTCAAAGATAGTTGTTATCAAATATCGCGCTGAA			
2187588	37	ATCTTCAAAGATAGTTGTTATCAAATATCGCGCTGAA			
2187655	36	GCTCAAGCGTCCAAGAAAGTTCAAGAAATTGTACTA			
2187721	37	TACTATGGTGTCTGCATTTGAGATACGCAAATAGCAA			
2187788	36	CCTCAATATCTTTCATTGCGACGAGAAATTTGCGAA			
2187854	36	AATCTATGATGATATAGAAGAAGACGACTTGTTGTT			
2187920	36	ATTAATGATTTTCTGAAATAAAGAAGCAGTTGCATA			
2187986	35	CTCAATATCTTTCATTGCGACGAGGAACTTGCGAA			
2188051	36	CCCCGTGGTAACTCATACCACCGACTATTCCACCGT			
2188117	36	TTTGAGTAATCATCGAATAATTATCGATTAATCATT			
2188183	36	TTATATGCATCATATTCTTAAAGTATTTTATTTGAA			
2188249	36	GACGAATACGGCGTTCATCATCGATAGTCGCGATGC			
2188315	36	CGGGCGGCCAACCGGTCACAACAAGAATAGACCGAT			
2188381	35	TTCCTTCCAGTCGCAGCTTAAGATACTGCGACTAC			
2188446	36	AATTTTCATCAGAGCATAAAAAAGGGCAAACTTTTT			
2188512	35	GATACGAGCACCAAGGCTGCGATACCGATTGCGTA			
2188577	36	CCTCGAGCAGATCTTCCTGCTCTTTGATGAGTGAGG			
2188643	35	GCAGGTTGATTATAATGTTGAAGATGCTTTAAGGG			
2188708	36	CTCGTATCGACTTTCAAGCAGGCTGGAGTGCAGCCT			
2188774	37	AATATGAGATCGGAAACAATTATAGTTGCGTCGATAT			
2188841	37	TCACAATAACCGACAAAATGTCTCGCGTAACGTACAA			
2188908	35	GTCGCCTGTTTTCTTGAACTCCTCACTGATTCGTA			
2188973	35	GATGGTGTCGACATCATACGACAATAGATCGTCGA			
2189038	35	GATGGTGTCGACATCATACGACAATAGATCGTCGA			
2189103	36	AATATTTTTCAAAAATTGTAAAACTTATTAAGTCA			
2189169	36	AAGTTGCTGATTGTCTTAGAATGAAAGGTTATGCTC			
2189235	35	ATAATATTACATAATGGCACACTGATGGTAAACTT			
2189300	34	CATGATGTACAAATATATCATGATCGTATACTAC			

Table 1. Cont.



Figure 2. BLAST results.

