Elucidation of Genes Responsible for Quorum Sensing in *Cronobacter sakazakii*

Enrollment Number - 123812

Name of student – Abhay Sharma

Name of supervisor- Dr. Gunjan Goel

Submitted in partial fulfillment of the Degree

5 Year Dual Degree Programme B.Tech – M.Tech

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY,

WAKNAGHAT, SOLAN (H.P.)

CERTIFICATE

This is to certify that the work which is being presented in the thesis titled **"Elucidation of Genes Responsible for Quorum Sensing in *Cronobacter sakazakii*"** in partial fulfillment of the requirements for the award of the degree of Master of Technology in Biotechnology and submitted to the Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by **Abhay Sharma (123812)** during a period from July 2016 to May 2017 under the supervision of **Dr. Gunjan Goel** Associate Professor, Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat.

The above statement made is correct to the best of our knowledge.

Date: - ……

Dr. Gunjan Goel

Associate Professor

Department of Biotechnology and Bioinformatics

JUIT Waknaghat.

ACKNOWLEDGEMENT

I would like to thank my thesis adviser **Dr. Gunjan Goel (Associate Professor)** of the Department of Biotechnology and Bioinformatics at Jaypee University of Information Technology Wakhnaghat, Solan (Himachal Pradesh). The door to Dr. Goel's office was always open whenever I ran into trouble spot or had a question about my research or writing. He consistently allowed this study to be my own work, but steered me in the right the direction whenever he thought I needed it.

I pay my gratitude to **Dr. R.S. Chauhan (Professor)** Head of Department of Biotechnology and Bioinformatics at Jaypee University of Information Technology Wakhnaghat, Solan (Himachal Pradesh). To provide students of Biotechnology and Bioinformatics an opportunity to choose desirable individual study and providing well knowledgeable faculty to guide the students throughout their course.

I would also like to thank **Mr. Sampan Attri (Ph.D Scholar)** of the Department of Biotechnology and Bioinformatics at Jaypee University of Information Technology Wakhnaghat, Solan (Himachal Pradesh) and **Mr. Deepak Sharma (Ph.D Scholar)** of the Department of Biotechnology and Bioinformatics at Jaypee University of Information Technology Wakhnaghat, Solan (Himachal Pradesh). They helped me throughout the Biology work and instructed well to use the laboratory and equipment's without their involvement it won't be easy to achieve results in my work.

Finally, I must express my very proud found gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my period of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Thank you.

Abhay Sharma

(123812)

Department of Biotechnology and Bioinformatics

Jaypee University of Information Technology

Waknaghat, Solan, Himachal Pradesh.

List of tables

Contents

INTRODUCTION

Quorum sensing (QS) is a mechanism which enables microorganism to communicate through chemical hormone-like molecules known as auto-inducers. The external degree of the auto-inducers enhances as a characteristic of growing cellular population. The bacteria detect and respond to the signal, when an auto-inducer concentration reaches a critical threshold. Bacteria use QS to regulate a variety of phenotypes, such as motility, virulence factor production, toxin production, exopolysaccharide production, and biofilm formation, which are essential to build a successful establishment of a symbiotic or pathogenic relationship with their respective eukaryotic hosts. QS changed into first described in the mechanism liable for regulation of bioluminescence in *Vibrio fischeri* and *Vibrio harveyi*[1], and when you consider that then shown to be a considerable mechanism of gene regulation in microorganism. One of the main reasons behind people who are affected with cystic fibrosis is QS in *Pseudomonas aeruginosa*. *Pseudomonas aeruginosa* genome sequence is available in the public domain. QS mechanism in *Pseudomonas aeruginosa* also deals with the expression of various virulence factors[2, 3].