367189         30         TATAGGAGGTTTCAAAAATCGAAAAATCGAA           367251         30         TATAGGAGTTTCCAGACCCAGCACCACCACC           367371         30         TATAGGAGTTCCAGACCCTCAAACCCA           367433         30         TACAAAATCGAAACATCGCAAAACCAAAACCAA           367433         30         TACAAAATCGAACAAACAAT           367551         30         CAATCGAATCTTTCGAGACTATTACCACAACAAT           367651         30         CAACAAAAGTGGAATCTTACCGAACAAT           367752         30         GCGGAAAACCATATAGCGAAACAACTAAA           367784         30         GCCTGGAAACCCATATGGAACAACAACAACAACAACAACAACAAACA	Spacer_Begin_Position	Spacer_Length	Spacer_Sequence		
36725         30         TATCAAGTICAGCCTICTITAAAGCTCCCG           36737         30         AAAATCGAAGCCAGCACCACCACCAC           367387         30         AAAATCGAAGCCAGCACCACCACCACCACCACCACCACCACCACC	367189	30	TATAGGAGGTTTCAAAATGGAAAAATCGAA		
36721         30         TATAGGAGTTCCAGACCAGCACCATCACC           36737         30         AAAATCGAATGTATCCAAGCAGAAGAACCA           367453         30         TACAAAATCGAATGTATCCAAGCTTC           367591         30         GATCTGTATTTCAAACGCAACAT           367651         30         CTAACAAAAGGTGAACCCGAATT           367651         30         CTACAAAATGTGAATCTCAATCGTAATCGCAACT           367782         30         CCGCGAAACCGCATATCGCGACCCCAGC           367984         30         CCTCGAAACCGCTATACGGACGCCAACTCC           367984         30         CCTTCGAAACCCCTATACGGACGCAACTCCC           367984         30         CCTTTGAACCCCCTATACGGACGCAACTCCC           367984         30         CCTTTGAACCCCCTATACGGACGCAACTACT           368945         30         CCTTTGAACCCCCTATACGGACGAACTAAT           368946         30         CCTTTGAACCCCCCAACTTCCCCCCT           368947         30         CCTACTCACACCCCCCTATTATAGGAACGACGAAAA           368918         30         TCTTTGTAAAAAAAGCCCCCTCCCCCCCCCCCCCCCCCC	367255	30	TATCAAGTTGAGCCTTCTTTAAAGCTCCGC		
367387         30         AAAATCGAATGTATCGCAAGATTCAACCA           367593         30         TGCAAAATCGAAGCGAAGAAAGGAACTTC           367593         30         GGTTCCAATCTTTTTGGAATGATTAACAAT           367593         30         CTACAAAAGGTGAAGTTGCGTAAT           36761         30         CTAACAAAAGGTGGAATTTACCGAAGCGCAGC           367792         30         CGCGGAAAAACTATATCGAAGCGCAGCAAGCTCC           3677848         30         CCTTCGAACCCCTATAGCGCACGCAAGCTCC           367794         30         CCCTTGCAACCCCTATAGCGCACGCAAGCTCC           3677890         30         CCTTCTAACCCTATTAGCGAACGCAAGCTATA           368046         30         CCTTCTAACCCTATTAGCGAACGCAAGCAAAA           36812         30         CTATCATTTGCAATCTATAACAAA           36813         30         CCATCTAACCCTATATGCGACGCAAGAAAA           36814         30         CTACCAACCCCTATATGCAGCACGCAAAA           368310         30         GCATCTACCCCGCATACTCCACAACGCAACGCAAAA           368310         30         CGATACTACCCCCGCAAGCTCG           368442         30         CTACCATCACCCCCCAAGCTCGCCCGCTG           368453         30         TCTCGACCTTGTGTGTGTCTGGAACTCCCGCCGG           368464         30         TTCCGACCTTGTGTGTGTCGACGCGCGGGGGGGGGGGGG	367321	30	TATAGGAGTTCCAGACCCAGCACCATCACC		
367433         30         TACAAAATCGAAGCAAGAAGGAAGCAACTTC           367585         30         GATTCTGTATTTCAAACGCGATGTTAACCAAT           367651         30         CTAACAAAAGCTGGAATTTACGCAACTAAT           36761         30         CTAACAAAAGCTGGAATTTACCGAACTAACGCCGGA           367717         29         AATTAGTTGTCATTGCAATCTCAATCGCAACTACGCCGGA           367782         30         CCGTGGAACGCCTATACGCACGCCAACCTCCT           367914         30         CCGTTGCAACGCCTATACGCACGCCAACTCCT           367980         30         CGTTCGAACGCCTATACGCACGCAACTCCT           367981         30         CCTTTTTGGAATGCACACCACGCAACTTCA           368046         30         CCATTTTTGGAACGCAACGCCAACTTAGG           368112         30         CATTCATCCTATTTGGAACGCCACAGCAAAA           368113         30         GAATGAAAAAGACACCACTAACTCCGCCGATCATTTG           368310         30         GAATGAAAAAGGCACAAGGAAAAAGCCCCTATTTT           368310         30         GAATGAAAAAGGCACAAGGAAAAAGCCCCTATTATGGAGGG           36842         30         GAATGATACCACCCCTCCAAACACCCC           368548         30         TTACTGCAAACCCAATCAACGCC           368574         30         CCACGACTTACACCCCCTAAAGTCC           368588         30         TACAATGATTAGGAATTCAAAGGC           36	367387	30	AAAATCGAATGTATCGCAAGATTCAAACCA		
367519         30         GGTTCCANTCTITTGGAAGCGATGTATAACAAT           367551         30         CTAACAAAAGGTGGAATTTACCGAACCAAT           367771         29         AATTAGTGTCATATTCGAAGCGCAGCAACCCCGAA           367782         30         CCCCGAAAACCTATATCCGTAATCTTCATAGCA           367784         30         CCCTCGAAACCCCTATAGCGACCGCAACCTCC           367984         30         CCTCTGAACCCCTATAGCGACCGCAACCTCC           367980         30         GGTTCCAATTTTTGGAAGCACACCCTATATACAAT           368046         30         CCTCTGAAACCCTATAGCCAACCACTATA           388112         30         CCTTTTTTGGAAGCACACCCTATTATAG           38812         30         CTATCATAGCACCCCTATTATAG           38813         30         CGAACCTATCCCCCTATTAGGACACCCCTATTATAG           38830         30         CGAACTCACCCCCCAATCAACCACCCCTATTAGGACACCCCT           388310         30         GGAATCCACCCCCTAATTAGGACACCCCTG           388310         30         CGAACTACCCCCCAATGCAACACCCCCTG           388310         30         CGACTAACTCACCCCCAATCAACCCCCCTG           388310         30         CGACTAACTCACCAAGCAAAAGGCACTGCG           388311         30         CGACTAACTCACCAAGCAAAAGGCACTGCG           388310         30         CGACTAACTCACCCCTAATACGCAGCAGCCGCTG <t< td=""><td>367453</td><td>30</td><td>TACAAAATCGAAGCAGAAGAAAGGAACTTC</td></t<>	367453	30	TACAAAATCGAAGCAGAAGAAAGGAACTTC		
367585         30         CATTCICTATTICAACCCCATCITTCCTATA           367717         29         AATTACTTICTCATTGAAGCTGAACCAT           367717         29         AATTACTTCTCATTGAAGCTGAACCCCGAA           367782         30         GCCGGAAACGCCTATATCTCATTACTGAAGCTGAACCCCGA           367948         30         GCTTGGAACGCCTATACGCACGCAACCTCC           367940         30         GCTTCGAACGCCTATACGCACGCAACAT           368946         30         CATTCTGTTTGGAATCCTATAGCACACGAACAAT           368112         30         CTTTTTTGTACTCCATTAAGCCACGCAACAT           36812         30         CATTCACATCCATTACGCACGCAACAAT           36812         30         CATTCACTATACAGCCACGCAACAAT           36813         30         CATTCACTACCACCCTATACGCACGCAACAAA           368510         30         GAATGAAAACGCCGCATACAACACGC           368510         30         GAATGAAAAGCGCATACACACGCAACAACGCA           368574         30         GCAGTAACTCACACGCGCAATCACACGCC           368574         30         GCAGTAACTCACACGCCAATCACAAGGC           368574         30         GCAGTAACTCACAATACGCCAACACAGCC           368575         30         CTTGCATCACATTCGCAATCACAAGGC           368574         30         GCAGTAACTCACAAGCCACATCACAAGGC           368575         30 <td>367519</td> <td>30</td> <td>GGTTCCAATCTTTTTGGAATGATTAACAAT</td>	367519	30	GGTTCCAATCTTTTTGGAATGATTAACAAT		
367651         30         CTAACAAAAGGTGGAACTTIACCGAACAAT           367717         29         AATTACTTGCTATTGGAAGCCGGA           367782         30         GCGGAAAAACTATATCGTAATCTCATAGA           367984         30         CCTTGGAACGCCTATAGCGAACGCAAGCTAG           367980         30         GGTCCAATCTTTTTGGAAGATTAACAAT           368046         30         CATCTAGAATCCTATAGGACGCAAGCAAGTAG           368142         30         CTTTTTTTTTGGAACCGCTATTGCGACGAAAAAAAGGCGAAAA           38810         30         GAATCTACCAGCAGAAAAAACGCCGCTATTACG           388310         30         GAATCTACCAGGAGCAAAAAACGCCCCT           388310         30         GCAATCAACCAGCCTTACGCACGAAAA           388310         30         GCAATCAACCAGCCCTATAGCGAAAAACGCCCCT           388310         30         GCAATCAACCAGCCCTTACGCAGAAAAACGCCCCT           388374         30         GCAATCAATCACACACACCCCCTT           38838         30         TTCCACAGCACTTAGCGAAACGCCCT           38870         30         CTTGCAGCAGCAAAAACTTCCAAGCCACTAAGGC           38871         30         CTTGCAGCAGCAAAAAGCTTCCAAGCCCCT           38872         30         CTTGCAGCGCTTGTGGTGAAAGGCAACGAACGCCCCT           38873         30         TTCCCAAAACTCCACTCAAAAGCCCACAAGGC           388772	367585	30	GATTCTGTATTTCAACGCGATGTTGCTAAT		
367712         29         AATTACTTCTCATTGAAGCTGAAGCCGGA           367782         30         GCGGAAAACTATATCGTAATCTTCATGAG           367948         30         GCTTGGAACGCCTATAGCGACGCAAGCTCCT           367940         30         GCTTCGAACGCCTATAGCGACGCAACGCAAGCTCC           367940         30         GCTTCTGAATCGCTATTGGAGGAAGAACAAT           368046         30         CATCTAGAATCCTATTGGAGGAGGAAGAA           368112         30         CATTTGTTTTGAACTCCGAGGAGAAA           368128         30         ATTACTTTTGCAATAAGGCACGCAGGAAAA           368310         30         GAATCTACCACCCCGCAATCACGAACGACGAAAA           368310         30         GAATCTACCACCCGCGATCACACAACGCA           36842         30         GCAATTACCCCGCGTCACCAACGAAAAAGCCA           368430         30         TCCAACATCACCGCGAGTAAATGCCCGCGTG           368541         30         GCACTTAATTGCAAAGCCACAGTAAATGCCGCGGGAGAAAAGCCGCG           368570         30         TCGTACACATACTCACAAGCCAATCAAGGC           368581         30         TACCCTCACCATTATACGAAACGCCAATCAAGGC           368590         30         TCGCAGCATCTCACAATAAGCTCACATCAGGCAGCATCACGCAGGAAC           368702         30         TCTTGCTCAACATAAGCTCCAACAGCATCACGAGGAAC           368903         30         TCCTCACCATTAGCAATACGCTCAACGACGACGACGACGA	367651	30	CTAACAAAAGGTGGAATTTTACCGAACAAT		
36782         30         GCGGAAAAACTATATCGTAATCGTCATAGG           367914         30         GCTGGAACCCCTATAGCGACGCAAGCTCC           367919         30         GCTCCAATCGCCAAGCTAAGCGACGCAAGCTAAT           368946         30         CATCTAGAATCCTATTAGGACGCAAGCTAAT           368917         30         ATTACTTTTGGAAAACCCTATTGGACACAAAA           368178         30         ATTACTTTTGGAAAAACCCCTATTGGACACAAAAA           368178         30         ATTACTTTTGGAAAAAACCCCTATTGGACAAAAA           368310         30         GAATCTACCACCCTGCATACCGACAAAAA           368310         30         GAATCTACCACCCCTGATCACCAACAGC           368310         30         GAATGAAAAAGGACAAGGAAAAAACGCCCTT           368310         30         GCACTACATCACCGCGATGAGTAAAAGCGCCCT           368310         30         GCACTACATCACCACCGCCAGTGAGAAAAACGCCCT           368310         30         GCACTACATCACAAAAACGCCCCCT           368310         30         TCCCACCCTTGCGAAAAACGCCCCCT           36840         30         TTCCACCCCTTGCGAAAAAGCTTCACAAGC           36840         30         TTCCACCACCATTAATACCAATCACAGC           36840         30         TTCCCACCATTAATCCACAAGCACCCCCCTTGAGGACAC           36840         30         TACCTCACCATTAATCCTCACATCACATCGAGTAGCCCCCCCC	367717	29	AATTAGTTGTCATTGAAGGTGAAGCCGGA		
367948         30         GCTGGAACGCCTATAGCGACGCAAGCTCC           367980         30         GGTTCCAATCTTTTGGAATGATTAACAAT           368906         30         CATCTAGCAATCCTATAGCGAAGGATAAT           368917         30         ATLACTTTTGGAATCATTAACAAT           368112         30         CCTTTTTTTGGAATCATAACAAT           368178         30         ATLACTTTTGGAATCCTATAGCGAAGGAAAA           368244         30         TCTTTGTATATAAAAAAGCGCGTATTAGCAGGAAAA           368300         30         GAATCTACCACCCTCAATACTCCGCACCTCT           368376         30         TGCAACATCACCGCGAAGCAAAGCAAAGCAACACC           368578         30         GCATTAATTGGAAGCACTACTACAAAGCCACTTG           368578         30         GCAGTAACTCACAAGCCACTTGCAAAAAGCCACTGTG           368574         30         GCAGAAGTTAATGCTGACAAGCAATCAACGC           368570         30         TGTCCACCTTGCGAAAAAGCCAATCAACGC           368571         30         TGCCACCTTAGCTGAAAAAGCCAATCAACGC           36870         30         TGTTCGATCAATTAGCAATTAGCAATCAACGC           368712         30         CGTTAACTCCAAAAGCCAATTGCAATAGCCCATC           368904         30         TGCCCCCAATTAAGCGAACTTACCCAAGCAGTA           369036         30         CCTATAACCACCTAACAGCCACCAATGAGCAGTA           369	367782	30	GCGGAAAAACTATATCGTAATCTTCATAGA		
36794         30         CGCTGGAACGCCTATAGCGCAGCAACGCTC           368046         30         CATCTAACAATCCTATTAGCAGTAATAACAAT           368046         30         CATCTAACAATCCTATAAGCAGTAATAACAAT           368112         30         CATCTAACAATCCTATAACGCACGAATGATTAAGA           368128         30         ATTACTTTTGAACTCCTATTAAGAGCACTAAGAATAGAT           368244         30         CATCTAACAACCCCCCCAATAACTCCCCCCTATT           368310         30         GAATCTAACCACCCCCGAAAAGCCCCCATT           368376         30         GCAACAACGACCAACAAGCACAAAAGCCCCCCT           368508         30         TGATTATTTGGAAGGCAATAGCCCCCT           368508         30         TGCATCACACCCCCAATCAAACGCC           368508         30         TGCAGCCTTGTGCAAAAGCCACCTGCAACAACGCC           368508         30         TGCAGCCTTGTCGAAAAGGCAATCAACGC           368508         30         TCCGACCCTTGTCGAAAAGCCACTCCCCCT           368508         30         TCCGACCCCCATTTATAAGCAATTCCCCT           368508         30         TCCGACCCCTTATAAAGCCATTCCCATCCCCCT           368640         30         TACCCTACCACCATTTATAGCACATCCACATC           36870         30         TGTTAGCTACAAATTCCCATTCCAATTCCAATCACCATCACATC           368904         30         TGCTACAACATCACCCTAATTCCAAGCATTCCAAAGCATC	367848	30	GCTGGAACGCCTATAGCGACGCAAGCTCCT		
367980         30         CGTTCCAATCTTTTGCAATGATTAACAAT           368916         30         CCATCTAGAATCCTATAAGCCACGAAGTAAT           368112         30         CCTTTTTTGCAATCCTATAGCCACGACGAAT           368113         30         ATTACTTTTCGAATAAAGCCGTATTATAG           368244         30         TCTTTGCAATAAGCCCGCACTATATAG           368310         30         GAATCACCACCCCCAATCACTCACAACACCC           368576         30         GCAACATCACCGCGACCACTACAACACCCCCCTAT           368578         30         GCAATGAAAGCACACGCCACCTGTGGAAAAGCGCACCGCC           368580         30         TGCATCTCACAAGCCACTTGCGACAAGCGCACCCCG           368574         30         GCAGTAACTCCAAAGCCACTTGCAAAGCCACTGCG           368570         30         TGCACCTCACCATGTGCGAAAAGCCACTGCGCACGCC           368571         30         CGACGAAGTTATTATTCTGAAGCACAGCC           368572         30         CTTTGGCAACAAGCCACATTCCAAGCCCC           368588         30         TACACTGCTGTGTGTGTGTGCGAAGCCC           368590         30         TGCTCCCAATAACCCACAGCC           368904         30         TCCCCCAACTGCGCAACAGTGCCCACAGCGCC           369030         30         TGTTACGTCCCAACAGCTGCCCACAGCGCC           369030         30         TTCCTGTTTTCGGCACCCAACTGCGCCCCCC           3693	367914	30	CGCTGGAACGCCTATAGCGACGCAAGCTCC		
368)46         30         CATCTAGAATCCTATAAGCGCATGAGGAAT           368112         30         CTTTTTGAACTCCTATTGCAACTCATAG           368178         30         ATTACTTTTGCAACCCATAG           368310         30         GAATCTACACCACCGCCACTGAAAA           368310         30         GAATCTACCACCCCCCAATACTCCGCCCTAT           36854         30         GCAACAACCACCACCGCGAAAAAGCCCCCCT           36856         30         GCACTAACTCACAACGCAATACTCCACCCCCT           368574         30         GCACTAACTTCACAACGCACTTGAACTTCAACGCC           368574         30         GCACTAACTTCACAACGCACTTGACGAATCAACGCC           368574         30         CCATTGACTACAACGCACTTGACGAACTCAACGC           368574         30         CCATTGACTACTAACGACTCAACTGC           368572         30         CTTTGGTATCAATTGGACTTCACACTG           368588         30         TACCATCACCACTATTATACTGCACTGAACTTCACGCATG           36870         30         TGTTACATCAATTAAGCAATTCCCATG           368970         30         TGTTACATCAAATACCATTGACTGACAGGACATTGAGACT           369903         30         TGTTACATCAAATTCCACATTCACATGACATGACG           369914         30         CTATCTACTTTGCGGACATGCACATGCACTGACGACGT           369936         30         TTCCTACAACTACCATTCCACATCTCACTTACGCCCCAGT	367980	30	GGTTCCAATCTTTTTGGAATGATTAACAAT		
368112         30         CCTTTTTGTAACTCCTATTGCAACAACGCCGTATTAG           368244         30         TCTTTGTATAAAAGTCCGCGATTATAG           368244         30         GAATCAACACCACCCCAATACTCCCCCCTATT           368310         30         GAATCAACACCACCCCCAATACTACCAACACC           368376         30         GCAACTCAACCCCCGATCACTAACAAACGC           368574         30         GCACTAATTGGAAGCCACTCACAACGCCT           368574         30         GCACTATTTGGAAGCCATCACAAGGCAATCCACAGGC           368576         30         TCCCACCCTTGGAAAAGGCAATCAACAGCC           368576         30         CCGACGAAATTATTATTCTGAACTTCGACGCGC           368576         30         CCGACAAAAGGCAATCAAAGGCAATCCAAGGC           368576         30         CCGACAATATAGCAAAGGCAATCAAGGCC           368772         30         TTGCTCTGCAAAATTGCATTCGTCTTGGAAGGAACCTAAGGAGC           368904         30         TACCCTCCACACTTATAAGCAAGAGTGCAAGGAAGGAGCGAACGAA	368046	30	CATCTAGAATCCTATAAGGCACGAAGTAAT		