The LuxI/R signaling system

The LuxI/R signaling system was the first one to be described in *Vibrio fischeri*[1]. The luciferase operon in *Vibrio fischeri* is regulated via two proteins, LuxI, that is responsible for the production of the Acyl Homoserine Lactone (AHL) which is an auto-inducer, and LuxR, that is activated through this AHL to promote transcription of the luciferase operon[4].As this explanation, homologs of LuxI–LuxR have been identified in other bacteria, and in all LuxI/R signaling systems, the bacteria produce an AHL, that binds to the LuxR protein and regulates the transcription of genes involved in a variety of phenotypic characteristics. These phenotypic characteristics include the production of antibiotics in *Erwinia*, motility in *Yersinia pseudotuberculosis*, and pathogenesis and biofilm formation in *Pseudomonas aeruginosa*, among others[5]. The AHL syntheses resembles the LuxI-type proteins. AHLs composed of a homoserine lactone ring connected through an amide bond to a acyl chain. These acyl chains differ in carbon wide variety from 4 to 18 and the third position may be changed or may not be modified. The different AHLs will be recognized by different LuxR-type proteins when different the acyl chains are available. To synthesize the homoserine lactone ring, S-adenosyl-methionine (SAM) acts as a substrate, which is used by LuxI-type proteins for AHL synthesis and the acyl chains come from lipid metabolism, which is carried by different acyl-provider proteins. Transcription of their target genes is regulated when AHL binds with the LuxR-type proteins which act as transcription factors. It has been shown that

AHL binding to these proteins balances them; differently, in the absence of signal, they are directed to degradation [6]. The LuxR-type proteins usually recognize a specific AHL. A specific AHL is recognized by the specific LuxR-type proteins, the LuxI/R signaling system has been primarily associated with intraspecies signaling.

The best characterized LuxI/R-type QS systems is in *Pseudomonas aeruginosa*. Several genes which are involved in biofilm formation, persistence within the host and colonization activates the QS in *Pseudomonas aeruginosa*. *Pseudomonas aeruginosa* is an opportunistic pathogen of immunocompromised individuals, including those with burns, human immunodeficiency virus, or cystic fibrosis[9]. *Pseudomonas aeruginosa* causes chronic colonization of the pulmonary airways which is related to the morbidity and motality in case of cystic fibrosis. QS is responosible for the production of virulence factors and biofilm development in *Pseudomonas aeruginosa*. Any change in the QS system disappears *Pseudomonas aeruginosa* virulence in plants and animals and inhibits biofilm formation[10]. The QS system of *Pseudomonas aeruginosa* is very complex and hierarchical. *Pseudomonas aeruginosa* produces two AHLs, N-(3-oxododecanoyl)-L-homoserine lactone (3OC12-HSL) and N-butanoyl-L-homoserine lactone (C4- HSL)[11]. LasR and RlhR transcription factors are activated when these two AHLs binds, respectively[9].

Some bacteria have the ability to disrupt QS signaling by degrading AHL auto-inducers. The soil bacterium *Bacillus* produces a lactonase enzyme that hydrolyzes the lactone ring of AHLs. AHL signaling is hindered by this lactonase enzyme by other bacterial species which makes *Bacillus* to compete in nature. Such mechanism, evidenced by the LuxR homologue TraR in the plant pathogen *Agrobacterium tumefaciens*, is the stability of LuxR-type proteins increases upon auto-inducer binding. In the absence of an auto-inducer the half-life of TraR is of few minutes. However, when AHL is present, the half- life of TraR enhances to over 30 minutes[6]. For folding of the polypeptide, AHL binding is required and indeed radiolabeled TraR was stabilized only, when its cognate AHL was added prior to labeling of the protein. Hence, when the concentration of AHL reaches the threshold (outside as well as inside the cell) can TraR bind it, fold, and start QS cascade. Other important mechanism that prevents "short-circuiting" of LuxI/R systems is active export of AHL signals. When a sufficient concentration of signal has achieved, which indicates high cell density, diffusion into the cell allows export which results in the circuit engagement. Active export to transverse the bacterial membrane is required by AHLs with long acyl side chains[11].

Literature Review

Quorum sensing (QS) was discovered around 25 years ago in oceanic bacterial species showing bioluminescence, *Vibrio fischeri* and *Vibrio harveyi*. In these species, manufacturing of light is handled by various enzymes which are regulated by different operons. The luminescence was able to arise best at high cell-population density in reaction to the buildup secreted signaling molecules referred to as auto-inducer[1]. Until now, there are only few different cases of regulation of expression of gene in bacteria in reaction to cell-cell communication were acknowledged. For example, manufacturing of antibiotic by *Streptomyces* spp.[12], attachment in *Enterococcus faecalis*[13], and maturation of fruiting body in *Myxococcus xanthus*[14] were all recognized to be regulated by cell-cell networking. These networking systems used by bacteria had been considered specific, and bacteria in trendy, as an entire were now believed to apply cell-cell conversation. Rather, the change of synthetic indicators among cells or organisms was assumed to be a eukaryotic characteristic. The cutting-edge advancement inside the field of cell-cell communication in microorganism has demonstrated that most microorganism likely communicate through chemical molecules to coordinate the conduct of the class. Further, now we understand that a great collection of various guidance of chemical signs is occupied, which different microorganism uses chemical sign to contact with their community. It seems clear that the capability to communicate inside and among species is vital for the survival of bacteria and intercommunication in natural habitats.

LUXI/R-TYPE QUORUM SENSING IN GRAM NEGATIVE BACTERIA

In the last decade QS mechanism had been reported in more than 25 species of Gram-negative bacteria. In each case besides the ones of *Vibrio harveyi* the QS systems identified in Gram-negative bacteria resembles QS system of *Vibrio fischeri* which is a symbiotic bacteria. LuxI and LuxR are the homologues of proteins in *Vibrio fischeri* which regulate QS. LuxI-like proteins are responsible for the synthesis of acylated homoserine lactone (AHL) signaling molecule known as auto-inducer, LuxI-like proteins are responsible. With an increase in cell population, there is also increase in the auto-inducer amount. The proteins like LUX-R bind respective auto-inducers which have completed a required concentration, and the target gene transcription is activated by LuxR auto-inducer complexes[4]. To vary cell-population, Gram-negative bacteria couple their gene expression by using QS mechanism. *Pseudomonas aeruginosa*, *Vibrio fischeri*, *Agrobacterium tumefaciens*, and *Erwinia carotovora* QS systems are the first-class understood among these 25 species of microorganisms. In those studies we have been able to find out both, the similarities and the

differences between the regulatory networks. The transformation of LuxI/LuxR circuit to modern networks depicts the variations in LuxI/LuxR-like network, which are adapted by bacterium species present in an environment.

The *Pseudomonas aeruginosa* LasI/LasR-RhlI/RhlR

Virulence System

In the opportunistic human pathogen *Pseudomonas aeruginosa*, QS is regulated by a hierarchical LuxI/LuxR circuit. In *Pseudomonas aeruginosa*, two pairs of LuxI/LuxR homologues, LasI/LasR[15] and RhlI/RhlR[16] are present. Both LasI and RhlI acts as an auto-inducer synthases which catalyzes the formation of the auto-inducers N- (3-oxododecanoyl)-homoserine lactone and N-(butryl)-homoserine lactone[11], respectively. The regulatory circuits act in tandem to govern the expression of a number of *Pseudomonas aeruginosa* virulence elements. *Pseudomonas aeruginosa* QS circuit features as described futher.

At excessive cell population density, LasR binds its specific AHL auto-inducer and collectively they bind at promoter factors straight away preceding the genes coding some of secreted virulence elements that are liable for host tissue destruction all through initiation of the contamination method. To determine pathogenicity encompass LasB depicts elastase; a LasA depicts protease; toxA depicts Exotoxin A; and aprA gene depicts alkaline phosphatase [15]. Opposite to the *Vibrio fischeri* LuxI/LuxR network, LasR which is specific to auto-inducer additionally turns on LasI expression. The expression of LasI generates a feedback loop. The complex of LasR auto-inducer additionally turns on the QS system of *Pseudomonas aeruginosa*[17]. Especially, it triggers the expression of rhlR. The auto-inducer produced via RhlI to which rhlR binds; this formation initiates the gene expression which can underneath the manipulation of the LasI/LasR system. Activation of RhlI also induces an auto regulatory loop similar to the LasI/LasR and LuxI/LuxR. However, the LasR-based auto-inducer, also prevents the binding of the RhlI-based auto-inducer, to its respective regulator RhlR[18]. These two systems provoke their cascades sequentially and in the perfect manner, which is ensured at the stage of manipulation of RhlI/RhlR autoinduction through the LasI/LasR system. A novel, extra auto-inducer has currently been verified to be involved in QS in *Pseudomonas aeruginosa*. As it is not always of the homoserine lactone class, this sign is noteworthy. Relatively, it is 2-heptyl-3-hydroxy-4-quinolone (denoted PQS for *Pseudomonas* Quinolone Sign)[19]. PQS partially controls the expression of LasB which is an