368178         30         ATTACITITICGAAAAAAAACGCGTATITATAG           368310         30         GAATCTACCACCCTCAATACCCGCGATAT           368310         30         GAATCTACCACCCCCCAATACACAAACAAC           368376         30         GTCAACATCACCACCCGCATCACCAAAACAACCC           368578         30         GCACTAAACATCACAAGGAAAAAACGTCCCATGCGT           368574         30         GCACTAAACTCACAAGCCAATCAAAGCC           368574         30         GCACTAAACTCACAAGCCAATCAAGGC           368574         30         GCACTAAACTCACAAGCCAATCAAGGC           368570         30         CGAGAAGTTATTATTCTGAACTCGAGCAACAGGC           368706         30         CGTTGGTATTATTAGGATTTCCTAAGGCAAC           368904         30         TACCATCACATTAGGAATCGCCAATG           368904         30         TACCACTCACATTAAGCAATCCGACATG           369905         30         CCTGTAAAGGAAGTCGACAGTAGGCAAAGGACGTA           369102         30         TGCTACCACATTAAGCCATGGAAGCAGGA           369103         GCACTAACATCCACAGCATAACCTCCCCCCCCCCCCCCC	368112	30	CCTTTTTTGTAACTCCTATTTGCAGCTATG		
36824430TCITIGTATTATAAAGTTACCAGGGAAAA36831030GAATCTACCACCCCTATT36837630GTCAACATCACCGGATCACTACAAACAGC36841230GAATGAAAAGGACAAAGGAAAAAGCTGCACAAACAGC36850830TGATTATTIGGAAGGCATGACAAAGCGCACCT36850930GCACTAACTCACAACCCACACAGGC36864030TCCGACGCTTGCGAAAAAGGCAATGCAGGC368670630CGAGAGTTATATTCGAACTTCACAAGGC36870730CTTTGGTATCAATTAGGATTTCCTAAAGTC36883830TACAATGATTAATCCAACAAGGCAATCCAATGG36890430TACCCTCACCATTATAAAGCAATTGCAATG36890430TACCCTCAAAAATCCCAATAGTTGGAAG36903630CCTGATAAGGAAGATTGCCAATAGTTGAAG36903630CCTCAAAAATCCCAATAAGTGGAAC36910230TTCCTCCAAATAATCCAATAGTGAAG3691230CTACTCAATATACCCTACAATAGTGAAG36913630CCAAAAGTTCCAATATACCTACAATGATC36936630AACCAATCTGTCCAAAGATACCTACCAAGTAGCTCC36936630AACCAATCGTGCCAATACCTACCATGTGTGT36936630AACCAATCGTGCGCAATACCCTACCAACGCTCC36936630AACCATTGATCGTGGTAGTTTTCGGACTCTGT36936630AACCATTGCTCGTAAGAATACCTTACCAGTTGT36936630AACCATTGCTCGTAGCTGTAGCTGTAGTGT36936630AACAATCGCTGCAGCTAGCTGGTGGT36936630AACATTGCTGCTGTAGCGTGTGTGTGTGTGTGTGTGTGTG	368178	30	ATTACTTTTCGAAAAAAGCCGTATTATAG		
368310         30         GAATCTACCACCCCCATACTCCCCCATT           368376         30         GCAATCACCCCCCCCCT           368442         30         GAATGAAAAGGACAAGGAAAAAACCTGCCCT           368578         30         TGATTATTGAAAGGCAATGAAAAGCTGC           368574         30         GCACGACAAGGAAAAAAGCTGC           368576         30         CCACGACAGCAAAGCAAAGGAATCAAAGC           368772         30         CCACGACGACTTTTATTCGAACTCCACATCG           368773         30         CGAGAACAAGTATATTTTTCGAACTCCACATCG           368774         30         CCACGACGACTTATTATTCCCAATCGCACATCACACCC           368772         30         CTTGGTATCAATAGGAAGTACATCCCATG           368838         30         TACACTCACCACTTATAAAGCAATACCATCCCATG           368904         30         TACCCTCACCATTATAAACCAATCCATG           368905         30         CCTCAATAGAAGACTGCAAAGAGCTA           369102         30         TGCTACACTCAACAATAAGCCCAACAGTG           369103         30         CTACTCACTCACCATTGGTACCCCCAGAGGACGT           369300         30         TTTCTCTGTTGCCAGTCCCAGAGCAGCT           369432         30         AACGAATAACCTTACCATTCGCACTTACGCCCCCC           369432         30         AACGAATAACCTTACCACTTCGAACCTCC           369438         0	368244	30	TCTTTGTATTATAAAGTTAGCAGAGGAAAA		
368376         30         GTCAACATCACCGCGATCATACAACAGC           368442         30         GATTATTTGGAAGGCATGACAGGAAAAAGCTGCCCT           368508         30         TGGATCACAAGGCAAGGAAAAGCGCTGCCGCT           368574         30         GCCAGTAACTCACAAGCCACTTGAGAGTTG           368576         30         CGCAGAAAGTATTATTCTGAAAGCATCACAGC           368706         30         CTTGGACACTTCACAAGCACATCACAGC           368707         30         CTTGGACAATTCACAAGCACATTCACATCG           368838         30         TACAATGATTGCTTGTTGTTGTGAGAAC           368904         30         TACACTCAAATAACCAATTCACAATCGAATCGCAATGAGCA           368905         30         CCTGATAAGGAAATTCGCAATGAAGCAAGGCAAGGAGGAAGGA	368310	30	GAATCTACCACCCTCAATACTCCGCCTATT		
36844230GAATGAAAAGGCAGAAGGAAAAAGCTCCCCT36857830GCACTAACCTACAAGCCACACTGCGAAAAAGCTG36864030TTCGACCCACAGCCACCACTTCAGAGCTG36864030CGAGAAGTTATTATCGAAAGGCACTCACAAGCC36877230CTTTGGTATCATTACGATTCCAAAGCC36877230CTTGGTATCAATTAGGAATCAGGAAC36890430TACCATCATGCTGTTGTTGTCTGTAGGAAC36897030TGTTACGTCAAAAAATCCAATGAGGTA36897030TGTTACGTCAAAAAATCCAATAAGGTAAGAG36910230TGCTACATCAAAAAATCCACAGTCAACAAGAGGTA36911830CCTACATCAAAAATCCCACAGTCAACGATGGCAAAGAGGTA36923430TTTCTTCTGTTTGTGCATGCCAACGTCCC36930030TTTCTTCTGTTTGTGCATGCCAACGTCCC36943230AACGAAAGTTAACCTTACCAATTAGGTAC36943830TTCCCAAAAGTTGCAAAATAATCATGGGT36943930TTCCCAAAAGTTGTAAGAAAATAATAACTACT36963030AGCGTAATAACCTTACCGACGTGCGGATAACGGT36963030AGCGTAATAACCTACCGACGCGCGT36963030AGCGTAATAGCAACGAAACG36996030ACCATATAGCAACCGAAACG36998130AACCATATAGCAACCGAAACGGAAACCG37002230GGGCATATTAGCAACAACTCA370035630AACCTTTCAATAGTAAAAACCTCAT37029030AACCTTTCAATAGCAAACTCAA37035630AACCTTTCAATAGCAAACCAAGGCAAACGG37042230CTTGGAAAAAGGTCTTGGTTAAAAAAACACAGCAAAGG37035630AACCTTTCAATATAAAAACCAAGTCA37042230CTTGCAAAAAGGTAAACGTAAAAAAGGTGTTGGA	368376	30	GTCAACATCACCGCGATCACTACAAACAGC		
36850830ICATIAITICGAAGCCAIGCIAAAIGCIG36857430GCAGTAACTCACAAGCCAAGCCAATCAAGCC36870630CGAGAAGTTATATTCGAAACTCAAAGC36870630CGTGGATCAATTAGGATTACGAATCGAAAGCC36877230CTTIGGATCAATTAGGATTCCTAAAGTC36883830TACAATGATTGCTGTGTGTTCTGAAAGTC36890430TACCACTGATTATAAAAGCAATTGCAATGGAAG36890430CCTCACCAATAAAAGCAATTGCAAGGAAG36903630CCTGATAAGGAAGATTGCGCAAAGAAGGTA36910230TGCTACACAAAATAACCCAGTGAAG36910330CTATCTACTCAATAAAACAGTGAAG36910430CTATCTACTCAATAAAACCCCAGAAGAGGA36910530CCAAAAGTTCACAGTCAAGCGCAAAGAGGTA36910830CTATCTACTTTTGGGACTAGCCAAACGTCG36930030TTTCTCTTTTGGGACTAGCCAAACCTCC36930030TTCCCAATAGTGGACTTAGCCCC36930030TTCCCAAAAGTGATGCGAATACCTCCCC36930030ACCAATGGTGGACTTAGCACTAGCACGAAACCGC36930030TTCCCAAAAGTGATGGAATAACCTCCCCCCC36936630AACAATCACCTACCACTAGCAATGCAACGCCCCCT36949830AACAATCACCCGTGAGGAATACCCCCCCCCC36956430AACAATCACCCCGAGAAAACAGAGTT36969630GGGCATATTATGCAGATATGCAACGAAACG36969630GAACATATGCTGCACTATGCAAACGAACCA36969630AACCTTTCAATAGTAGCAACCAGAGTACGCAAACCAG36969630AACCTTCCAATAGTAGCAACCACCAGAACCAG36969630GAACATATGCTGCAATTATAGCAAGCAACCAG36969630GAAAATTGCA	368442	30	GAATGAAAAGGACAAGGAAAAAGCTGCCCT		
36857430GCAGTAACTCACAAGCCAATCCAAGCC36864030TTCGACCTTCTCGAAAGGCAATCAAGGC36870630CGAGAAGTTATTATTCTGAAACTTCACATCG36877230CTTTGGTATCATTAGGATTCGAAGCA36883830TACCAATCATGGTGTGTGTTGTCTGAGGAAC36890430TAGCCTCAACACATTAGGAATGGCAAGGAAGGATGGAAG36990530TGTTACGTCAAAAAACCAATAAGCAATGGAGAG36903630CCTGATAAGGAAGAATGGCGAAAGAAGGTA36910230TGCTACATCAAATAACCATCAAATGGAAG36918830CCAAAAGTTCACAATAACCAAAGTGAAGAGGTA36930030TTTCTTCTGTTTGTCCATGCTACAAACTCC36930030TTTCTTCTGTTTGTCCATGCTCAAACCTCC36936630AACAAGTTGGAACCTTAACGAACGCCCCA36943230AAGGGAATAACCTTACCATTCGTGTTAGG36949830TTCCCAAAAGTTGAACGCCGCGT36969630GGGCAATTATGCAAGCAACGAACGC36969630GGGCAATTATGCAACGAACGCACGAACG369989430AACAATAGCCGCGCGTTAGAGGAATACCGCAGGAATACGCAACCG369989430AACATATGCTGCGCTCTTTCTCGAGTACTC36989430AACATATGCTGCGCCTTTTCTCGAGTACTC36989430AACATATGCTGCGCCTTTTCTCGAGTACTC36989430AACATATGCTGCGCCTTTTCTCGAGTACTC36989430AACATTTGGAAAAGAATAAAAACAACGCAG37002630GGCCATATTAGCAACGGAAACCGAAACCG37012830CTTGGAAAGGAATTAACAAGTCAAAAAAAAAAAAAAAAA	368508	30	TGATTATTTGGAAGGCATGAGTAAATGCTG		
368640       30       TICGACCCTIGTCGAAAGCCAATCAAGCC         368776       30       CGTAGCAATTATTCTCGAACTTCACATCG         368772       30       CTITGGTATCAATTAGCATTTACTCTCAATCG         368838       30       TACAATGATTGCTTGTTGTTCTGATGGAAC         368904       30       TACCATGCTCACAATAAGCAATTCGCATG         368970       30       TGTTACCTCAAAAAATCCCATAAGGAAGGTA         369036       30       CCTAAAGAAGATTGCCAAAGAAGGTGAAGG         369102       30       TGCTACATCAAAAAACCCTACCAGATGGAAG         369188       30       CCAAAAGTTCGCAATCACACTGCCAAGTGAAG         369300       30       TTTCTTCTGTTTTGGCAATCCCAATTGGTAC         369300       30       TTTCTCTGTTTTGGCAATCCAACTCCC         369300       30       TTCCTCATTTTGGGAACCTTAACTACTGCC         369300       30       TTCCTCATTTTGGGAACCTTACCATTCGCC         369300       30       AACAATCTTCGCATTCGCAAACTCC         369366       30       AACAATCTTACCATTCGCAACGCAACTTATGCAACGAAACG         369498       30       AACAATCAGCGTGAGGGAATACGCCGCGT         369564       30       AACAATTGCAACAACAACAACGAACG         369696       30       GGCATATATGCAACAACAACGAACG         369894       30       AACAATCTTCACTTGCAATAGCAACGAAACG         369894	368574	30	GCAGTAACTCACAAGCCACTTTGAGAGTTG		
36870630CGAGAAGTIATIATICIGAACTICACATCG36877230CTTIGGTATCAATTAGGATTICCTAAAGTC36883830TACAATGATTGGCTTGTGTTGTGTCGAGGC36890430TAGCCTCACCATTATAAAGCAATTCGCATG36890530CCTGATAAGGAAGATTGGCGAAAGAAGGTA36903630CCTGATAAGGAAGATTGGCGGAAAGAAGGTA36910230TGCTACAACAATAACCCTACAAGTAGCCT36923430CTATCTACTTTTGGCAACCCTAATTGGTAC36930030TTTCTTCTGTTGTGCATGCCAAACCTCC36936630AACAATGTGTGATTTTCGGAACCATTAGTCC36938030TTCTCTGTTGTGCATGCCAACCTTAGTCCC36943830TTCCCAAAAGTTGAGCTGATAACCTTACGCCGGT36943830AACGGAATAACCTGCCGTGATACGGTGT36954930AACAATGACCCGTGAGGGAATAACCACCTC36963030AGGTAATGCAGAAGAATTTATAAAATAACAACTG36963030GGGCATATTATGCAAGAAAATAATAACAACG36963030GGGCATATTATGCAAGAAAGTGCAAACG36963030AACATTGCAGAAAAATAAAAACAAGTCA36963030AACATTGCAGAAAAATAAAAACAAGTCA36982830GAACATATGCCGCTGCTTCTCGAGTAACCAA36996030ACCTTTCAATAGTAGCAACGAAAGTCA37002630GTGCATATATGCAACATTAACCTACCTAT3702430AAACCTTGCAGAATTGCAACTAAACAAGTCA37035630AAACTTCGAGGTACTAAAAAAAAAAAAAAAAAAAAAAAA	368640	30	TTCGACGCTTGTCGAAAAGGCAATCAAGGC		