elastase gene with respect to Las and Rhl QS systems. Transcription of RhlI and LasR which is induced by PQS is required for the expression of PQS. This information implies that relation between the Las and Rhl circuits is a PQS. PQS initiates the Rhl cascade by way of allowing the manufacturing of the RhlI-mediated auto-inducer, after the formation of the LasI/LasR system cassette. At present, studies on QS in *Pseudomonas aeruginosa* have found out that QS is important for correct biofilm formation. Significantly, *Pseudomonas aeruginosa* LasI mutants do no longer turn into complete biofilms. Precisely, they end biofilm formation on the microcolony stage[5]. Wild-type biofilm production by using the exogenous addition of the LasI-based AHL may be complemented by these mutants. *Pseudomonas aeruginosa* is the number one pathogen discovered within the lungs of people affected with cystic fibrosis (CF), and when CF sputum samples are examined under microscope, they suggests that *Pseudomonas aeruginosa* present in major amount in biofilms in vivo. Certainly, the LasI and the RhlI-directed auto-inducers were detected in sputum samples of CF patients[9].These facts suggest that biofilm formation through *Pseudomonas aeruginosa* will be vital for lung colonization, and consequently antimicrobial remedies designed to intervene with QS, and by means of analogy with biofilm formation, may be used within the remedy of CF.

Functions of the LuxI and LuxR Family of Proteins

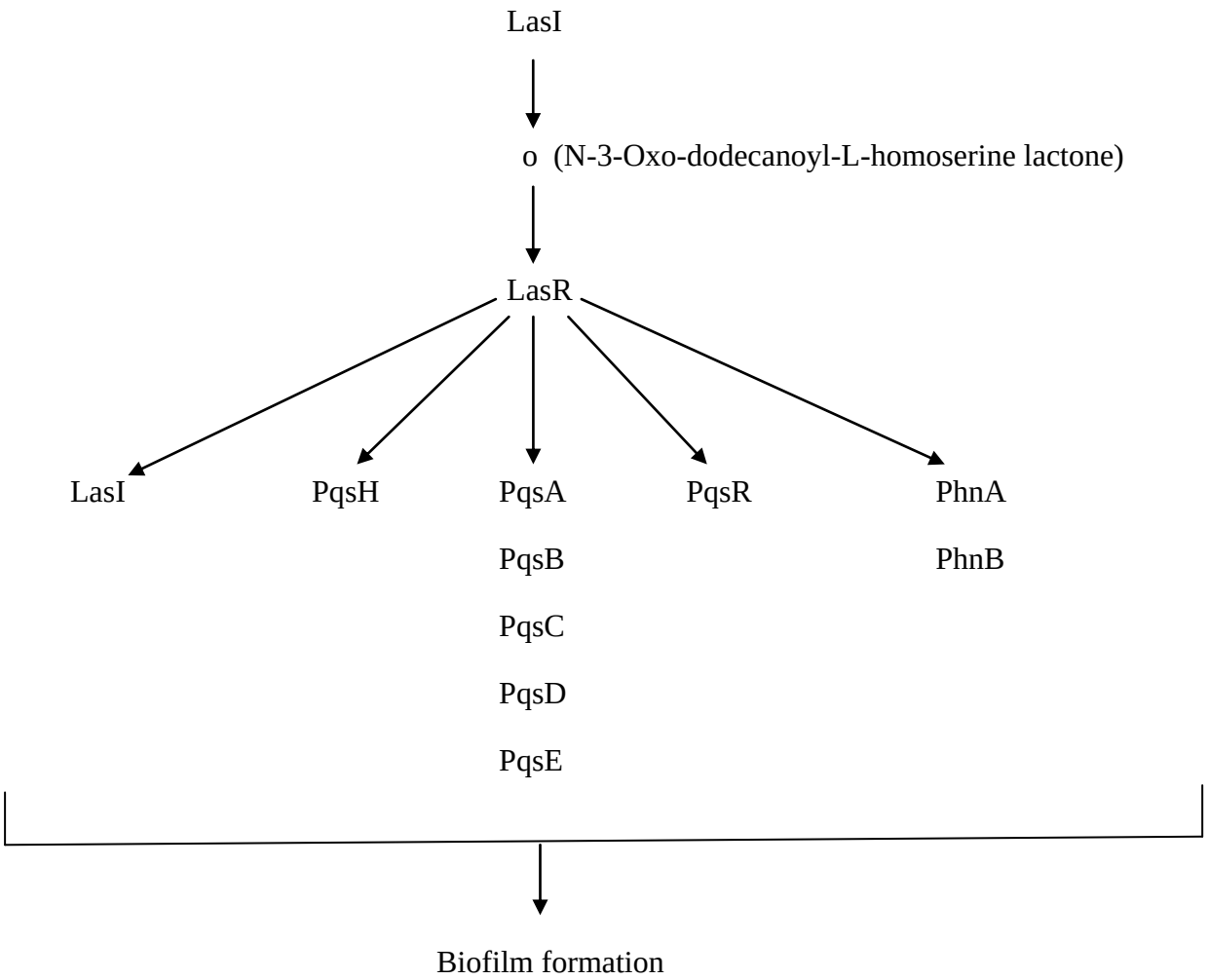
Previously conducted research shows that the fundamental capabilities and bio-chemical nature of motion of the diverse LuxI-like and LuxR-like proteins are equal to the ones of LuxI and LuxR of *Vibrio fischeri*. Research on *Vibrio fischeri* LuxI and LuxR, increased our knowledge about AHL biosynthesis and quorum sensing transcriptional activation.

LUXI FUNCTION- Substrates for the LuxI-type enzymes are S-adenosyl methionine (SAM) and acyl-acyl carrier protein (Acyl-ACP). Acyl-ACP is synthesized in fatty acid synthesis which is additionally acted on with the aid of the LuxI enzymes in AHL biosynthesis. The LuxI proteins attach a specific Acyl-ACP to SAM via amide bond which is formed between the acyl side chain of the acyl-ACP and the amino group of the homocysteine moiety of SAM. The formation of the AHL autoinducer is effected from inside by lactonization of the ligated and with the discharge of methylthioadenosine[20].

LUXR FUNCTION- For binding of a specific AHL, binding of specific target gene promoters and transcriptional activation, LuxR-type proteins are responsible. Various domains are present in *Vibrio fischeri* LuxR protein. For DNA binding and transcriptional activation, the amino-terminal domain binds to the AHL and the carboxyl-terminal domain is needed[21]. DNA binding is stopped by amino-terminal domain via the carboxyl-terminal domain. So, when LuxR is bound to an auto-inducer this inhibitory feature is removed. To increase an amount of the LuxR protein, residues within the carboxyl-terminal domain are also required and LuxR multimers are the species that binds promoter DNA sequences[22]. The carboxyl-terminal domain of every LuxR homologue carries a highly conserved helix-flip-helix that is responsible for DNA binding. In both case, the LuxR-type of proteins bind a comparable DNA promoter detail termed the "lux box." The "lux box" consists of a 20-basepair palindromic DNA sequence situated at kind of-40 from the start site page of transcription of a given target gene[2]. It is hypothesized that the target specificity in these structures derives from the selectivity of a specific LuxR-kind protein for its cognate auto-inducer, because the DNA reputation elements within the LuxR-type proteins are conserved. Data helping this statement comes from the reality that LuxR homologues had been proven with a purpose to activate the expression of noncognate target genes. Activation of the expression of the downstream heterologous gene is done by a specific LuxR-type protein with its auto-inducer, initiating transcription by binding to different Lux boxes. Although the LuxIs produce a extraordinarily related family of molecules, in standard, the AHL auto-inducers are not able to move-stimulation of a non-cognate system. Consequently, the LuxRs are extraordinarily sensitive to alterations within the acyl side chains of the auto-inducers. Evidence for this supposition comes from numerous research demonstrating that handiest compounds that are carefully related to the true auto-inducer are able to induce weak activation of gene expression, whereas compounds with less similarity are not active. Furthermore, in numerous instances auto-inducer analogues inhibit cognate auto-inducer binding to LuxR and, therefore, inhibit target gene activation[23]. The biofilm network well explains primitive homeostasis, a primitive circulatory system, genetic material exchange, and metabolic cooperation. Biofilms are pretty widespread and hard to absolutely eliminate. Biofilms can exist on all styles of surfaces in food flora ranging from plastic, glass, metallic, wooden, and food merchandise. It is thought that biofilm formation is a part of the survival techniques of microorganisms in adverse environments. Inactivation of micro organism via traditional techniques inclusive of use of antibiotics and disinfectants are often useless towards biofilm microorganism. Bacteria attach to available surfaces in industrial