3687/230CHILGGIAICAAHIAGGATHCCHAAAGTC36883830TACAATGATIGCTTGTTGTTCTGAAGGAAC36890430TGTTACGTCAAAAAATCCAATAAGTGAAC36890430TGTTACGTCAAAAAATCCAATAAGTTGAAG36903630CCTCATAAGGAAATCCAATAAGTTGAAG36910230TGCTACAATCAAATAACCCTACAAGGTAAG36910330CCTAATAGGAACATCCCGAAAGAAGT36914830CCAAAAGTTCACATCCAACGTACCGACGT36930030TTTCTTCTGTTTGCCATGTCCAAACCTCC36930030TTTCTTCTGTTTTCGGACTCGATACGCT36936630AACAATGTGTGGTTTATCGGACTTAGTCCC36936630AACGAATAACCTTACCATTCTGTCTTATG36949830TTCCCAAAAGTTGAGGGAATACCCCCGCGT36956430AACAATCAGCCGTGAGGGAATACCGCCGCGT36965030GGCATATTATGCAGATATGCAACGAAACG3696630GGCCATATTATCAAGAGAAAAAAAAAAACAACTT36969630GGCCATATTATCAAGGAACAACGC36976230CTTGGAAAAGAATTTATCAACGCAACGCAAACG36989430AACATTGCAGCATCGCGCAAACCAG36989430AACCTTTCAATAGTAGCAACGGCAAACCAG37002630GGCCTCTATATGCAGATAGCAACCAACGAACCAG37015830CTTGGAAAAGAATTTATAAAATGCAAACTATA37029030AAACTTTGAGGTACTAAATGAAACAAGTCAT37035630TATCTCAAAAAAAGGTCTATACTAGCAACGCAACGCA37035630ATATCTTTTCCAGGTAAAGAAGGAAGG37055430TAATCCAAAAAAAGGTCCACCGCGTT37088430TTGCAGCTCAACGACACGGCGCCCTT37088430TTGCAGCTCAACACACGTGGG	368706	30	CGAGAAGTTATTATTCTGAACTTCACATCG		
36883830TACAATGATIGCTIGTIGTIGTIGTIGTIGTIGGAG36890430TGTTACGCATTATAAAGCAATTGGCAATTGGCATG36903630CCTGATAAGGAAGATTGGCGAAAGAGGTA36910230TGCTACATCAAATAACCCTACAAGTGAAG36910330CCTACATCAAATAACCCTACAAGTGGAGG36910430CCAAAAGTTCACAGTCACAGTGGAGGT36910530CCAAAAGTTCACAGTCATCCGAGTAGACGT36930630CTATCTACTTTTGGGAACCCTAATTGGTAC36936630AACAATGTGTGATTTTCGGATCTCCGAAACCCTCC36936630AACAATGTGTGATATTACCATTCGTTATG36943230AAGGGAATAACCTTACCATTCGTTATG36956430AACAATCAGCCGTGAGGGAATACGCCGCGT36963030AGGTTAATGATGAAAAAAATAATAATAACTACT36963030AGGTTAATGATGAAAAAAATAATAACTACT36963030AACAATCAGCCGTGAGGGAATACGCCAGGT36963030AACAATTGAGGATACTTACAACGAAACG36963030AACAATTGCAGATATGCCAGCAAACG36963030AACATTGGAGAAAAAAAAAAAATAATAATAACTACT36963030AACATTGCAGATATGCCAGCAAACG36976230CTTGGAAAAGAATTTATAAAACAACTCA36988430AAACTTTGAGGGATCAAAAAAACGCAAGCA37002630GCCTTAGTTACTTTACCAATAAAACAACTCA37029030AACCTTGGGGAAATTGACAATGCAAAGGAAACG37035630ATATCTTTTGTGCGTTAAAGTAAAAAAAGTCA37042230TTTGAAATTCCACACGAAGGTTACAAGGAAGGT37048830CTTGGAAAAAGGCCCACCC37055430TAATTCAAAAAAAGGCCCACCCCCTT37086630ATATCCAAAA	368772	30	CITIGGIAICAATIAGGATITCCTAAAGIC		
36890430TACCTICAACATITAAAACATGCATGG36897030TGTTACGTCAAAAAATCCAATAAGTGAAG36903630CCTGATAAGGAAGATTGGCGAAAGAGGTA36910230TGCTACATCAAATAACCTACAAAGTGGAAGGTA36918830CTATCTACTTTTGGGAACCCTAATTGGTAC36930030TTTCTTCTGTTTGTCCATGTCCAAACCTCC36936630AACAATGTGGATTTTTCGGACTTAGCCC36949830TTCCCAAAACTTACCATTCTGTTTGTG36956430AACAATCAGCGGGACTAACGATTGGT36966630AACAATCAGCCGTGAGGGAATAACCTCCC36956430AGCGAATAATGATGAAAAAATAATAACTACT36966630GGGCATATTATGCAGAGTAATGCAACGATTGGT36956430ACCTTTCGAAAAAAAATAATAACTAACT36966630GGGCATATTATGCAGATAGCAACGAAACG369676230GGGCATATTATGCAGATAAGCAAAGGATT36989430AAACTTTGAGGTACTAAATAAAACAAGTCA36998030ACCTTTCAAAGTAACCTAACAAGCA37002030ACCTTTCAAAGTAACCAACAAGTCA37092130GGGCATATTATGCAGATAGCAACGAAACG37035630ATACCTTTACCTATACTTTACCTAACTACCAACGAAAGG37035630ATACCTTTACCAATGTAAATAAAAAAAAAAAAAAAAAAA	368838	30	TACAATGATTGCTTGTTGTTCTGATGGAAC		
3689/030IGITACGICAAAAAAICCAAIAAGITGAAG36903630CCTGATAAGGAAGATGGCGAAAGATGAAGGA36910230TGCTACATCAAATAACCCTACAAGTGACGCGA369123430CCTATCTACTTTGGGAACCCTAATTGGTAC36930030TTTCTTCGTTTGTCCATGTCCAAAACCTCC36936630AACAATGTGTGGATTTTTCGGACTTAGTCCC36943230AACGGATAACCTTACATTGGTGATGGT36943230AACGACCGCGTGAGGGAATACGCCGCGT36948330TTCCCAAAAAGTTGATGCGATAGCCGCGGT369494830AACAATCAGCCGTGAGGGAATACGCCGCGT36965430AACAATCAGCCGTGAGGGAATACGCCGCGT36966030GGCCATATTATGCAGATATGCAACGAAACG369676230CTTGGAAAAGAATTATAAAATAACAAGTCA36988430AAACATTGCTCGGCCAAACCAACGAACGA36996030ACCTTTCAATAGTAGCAACGAACCAG36996030ACCTTTCAATGATGACAACAAGCAG37002630GCTCTAGTTACTTACGAAAACAAGTCA37002630CTTGGAAAAGAATTTATAAAAACAAGTCA37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37035630ATATCCTTTAGCTTAACTTATGCAAAAAA37042230TAATTCCAAAAAAAGGCTTTGCAAGGAGGTT37048830GAAAATCCAAGCGCGTTCCACTGGAGAGGTT3705430TAATTCCAAAAAAGGCCTTGGAGAGGCTT37055430TATTCCAGCTAACGACCGCGCACC37075230TTGCGCAACACGGGCGTTCCACC37075230TTGCCACCTGCAACAACGGCGCTTCCACC37084830CTCGCTCAGCCCTCAAACAGGTCACCACC <td>368904</td> <td>30</td> <td>TAGCCTCACCATTATAAAGCAATTCGCATG</td>	368904	30	TAGCCTCACCATTATAAAGCAATTCGCATG		
36903630CCICAIAAGGAAGAI ICGCGAAAGAAGI36910230TGCTACATCAAATAACCCTACAAGTGAAG36910330CTATCTACTTATGGAACCTTACCAGTGAAG36923430CTATCTACTTTGGGAACCCTAAATGACCTAC36930030TTTCTTCTGTTTGCCATGTCCAAACCTCC36936630AACAATGTGATGTGATTTTTCGGACTTAGTCC36943230AACGACCTACCAACCTGCCATGTGCTATG36943830TTCCCAAAAGTTGATGCTGATACGATTGGT36949830ACCATCACCGGGAATAACCCCCGCGGT36963030AACAATCAGCCGTGAGGGAATACCCCCGCGT36964030GGCCATATTATCCAGATATGCACCGACGT36969630CGTGGAAAAGAATATAAAATAATAACTACT36969630GGACATATGCTCGCTCTTTCTCGAGTACTC36982830GAACATATGCTGCGCTCTTTCTCGAGTACTC36989430AACCTTTCAATAGTAGCAATCGCGAAACCAG37002630GTCCTAGTACTAAATAAAACAAGTCA37002930AACCTTTGAGAATTGACAACCAGAAACG37015830CTTGGAAAAGAATTTAAAATGCAAACGAACG37029030AAACTTTGAGCATTAGCATTAACTATTCT37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37042230TTTGAAAATGCAGGCGCTTCCACTGGAGAGGTT37048830GAAAATGCAGGCGGTTCCACTGGAGAGGTT3705430TAATTCAAAAAAAGGTCTTGGATGAAAGG37055430TTAGTAACGCCCCCGGTAAGCCCACC37055430TTGGTAATGCAGGCGAATGCACGCCGTT37066030ATTGCGCCCAGGGAATGCACGCCGTT37081830TTGCGCCAGGCACAACAGGTGGCTGTCCACCGCGTT37088430TTGCTAATCAAAAGGCG	368970	30			
36910230ICC IACAICAAIAAACCUACIACAGUIGAG36916830CCAAAAGTTCACCGTCATCCGGATGACGGT36923430CTATCTACTTTTGGGAACCCTAATTGGTAC36930030TTTCTTGTTTTGGCAAGCCTACCAACCTCC36936630AACGATGTGGGATTTTCGGACTTATG36943230AAGGGAATAACCTTACCATTCTGTTTATG36943230AAGGGAATAACCTTACCATTCTGTCTTATG36943830TTCCCAAAAGTTGATGCTGATACGATTGGT36956430AACGATCAGCCGTGAGGGAATACGCCGCGT36965030GGGCATATTATGCAGATATCACAGCAACG36966030GGGCATATTATGCAGATATCCAACGAAACG36976230GAACATATGCTGCTCTTTCTCGAGTACTC36989430AAACTTTGAGTGAAAAAAATAATAACTACT36989430AACTTTGAGTACAAATAAAAAAACAACTCA36996030ACCTTTCAATAGTAGCATCGGGCAAACCAG37002630GTCTCTAGTTAACTAAAAAAACAAGCA37015830CTTGGAAAAGAATTATAAAATGCGAAAGTA37029030AAACTTTGAGCATCAGCAAAGCAAGTA37035630ATACTTTAGCAAATGAAAAAAAAAAAAAAAAAAAAAAAA	369036	30			
36916830CCAAAAG II CAAG ICA ICCAG ICA ICCGAG IAGACGI36923430CTATCTACTTTTGGG AACCCTAATTGGTAC36930030TTTCTTCTGTTTGCCATGTCCAAACCTCC36936630AACAATGTGTGATTTTCGGACTTAGTCCC36949830TTCCCAAAACTTGCTGATACGATTGGT36956430AACAATCAGCCGTGAGGGAATACGCCGCGT36963030AGGTTAATGATGAAAAAATAATAACTACT36969630GGGCATATTATGCAGAACGAACGAACG36969630GACATATGCCCCCTTTCTCAGGAACGAACG36982830GAACATTGCCCCCTTTCCTCGAGTACTC36989430AACATTGCCCGCTCTTTCCTGAGTACTC36989430AACTTTGAGGTACTAAATAAACAAGTCA37002630GCTCTAGTTACTTAGCACACGAAACCG37002830CTTGGAAAAGAATTTATAAAATGCGAAACG37015830CTTGGAAAAGAATTTATAAAATGCGAAACTA37029030AAACTTTGCGTAAATAAAACAAGTCA3702430ATACTTTGCGTAAAATAAAACAAGTCA37035630ATATCTATGTGCATAAATAAAACAAGTCA37042230TTTGCAATTGCCAACGAAAGG37042330AAACTTCCCAAATGCAATTGCTAAAAAAAAAAAAAAAAA	369102	30			
36923430CHARCHACCCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACCHARACHAR	369168	30			
36930030ITICITICITICITICGATGICCAAAACCICC36936630AACAATGGTGATTTTTCGGACTTAGTCCC36943230AACGGAATAACCTTACCATTCTGTCTTATG36949830TTCCCAAAAGTGATGATGATGCGGACTACGATGGT36956430AGGTAAATAATAACAGCCGTGAAGGAAAACAATAATAACTACT36969630GGGCATATTATGCAGAATAGCAGAAACG36976230CTTGGAAAAGAATTATAAAACGGAAGTT36982830GAACATATGCTGCTCTTTCTCGAGTACTC36989430AACCTTTGAGGTACTAAAAAAACAAGTCA36996030ACCTTTCAATGAGCACACGGGCAAACCAG37002630GTCTCTAGTTACTTAGCAGATATGCAACGAAACG37012630CTTGGAAAAGAATTGCAACGAAACG37015830CTTGGAAAAGAATTTAAAAATGCGAAAGTT37022430ATGCGATATATGCAGATATGCAACGACAA37035630ATATCTTTTGTCGTTAAAGTAAAAAAAAA37042230TTTGAAATTCCCAAATGCAATGCAATGTTTC37048830GAAAATCCAGGGGTTCCACTGGAGAGGT37055430TAATTCAAAAAAAGGTTCTGGTTGAAAGG37062030AACCCGCCCCGGGAATGCACCGGCCCTT37081830CTCGCTCGAGCACAACGGCGCTCCACC37081830TTTCCAGCTAAGCAACAGGTGCACCC37081830TTTCCAGCTAGAGCACACAAGTTAAGGG	369234	30			
36936030AACAAGATAACCTTACCATTAGTAGTCCCC36943230AAGGAATAACCTTACCATTCGTCTTATG36949830TTCCCAAAAGTTGATGCTGATACGATTGGT36956430AACAATCAGCCGTGAGGGAATACGCCGCGT36963030AGGTTAATGATGAAAAAAATAATAACTACT36969630GGCCATATTATGCAGATATGCAACGAAACG36976230CTTGGAAAAGGATATGCAACGAAGTAC36989430AAACTTTGAGGTACTAAATAAAACAAGTCA36989430AAACTTTGAGGTACTAAATAAAACAAGTCA36996030ACCTTTCAATAGCAGCAGCAACCAG37002630GCCTGTAGTAATTATAAAATCCGAAACG37015830CTTGGAAAAGAATTTATAAAATGCGAAAGTA37022430ATGCGATATATCATGAGTACTCAATTCT37029030AAACTTTGAGGTACTAAATAAAAAAAAAAAA37042230TTTGAAATGCAGGCGGTCCAACTGAGATTT37048830GAAAATGCAGGCGGTTCCAATGGAGATT37055430TAACTTTGAGTAAAAAAAGGTCTTTGGTTAAAAGG37055230ATGCGCCCCGCGGAAACCGCCCGTT37081830CTCGCTCGAGCACAACGGCGCTCCACC37083430TTCCCACAAAGGCAATTGCAACGCCACC37084430TTCCAACTAGCACACAGGTGCCACC37084430TTCCCACAAAGGCAACAAGTTAAAGTACCACC	369300	30			
36943230AAGGGAAIAACCI IACCAITGCIGITAIG36949830TTCCCAAAAGTIGATGCGATACGATTGGT36956430AACAATCACCCGTGAGGGAATACCCCGCGT36963030AGGTTAATGATGAAAAAAATAATAACTACT36969630GGCCATATTATGCAGATATGCAACGAAACG36976230CTTGGAAAAGAATTTATAAAATGCGAACTA36982830GAACATATGCTCGCTCTTTCTCGAGTACTC36989430AACCTTTCAAGGTACTAAATAAAACAAGTCA36996030ACCTTTCAATAGTAGCATCGGGCAAACCAG37002630GGCCATATTATCAGGAACTGCAACGAAACG37009230GGGCATATTATGCAGATATGCAACGAAACG37015830CTTGGAAAAGAATTTATAAAATGCGAACGTA37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37035630ATATCTTTGCGTTAAAGTACAATGAAAAAA37042230TTTGAAATTCCACAAAGTAATTTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGTT37055430AACCCCCCCCTGCGGAATTGCACGGCCCGTT37062030AGCCCCCCCTGCGGAATTGCACGGCCCCGTT37068630ATTGAGCGTCAAGCACCACCGGTAGCCCACC37075230TTGGTTATTCGACTTTGAGGTGCCACC37081830CTCCGCTCGAGCACAACAGGTGGCTGTCCAC37084430TTTCCAGCTAAGCACAAAAGGGGTTTATAGGG	309300	30			
36949630FITCCAAAAGTIGATACGATGGCIGATACGATGGT36956430AACAATCAGCCGTGAGGGAATACGCCGCGT36963030GGGCATATTATGAAAAAAATAATAACGCCGCGT36969630GGGCATATTATGCAGATATGCAACGAAACG36976230CTTGGAAAAGAATTTATAAAATGCGAAGTT36982830GAACATATGCTCGCTCTTTCTCGAGTACTC36989430AAACTTTGAGGTACTAAATAAAACAAGTCA36996030ACCTTTCAATAGTAGCATCGGGCAAACCAG37002630GTCTCTAGTTACTTAAAACTGCAACGAAACG37009230GGGCATATTATGCAGATATGCAACGAAACG37015830CTTGGAAAAGAATTTATAAAATGCGAAGTT37022430ATGCGATATATCTATGACTTTACCTATTCT37029030AAACTTTTGGTTAAAATTGCAAAGAAGTCA37035630GAAAATCCTGCTAAATTCAAATAAAAAAGAGTCA37042230TATCCTTTGGTAAAGTTAGAAAGG37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAAATGCCACGGCCCGTT37068630ATTGGTATATCCAACGACCCCGGTAAGCCCACC37075230TTGGTATATCGACTTTGATTGAGTACCAACGGGGCTGCCAC37081830CTCCGCTCGAGCACAACAGGTGCCCACC37084430TTTCCAACTGAGCACTCAAAGTTATAGGG	369432	30			
369:0430AACAAACAGCGGGGAAAAAGGAAAACG36963030AGGTTAATGATGAAAAAAAAAAAAACAACTACT36969630GGGCATATTATGCAAGAAACGAACG36976230CTTGGAAAAGAATTTATAAAAATGCGAAGTA36988430AAACTTTGAGGTACTAAAAAAAAAAAAACAAGTCA36989430AAACTTTCAATAGTAGCATCGGGCAAACCAG36996030ACCTTTCAATAGTAGCATCGGGCAAACCAG37002630GTCTCTAGTTACTTTACGTATAAAACCAGG37015830CTTGGAAAAGAATTTATGCAACGAAACG37029030ATGCGATATATCTATGACGTACTAAAAAAAA37029030AAACTTTGAGGTACTAAAATAAAACAAGTCA37042230TTTGAAATTCCCCAAATGTCAATGTAAAAAA37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AACCCGCCCTGCGGAATTGCACGGACGATT37068630ATGCGCTCAAGCACCCCGCTAAGCCCACC37075230TTGGTAATTCGAAGCACCCGGTAACCACC37088430TTTCCAAGCACCAAGGTGTGCCAC37088430TTTCCAAGCACCAAAGTTTAAGGG	369498	30 20			
309000300AGGCATATIGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAA	369304	30			
3697630CGGGCATATTATAGCAGATAGCAACGAAAGG36976230CTTGGAAAAGAATTTATAAAATGCGAAGTT36982830GAACATATGCTCGCTCTTTCTCGAGTACTC36989430AAACTTTGAGGTACTAAATAAAACAAGTCA36996030ACCTTTCAATAGTAGCATCGGGCAAACCAG37002630GTCTCTAGTTACTTACGTATAAACGAACG37015830CTTGGAAAAGAATTTATAAAATGCGAAAGTCA37022430ATGCGATATATCTATGACTTACCTATTCT37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37035630ATATCTTTGTCGTTAAAGTAGAAAAAA37042230TTTGAAATTCCCCAAATGTCAATTGTTTC37042230ATATCTTTGTCGTTAAAGTAGAAAAAA37042230ATATCCTTTGTCGTTAAAGTAGAAAAAA37042230TATGCAGGCGGCTTCCACTGGAGAGGTT3705430GAAAATGCAGGCGGTTCCACTGGAGAGGGT37066630ATGCCGCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTATATCGACTTTGATTGAGCTATC37081830TTTCCAGCTAGAGCATCAAAGTTATAGGG	260606	30			
3070230CHROMARAGATHATAAATGCGAAGT36982830AAACTTTGAGGTACTCCGCTCTTTCTCGAGTACTC36989430AAACTTTGAGGTACTAAATAAAACAAGTCA36996030ACCTTTCAATAGTAGCATCGGGCAAACCAG37002630GTCTCTAGTTACTTTACGTATAAACTCTAT37009230GGGCATATTATGCAGATATGCAACGAAAGG37015830CTTGGAAAAGAATTTATAAAATGCGAAGTT37022430ATGCGATATATCTATGACTTTACCTATTCT37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37035630TTTGAAATTCCCCAAATGTCAATGTAAAAAA37042230TTTGAAATTCCCCAAATGTCAATGTTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGTT37055430AAGCCCGCCCTGCGGAATTGCACGGCCGTT37068630ATGCGCCACACCCGGTAAAGCACCCCGTT37068630ATGGGTCAAGCACCCGGTAAGCCCACC37081830CTCGGTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGGTTAAGGG	369090	30			
360360GAACATATGGGTGGCTATATAGAGTGG36989430AAACTTTGAGGGTACTAAATAAAAACAAGTCA369896030ACCTTTCAATAGTAGCATCGGGCAAACCAG37002630GTCTCTAGTTACTTTACGTATAAAACTCTAT37009230GGGCATATTATGCAGATATGCAACGAAACG37015830CTTGGAAAAGAATTTATAAAATGCGAAGTT37022430ATGCGATATATCTATGACTTACCTATTCT37029030AAACTTTGAGGTACTAAAAAACAAGTCA37035630ATATCTTTGTGGTTAAAGTAGTAAAAAA37042230TTTGAAATTCCCAAATGCAAATGATTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGTT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37068630ATGCGGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTATAGGG	369828	30	CAACATATCCTCCCTCTTTCTCCACTACTC		
3696030AAACTITICAATAGTAGCATCGGGCAAACCAG36996030GTCTTCAATAGTAGCATCGGGCAAACCAG37002630GGCCATATTATGCAGATATGCAACGAAACG37009230CTTGGAAAAGAATTTATAAAATGCGAAGTT37022430ATGCGATATATCATGACTTTACCTATTCT37029030AAACTTTGAGGTACTAAAGAACAAGTCA37035630TTTGAAATTCCCCAAATGTCAATGATAAAAAA37042230TTTGAAATTCCCCAAATGTCAATGTTTTC37042230TAATCCAAAGGCGGGTTCCACTGGAGAGGTT37055430GAAAATGCAGGCGGGTTCCACTGGAGAGGGT37062030AACCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTATAGGG	369894	30			
3070030ACCTITCAGTACATICGAGCAAACCAAG37002630GGGCATATTATGCAGATATGCAACGAAACG37009230GGGCATATTATGCAGATATGCAACGAAACG37015830CTTGGAAAAGAATTTATAAAATGCGAAGTT37022430ATGCGATATATCTATGACTTTACCTATTCT37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37035630TTTGAAATTCCCCAAATGTCAATTGTTTTC37042230TTTGAAATTCCCCAAATGTCAATTGTTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGTT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37068630ATGCGCCCCGGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	369960	30			
37002030GGGCATATTATGCAGATATGCAACGAAACG37009230CTTGGAAAAGAATTATGCAACGAAACG37015830CTTGGAAAAGAATTTATAAAATGCGAAGTT37022430ATGCGATATATCTATGACTTTACCTATTCT37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37035630ATATCTTTTGTCGTTAAAGTTAGTAAAAAA37042230TTTGAAATTCCCCAAATGTCAATTGTTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGGTT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370026	30	CTCTCTACTTACTTTACCTATA A ACTCTAT		
37015830CTTGGAAAAGAATTTATAAAATGCGAAGTT37022430ATGCGATATATCTATGACTTTACCTATTCT37029030AAACTTTGAGGTACTAAATAAAACAAGTCA37035630ATATCTTTTGTCGTTAAAGTTAGTAAAAAA37042230TTTGAAATTCCCCAAATGTCAATTGTTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGGT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370020	30	GGCATATTATGCAGATATGCAACGAAACG		
37010030ATGCGATATATCTATGACTTTACCTATGT37022430AAACTTTGAGGTACTAAATAAAAACAAGTCA37029030AAACTTTGAGGTACTAAATAAAAACAAGTCA37035630ATATCTTTGTCGTTAAAGTTAGTAAAAAA37042230TTTGAAATTCCCCAAATGTCAATTGTTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGGTT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370158	30	CTTGGAAAAGAATTTATAAAATGCGAAGTT		
37022030AAACTTTGAGGTACTAAATAAAAACAAGTCA37035630ATATCTTTGTCGTTAAAGTTAGTAAAAAA37042230TTTGAAATTCCCCAAATGTCAATTGTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGGTT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370224	30	ATGCGATATATCTATGACTTTACCTATTCT		
37035630ATATCTTTGTCGTTAAAGTTAGTAAAAAA37035630TTTGAAATTCCCCAAATGTCAATTGTTTC37042230TTTGAAATCCCCAAATGTCACTGGAGAGGGTT37048830GAAAATGCAGGCGGTTCCACTGGAGAGGGTT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACCGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370290	30	ΑΑΑΓΤΤΓΓΑΩ		
37042230TTTGAAATTCCCCAAATGTCAATTGTTTTC37048830GAAAATGCAGGCGGTTCCACTGGAGAGGTT37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370356	30	ATATCTTTTGTCGTTAAAGTTAGTAAAAAA		
37048830GAAAATGCAGGCGGTTCCACTGGAGAGGGTT3705430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370422	30	TTTGAAATTCCCCAAATGTCAATTGTTTTC		
37055430TAATTCAAAAAAAGGTCTTGGTTTGAAAGG37062030AGCCCGCCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370488	30	GAAAATGCAGGCGGTTCCACTGGAGAGGTT		
37062030AGCCCGCCTGCGGAATTGCACGGCCCGTT37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGATTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370554	30	TAATTCAAAAAAGGTCTTGGTTTGAAAGG		
37068630ATTGAGCGTCAAGCACCCGGTAAGCCCACC37075230TTGGTTATTCGACTTTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370620	30	AGCCCGCCTGCGGAATTGCACGGCCCGTT		
37075230TTGGTTATTCGACTTTGAGCTATC37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370686	30	ATTGAGCGTCAAGCACCCGGTAAGCCCACC		
37081830CTCGCTCGAGCACAACAGGTGGCTGTCCAC37088430TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370752	30	TTGGTTATTCGACTTTTGATTTGAGCTATC		
370884 30 TTTCCAGCTAGAGCATCAAAGTTTATAGGG	370818	30	CTCGCTCGAGCACAACAGGTGGCTGTCCAC		
	370884	30	TTTCCAGCTAGAGCATCAAAGTTTATAGGG		