environments, and can become vast biofilm. Microflora forming biofilms consist of *Pseudomonas* spp., *Escherichia coli, Salmonella* spp., *Klebsiella* spp., *Campylobacter* spp., and *Salmonella* spp. These microorganism are of unique importance in equipped-to-consume and minimally-processed meals products, in which microbiological manage is not carried out within the terminal processing step[24]. Biofilm microbiology, by no means new, has experiencing full-size growing pains. One of the underlining motives is the shortage of standardized techniques for culturing biofilms groups. The genus *Cronobacter* accommodates the sixteen bio corporations of the rising opportunistic pathogens known formerly as *Enterobacter sakazakii*. *Cronobacter* spp. are occasional contaminants of milk powder and powdered infant formulation and constitute a good sized hazard to neonates. For this take a look at fourteen isolates of *Cronobacter sakazakii* had been selected.

AHL Pathway in *Pseudomonas aeruginosa*



Genome Neighborhood Network

The Genome Neighborhood Network allows a person to test and examine the interplay with genome community data for massive data of protein sequences, involving whole protein database. EFI-Enzyme Similarity Tool (EFI-EST) is a tool which is used to sequence datasets. It is a web tool which is used to form Sequence Similarity Networks (SSNs) which might be visualized and analyzed with the help of Cytoscape. An SSN that is to be inserted in the .Xgmml record layout, segregated into ability isofunctional households by way of filtering with the precise alignment rating is the input for Genome Neighborhood Network Tool. The genome neighborhood proteins inside an orf window on both sides of the input queries are collected from sequence files for bacterial and fungal genomes in the European Nucleotide Archive (ENA) database. Genome Neighborhood Network Tool generates two formats of the Genome Neighborhood Network (GNN) in addition to a coloured model of the input SSN that aids evaluation of the GNNs. GNNs tells about the UniProt accession IDs for the queries and the neighbors, the Pfam families for the neighbors, and each the question-neighbor distances (in orfs) and co-incidence frequencies. By using Cytoscape the GNNs and coloured SSN are downloaded, visualized, and analyzed.

GNNs can be filter out for a range of query-neighbor distances and co-incidence frequencies which permits the identification of functionally of proteins, with shorter distances and super co-incidence frequencies. With the identities of the Pfam families, the in vitro enzymatic activities of the queries and neighbors are interpreted and are expecting the reactions within the metabolic pathway.

Cytoscape

Cytoscape is a bioinformatics software program platform that is used for visualizing molecular interplay networks and integrating with gene expression profiles. Plugins are to be had for network and molecular profiling analyses, new layouts, additional file format support and connection with databases and searching in massive networks. This is normally used for the biological research applications; it is miles agnostic in terms of utilization. To download, visualize and analyze network graphs of any kind regarding nodes and edges Cytoscape is used.