**Table 2.** The CRISPR array identified with sequences of *T. denticola*.



**Figure 3.** BLAST results similar to those of the pectobacterium phage with an e value of 0.001%; 100% query cover 60%.

Table 3. Identified CRISPR arrays with spacers of Tanerella forsythia.

SPACER ID	POSITION	SEQUENCE		
2508368	36	ACAGAAACTTCTTTTCCTGCAAGATTAAATAATACA		
2508434	37	GGTATGTATAAATCTACACGTCTTGGGTTTTCTAATA		
2508501	40	AAAATTATCTTTGATAACTTTAAGAATCTTTTTGTCTTCT		
2508571	40	CCATTACTGCGCGGGCGGCGATGCAGGAGAACCCGGACCG		
2508641	39	TAAAAGGTTAAAAGTTAAATGAAGAAAGACATAAAACGA		
2508710	36	GTCTTTGGAGGCCTTTACTCTTTTAAAAATGCCCGA		
2508776	34	TCTCTTGTGTAGTTATAACACACAATTGAGTCAT		
2508840	35	CCTCTTCGTAATACGGCTCTATATCGAGCTCTCTG		
2508905	38	CCCCGAAGGCGCGGCCGTTCCACTCTAAGGTGTCGACT		
2508973	36	GTTAGATGATAACTTCCGTCTTCTCCGAGCATCATC		
2509039	36	TTCACGGGGGTAAAGCCCGCCCTTACGGGGAACTA		
2509105	37	AAAAGATTTGCTATATTGTGAAAAATTTAAAAAAAG		
2509172	39	TTCACGGGGGTAAAGCCCGCCCTTACGGGGAACTACTC		
2509241	38	TGCAGCCGTTGTCTCGCAAAAAGCAGCGCCTTCTAAAA		
2509309	36	CCTGCCCTGCGTGGGGGGCTTTGGTGTGCGCAGCCTG		
2509375	37	CTTCAAAAGCGGCAGCGCGCTTTTTGAGAAGGCGCTG		
2509442	37	TTTGTGTTATACGGAGATTATACACGGTGGGTGTGGG		
2509509	37	TGCGCGCATACGTTCTTCTACGTACGCCTTTTTGGTT		
2509576	38	AAAAGATTTGCTATTTGTAATAGCAAAGAACTATGAA		
2509644	35	TAAGAGGTTATCTCCGTCCGCACGAGTTCGGACGT		
2509709	36	CATACGTTGCACGAATGTCCGCATCAAAAGCGGCAG		
2509775	36	GTTATTCTCCCAATCACCGAACCTGTCAACAAAAGA		
2509841	38	CTATAACTGCAGCCCGCACGGCTGCACGGTTGGCGTAC		

SPACER ID	POSITION	SEQUENCE			
2509909	37	GTCTGGTCGTTTTTTCTGTATTGGGGGGGGGGGCGCATT			
2509976	36	ACGTTAACACGTGCACCGCCAACGCGGGCAGAAAAG			
2510042	35	CGTATCTGAAACGTCCGAATTGGTGCGGACGGAGG			
2510107	36	GCCACCACCGACCCTGCACGGGCGGTTTGGCTCGCG			
2510173	37	ACATAGCAAACCTTGTTGGATAATATTCTTCGTCTAT			
2510240	36	GAACAATAACAAACAGGATTGGGCGATCATACTTTT			
2510306	36	CTCGGATTTATCCATTACGTCGGCCGCATCGATTTG			
2510372	38	CCAACGCGGGGAAGAAGGCGGATGAGGTTTCGGAACAA			
2510440	37	CGCCGAGAAGGTCTCCGTCCTCGTCAAAAACAGACTC			
2510507	37	GCTCGGGGATCGATATAATAGAATGGTAGAGGAGTGG			
2510574	38	TACGGGGTCACCTCCGTTCGCATGAGTTTGGACGTTTC			
2510642	36	CATATAACAGTGCGACGAGACAGGCTGCGCACACCA			
2510708	37	GGGGTAAATCATGTAAAACGAACAATTTTAGAATATA			
2510775	36	CGTAGCGCTATTATGCGCGGGTATGGCGCGGCTCTGC			
2510841	36	CTTGCAGTAGGGGCAGTGAAAATTACACGACCCGAA			
2510907	39	CTATTACCGCCGCCCGCACGGGCGCACGGCTGGCGTACG			
2510976	38	GAATAAAAATCAGCGTAAAAACATATATGCTTACACGA			
2511044	38	TTTTTTGTGGCCGTTTCGGCCTTATCTGTATCTACTGT			
2511112	36	AGCCGACGCGATATATTCGTGCCCGGCAAACAAATC			
2511178	36	CGCCGTACACGTGGTGATATTTATGAAACAAATTTT			
2511244	36	GATATATTTTTCTAAACATTTTCATGACTTTATCGG			
2511310	36	TGATAACTTTTAAAGTCTTTTTATCCTCTGCATTCA			
2511376	37	ATTGTACACCTTGTCTACTTGCTCATCAGGGAAATTC			
2511443	37	GTAATCCTCACACTGCCGATTATGGCTGAGTAATCTA			
2511510	36	AGTTGATGCAAAAGGTTTATCGAAACACCTTCCTTC			
2511576	37	CCTTATCTCAATGGTTTAAGGAAGACGTTGCGCGCAT			
2511643	36	AGCCTATCTAAACGCCAAATAACTAATTTGTCTCCT			
2511709	37	TCGGTTTGGTGATAGTTGGTAACAAGATTTTAAAACA			
2511776	34	ATTACATCAGGATAGATGGTGTGGCGTATGAAAA			
2511840	37	GAAGCTGGTAACCGCGGATGTCGCCTTCGGCCTCGAA			
2511907	38	CCGCTACCCAAAAAACCCTCGACAAAATGCCCGAAATC			
2511975	37	CACTGGGGCACATAGTCCTGATCTCCTCGGCGAGTAT			
2512042	37	GAAAAGTGCTAATGGTAAAGTTAAACTTTCTCCTTTA			
2512109	35	TTAGATGATCTTGCGGACTACAGTGCCGAATTTGA			
2512174	36	GCAAACGGGGCCTTTTGTGCCATGTCACCAGTGCGG			
2512240	37	AGAGACGTTCCTCCTGCGTTGTACAAGATACTCTGTA			
2512307	34	AATTTAAGTGATGATGTAAAAACTGGTATTAATA			
2512371	35	TTAGATGATCTTGCGGACTACAGTGCCGAATTTGA			
2512436	36	TTCGCAATCGCTTCGACGGCAAAAGTGCCGCCCAGG			
2512502	35	TGTTTATGTACGACTCTTTCAGTTTTAACTGCTTT			
2512567	39	GATATAATGCTTACAATCGAGGCATTTAATAACATTAAA			
2512636	39	CAATGTGTACGTTGTCATTGCCGGTATATTAGGATTCTT			
2512705	35	AAGTGCCTTTGCCGTTAATTTTGTCAATGAGTTTC			
2512770	38	GCGAAATCCTGCTCGTGCAGGGTGGAAGCAAGAATATC			
2512838	38	AGCGTCATAGCATTCACACCGGCAGCACCAGTTACGAA			
2512906	37	GTCGGACACGATGGGCAGAAATTCTTCTTCGACCATG			
2512973	35	CCGTTAAATTGTCTGGCAAGGACGTGACGCCGGTA			
2513038	36	CGTTGTAGAAATCGTCTTCGTTGATAACAAGTGTTA			
2513104	37	GCTTCTTCGATTCTTCTTTTACCTCTCCGGTTTCCGT			

*Tannerella forsythia*: complete genome length, 3405521; CRISPR rank in the sequence, 4; Crispr\_begin\_position: 2508368  $\rightarrow$  Crispr\_end\_position: 2513104.

Machine learning is a rapidly developing technology today. It is used for image recognition, speech recognition, traffic predictions and product recommendation, and now, this technology is developing in medicine. CRISPR-Cas systems identify and destroy intruder DNA by matching spacers to viral protospacers [40]. These CRISPR array spacers can also be predicted. Modeling CRISPR sequences reveals accurate predictions of 100 percent for Random Forest, Neural Networks and Support Vector Machines.

(Table 4, Figure 4) This prediction can help identify and classify red complex spacers in future studies.

Model	AUC	CA	F1	Precision	Recall	LogLoss	Specificity
Random Forest	1.000	1.000	1.000	1.000	1.000	0.097	1.000
Neural Network	1.000	1.000	1.000	1.000	1.000	0.021	1.000
SVM	0.938	1.000	1.000	1.000	1.000	0.387	1.000

Table 4. Accuracy of nearly 100% in all algorithms.



Figure 4. ROC curve of CRISPR-predicted sequences.

The application of phage therapy in dentistry is still in its infancy and requires exploration. Bacteriophages are against many bacteria that are present in biofilms [41]. Understanding bacteria–bacteriophage interactions and specificity is the first step in expanding its applications in dentistry. In this study, we aimed to identify the bacteriophages interacting with red complex bacteria in periodontitis. The formulation of phage-based cocktails derived from different phages could surpass bacterial resistance against a single bacteriophage. Phage-based products can be developed based on the virulence of the phage against the bacteria and the profile of the CRISPR-Cas of the bacteria [42,43]. In that way, our study is the first step toward understanding the therapeutic possibilities to treat periodontitis. Our results show that the CRISPR-Cas of red complex bacteria target phages such as cellulophages (*P. gingivalis*), bacillus phages (*T. denticola*), and pectophages (*Tannerella*) in periodontitis. Our results suggest that red complex bacteria are resistant to phages such as cellulophages, bacillus phages, and pectophages in periodontitis. There are currently only a few clinical uses for CRISPR in periodontics. However, the possible clinical use of CRISPR is to target the periodontal biofilm and create new methods for lowering or

getting rid of periodontal infections. Additionally, CRISPR can change the transcriptome and gene expression of genes that contribute to the development of periodontitis.

#### 5. Conclusions

The current study highlights that viral phage communities cannot modify sub-gingival bacterial environments, as they have acquired immune mechanisms via CRISPR-Cas to kill the virus and competitively infect periodontal pockets. Future research can be simplified, reducing time and effort by using predictive modeling of red-complex-based spacer analyses.

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

Funding: This research received no external funding.

Institutional Review Board Statement: IHEC/SDC/FACULTY/PERIO/020.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

#### References

- 1. Sela, M.N. Role of Treponema denticola in periodontal diseases. Crit. Rev. Oral Biol. Med. 2001, 12, 399–413. [CrossRef] [PubMed]
- Könönen, E.; Gursoy, M.; Gursoy, U.K. Periodontitis: A multifaceted disease of tooth-supporting tissues. J. Clin. Med. 2019, 8, 1135. [CrossRef] [PubMed]
- 3. Bibi, T.; Khurshid, Z.; Rehman, A.; Imran, E.; Srivastava, K.C.; Shrivastava, D. Gingival crevicular fluid (GCF): A diagnostic tool for the detection of periodontal health and diseases. *Molecules* **2021**, *26*, 1208. [CrossRef] [PubMed]
- Shrivastava, D.; Srivastava, K.C.; Ganji, K.K.; Alam, M.K.; Al Zoubi, I.; Sghaireen, M.G. Quantitative assessment of gingival inflammation in patients undergoing nonsurgical periodontal therapy using photometric CIELab analysis. *BioMed Res. Int.* 2021, 2021, 6615603. [CrossRef]
- Ximénez-Fyvie, L.A.; Haffajee, A.D.; Socransky, S.S. Microbial composition of supra-and subgingival plaque in subjects with adult periodontitis. J. Clin. Periodontol. 2000, 27, 722–732. [CrossRef] [PubMed]
- Shrivastava, D.; Srivastava, K.C.; Dayakara, J.K.; Sghaireen, M.G.; Gudipaneni, R.K.; Al-Johani, K.; Baig, M.N.; Khurshid, Z. BactericidalActivity of Crevicular Polymorphonuclear Neutrophils in Chronic Periodontitis Patients and Healthy Subjects under the Influence of Areca Nut Extract: An In Vitro Study. *Appl. Sci.* 2020, *10*, 5008. [CrossRef]
- Assiri, K.I.; Sandeepa, N.C.; Asiri, R.S.; Mulawi, S.A.; Najmi, S.M.; Srivastava, K.C. Assessment of Oral–Systemic Disease Association amongst Dental Patients: A Retrospective Panoramic Radiographic Study. J. Contemp. Dent. Pract. 2020, 21, 748–755. [PubMed]
- Shrivastava, D.; Natoli, V.; Srivastava, K.C.; Alzoubi, I.A.; Nagy, A.I.; Hamza, M.O.; Al-Johani, K.; Alam, M.K.; Khurshid, Z. Novel approach to dental biofilm management through guided biofilm therapy (Gbt): A review. *Microorganisms* 2021, 9, 1966. [CrossRef]
- 9. Lin, D.M.; Koskella, B.; Lin, H.C. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World J. Gastrointest. Pharmacol. Ther.* **2017**, *8*, 162. [CrossRef]
- Gummalla, V.S.; Zhang, Y.; Liao, Y.-T.; Wu, V.C. The Role of Temperate Phages in Bacterial Pathogenicity. *Microorganisms* 2023, 11, 541. [CrossRef]
- 11. Pinto, G.; Silva, M.D.; Peddey, M.; Sillankorva, S.; Azeredo, J. The role of bacteriophages in periodontal health and disease. *Future Microbiol.* **2016**, *11*, 1359–1369. [CrossRef]
- 12. Egido, J.E.; Costa, A.R.; Aparicio-Maldonado, C.; Haas, P.-J.; Brouns, S.J. Mechanisms and clinical importance of bacteriophage resistance. *FEMS Microbiol. Rev.* **2022**, *46*, fuab048. [CrossRef]
- 13. Yeh, T.-K.; Jean, S.-S.; Lee, Y.-L.; Lu, M.-C.; Ko, W.-C.; Lin, H.-J.; Liu, P.-Y.; Hsueh, P.-R. Bacteriophages and phage-delivered CRISPR-Cas system as antibacterial therapy. *Int. J. Antimicrob. Agents* **2022**, *59*, 106475. [CrossRef] [PubMed]

- Soto-Perez, P.; Bisanz, J.E.; Berry, J.D.; Lam, K.N.; Bondy-Denomy, J.; Turnbaugh, P.J. CRISPR-Cas system of a prevalent human gut bacterium reveals hyper-targeting against phages in a human virome catalog. *Cell Host Microbe* 2019, 26, 325–335.e325. [CrossRef] [PubMed]
- 15. Johnson, C.M.; Grossman, A.D. Integrative and conjugative elements (ICEs): What they do and how they work. *Annu. Rev. Genet.* **2015**, *49*, 577–601. [CrossRef] [PubMed]
- 16. Alseth, E.O.; Pursey, E.; Luján, A.M.; McLeod, I.; Rollie, C.; Westra, E.R. Bacterial biodiversity drives the evolution of CRISPRbased phage resistance. *Nature* **2019**, *574*, 549–552. [CrossRef]
- 17. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **2007**, *315*, 1709–1712. [CrossRef]
- 18. Jansen, R.; van Embden, J.D.A.; Gaastra, W.; Schouls, L.M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **2002**, *43*, 1565–1575. [CrossRef]
- 19. Santiago-Rodriguez, T.M.; Naidu, M.; Abeles, S.R.; Boehm, T.K.; Ly, M.; Pride, D.T. Transcriptome analysis of bacteriophage communities in periodontal health and disease. *BMC Genom.* **2015**, *16*, 549. [CrossRef]
- Jore, M.M.; Brouns, S.J.; van der Oost, J. RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. *Cold Spring Harb. Perspect. Biol.* 2012, 4, a003657. [CrossRef]
- 21. Coordinators, N.R. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **2012**, *41*, D8–D20. [CrossRef] [PubMed]
- 22. Gussow, A.B.; Park, A.E.; Borges, A.L.; Shmakov, S.A.; Makarova, K.S.; Wolf, Y.I.; Bondy-Denomy, J.; Koonin, E.V. Machinelearning approach expands the repertoire of anti-CRISPR protein families. *Nat. Commun.* 2020, *11*, 3784. [CrossRef] [PubMed]
- 23. Konstantakos, V.; Nentidis, A.; Krithara, A.; Paliouras, G. CRISPR–Cas9 gRNA efficiency prediction: An overview of predictive tools and the role of deep learning. *Nucleic Acids Res.* **2022**, *50*, 3616–3637. [CrossRef]
- 24. Mohanty, R.; Asopa, S.J.; Joseph, M.D.; Singh, B.; Rajguru, J.P.; Saidath, K.; Sharma, U. Red complex: Polymicrobial conglomerate in oral flora: A review. J. Fam. Med. Prim. Care 2019, 8, 3480.
- Butera, A.; Gallo, S.; Pascadopoli, M.; Taccardi, D.; Scribante, A. Home oral care of periodontal patients using antimicrobial gel with postbiotics, lactoferrin, and aloe barbadensis leaf juice powder vs. conventional chlorhexidine gel: A split-mouth randomized clinical trial. *Antibiotics* 2022, 11, 118. [CrossRef] [PubMed]
- Gallo, S.; Scribante, A. Ozone therapy in dentistry: From traditional applications towards innovative ones. A review of the literature. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Sanya, China, 28–30 October 2020; p. 012001.
- 27. Kim, J.; Amar, S. Periodontal disease and systemic conditions: A bidirectional relationship. Odontology 2006, 94, 10–21. [CrossRef]
- Martínez-García, M.; Hernández-Lemus, E. Periodontal inflammation and systemic diseases: An overview. *Front. Physiol.* 2021, 12, 709438.
- 29. Hajishengallis, G. Periodontitis: From microbial immune subversion to systemic inflammation. *Nat. Rev. Immunol.* **2015**, 15, 30–44.
- 30. Radaic, A.; Kapila, Y.L. The oralome and its dysbiosis: New insights into oral microbiome-host interactions. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1335–1360.
- 31. Chen, T.; Olsen, I. Porphyromonas gingivalis and its CRISPR-Cas system. J. Oral Microbiol. 2019, 11, 1638196. [CrossRef]
- 32. Richter, C.; Chang, J.T.; Fineran, P.C. Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems. *Viruses* **2012**, *4*, 2291–2311.
- Hille, F.; Charpentier, E. CRISPR-Cas: Biology, mechanisms and relevance. *Philos. Trans. R. Soc. B Biol. Sci.* 2016, 371, 20150496. [CrossRef]
- 34. Sorek, R.; Kunin, V.; Hugenholtz, P. CRISPR—A widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* **2008**, *6*, 181–186. [CrossRef] [PubMed]
- Ly, M.; Abeles, S.R.; Boehm, T.K.; Robles-Sikisaka, R.; Naidu, M.; Santiago-Rodriguez, T.; Pride, D.T. Altered oral viral ecology in association with periodontal disease. *MBio* 2014, *5*, e01133-14. [CrossRef] [PubMed]
- 36. Szafrański, S.P.; Slots, J.; Stiesch, M. The human oral phageome. Periodontol. 2000 2021, 86, 79–96. [PubMed]
- 37. Shanmugasundaram, S.; Nayak, N.; Puzhankara, L.; Kedlaya, M.N.; Rajagopal, A.; Karmakar, S. Bacteriophages: The dawn of a new era in periodontal microbiology? *Crit. Rev. Microbiol.* **2023**, 1–12. [CrossRef]
- Landsberger, M.; Gandon, S.; Meaden, S.; Rollie, C.; Chevallereau, A.; Chabas, H.; Buckling, A.; Westra, E.R.; van Houte, S. Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. *Cell* 2018, 174, 908–916.e912.
- Marraffini, L.A. CRISPR-Cas immunity against phages: Its effects on the evolution and survival of bacterial pathogens. *PLoS Pathog.* 2013, 9, e1003765. [CrossRef]
- Mitrofanov, A.; Alkhnbashi, O.S.; Shmakov, S.A.; Makarova, K.S.; Koonin, E.V.; Backofen, R. CRISPRidentify: Identification of CRISPR arrays using machine learning approach. *Nucleic Acids Res.* 2021, 49, e20.
- 41. Steier, L.; De Oliveira, S.D.; De Figueiredo, J.A.P. Bacteriophages in dentistry—State of the art and perspectives. *Dent. J.* **2019**, *7*, 6. [CrossRef]

- 42. Abedon, S.T.; Danis-Wlodarczyk, K.M.; Wozniak, D.J. Phage cocktail development for bacteriophage therapy: Toward improving spectrum of activity breadth and depth. *Pharmaceuticals* **2021**, *14*, 1019. [CrossRef] [PubMed]
- Yang, Y.; Shen, W.; Zhong, Q.; Chen, Q.; He, X.; Baker, J.L.; Xiong, K.; Jin, X.; Wang, J.; Hu, F. Development of a bacteriophage cocktail to constrain the emergence of phage-resistant Pseudomonas aeruginosa. *Front. Microbiol.* 2020, 11, 327. [CrossRef] [PubMed]

**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.