Methodology

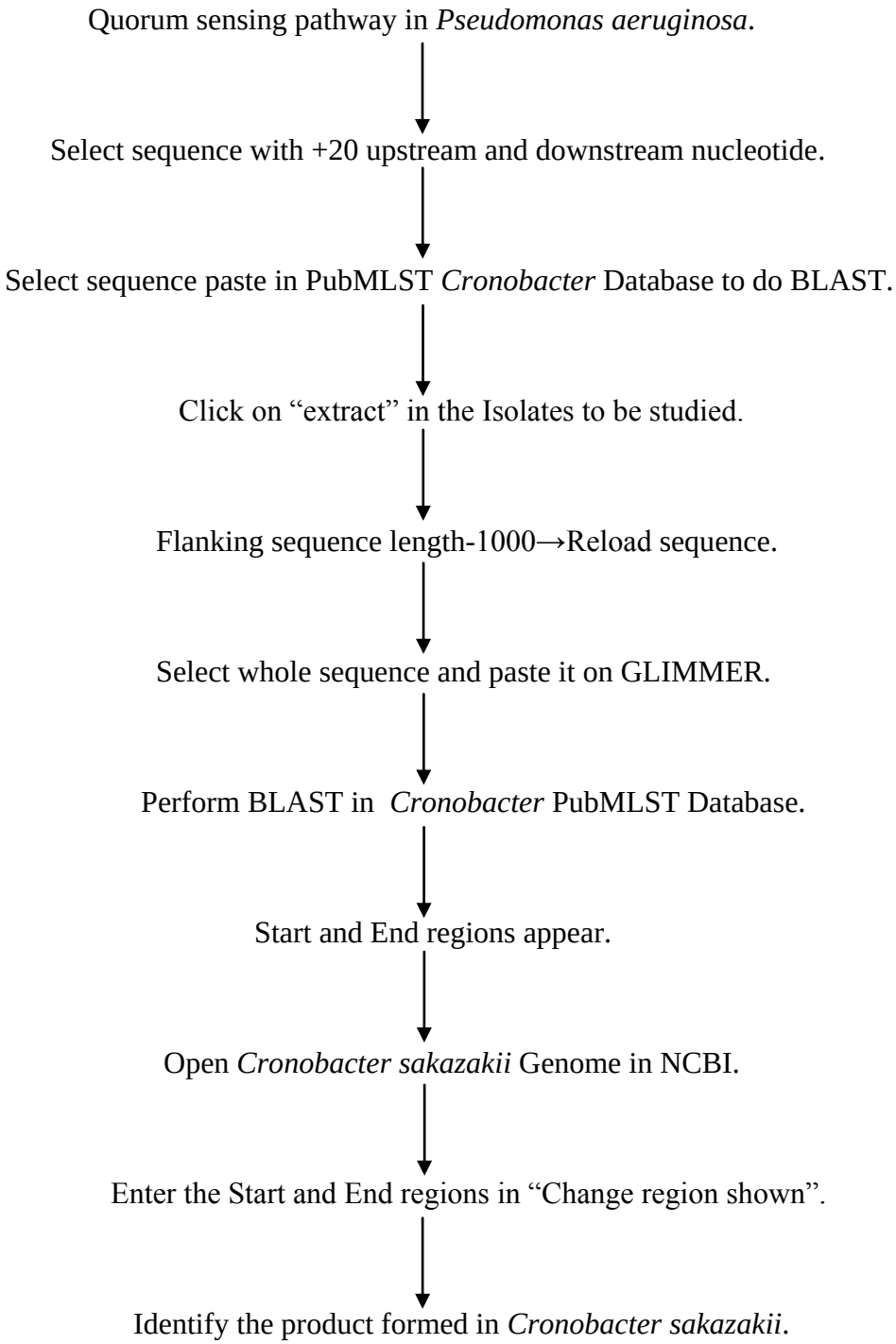
Firstly, the pathway which is responsible for Quorum Sensing mechanism in *Pseudomonas aeruginosa* was searched in the Kegg Pathway Database. Then, clicked on the gene (LasI) and PAE: pa1432 (LasI) was selected with +20 nucleotide sequence upstream and downstream both. The whole nucleotide sequence was selected and pasted it on *Cronobacter* MLST Database to do BLAST (Basic Local Alignment Search Tool). Under the "Isolates" all isolates and under "Include in results table" species were selected respectively and submitted for the BLAST. The result of the BLAST gave- Isolate id, Isolate, Species, % identity, Alignment length, Mismatches, Gaps, Seqbin id, Start, End, Orientation, E-value, Bit score. Then the isolates which are previously reported in *Cronobacter sakazakii* were selected. The isolates which were selected "Isolates: 658, 978, 680, ES15, SP291, ES 713, ES35, G-2151, E764, HPB5174, NBRC102416T, 696, 701, 716". Clicked on "extract" of the respected isolates and extracted sequence appeared on the screen. The Flanking sequence length- "1000" was done and the sequence was reloded again. The whole sequence was selected and pasted in the FASTA format into the "GLIMMER" to find the gene of interest in the Genome of *Cronobacter sakazakii* and run it. The run gave "GLIMMER orfID, Start, End, Frame and Predictions score", then on the basis of "Highest Predictions score" the start and end regions of the sequence to be selected. The selected length of sequence was copied and pasted in the *Cronobacter* PubMLST Database to do BLAST. Respective Isolate and Species were selected and submitted for the BLAST. Now start and end regions appeared with 100% identity.

Cronobacter sakazakii complete genome with RefSeq-NZ_CP011047.1 in NCBI was opened. The start and end region under "Change region shown" was inserted and selected "Update View". The entered region length gave the "Product" in *Cronobacter sakazakii* for the gene selected (LasI).

Repeated this methodology for each Gene and Isolate to get the product in *Cronobacter sakazakii*.

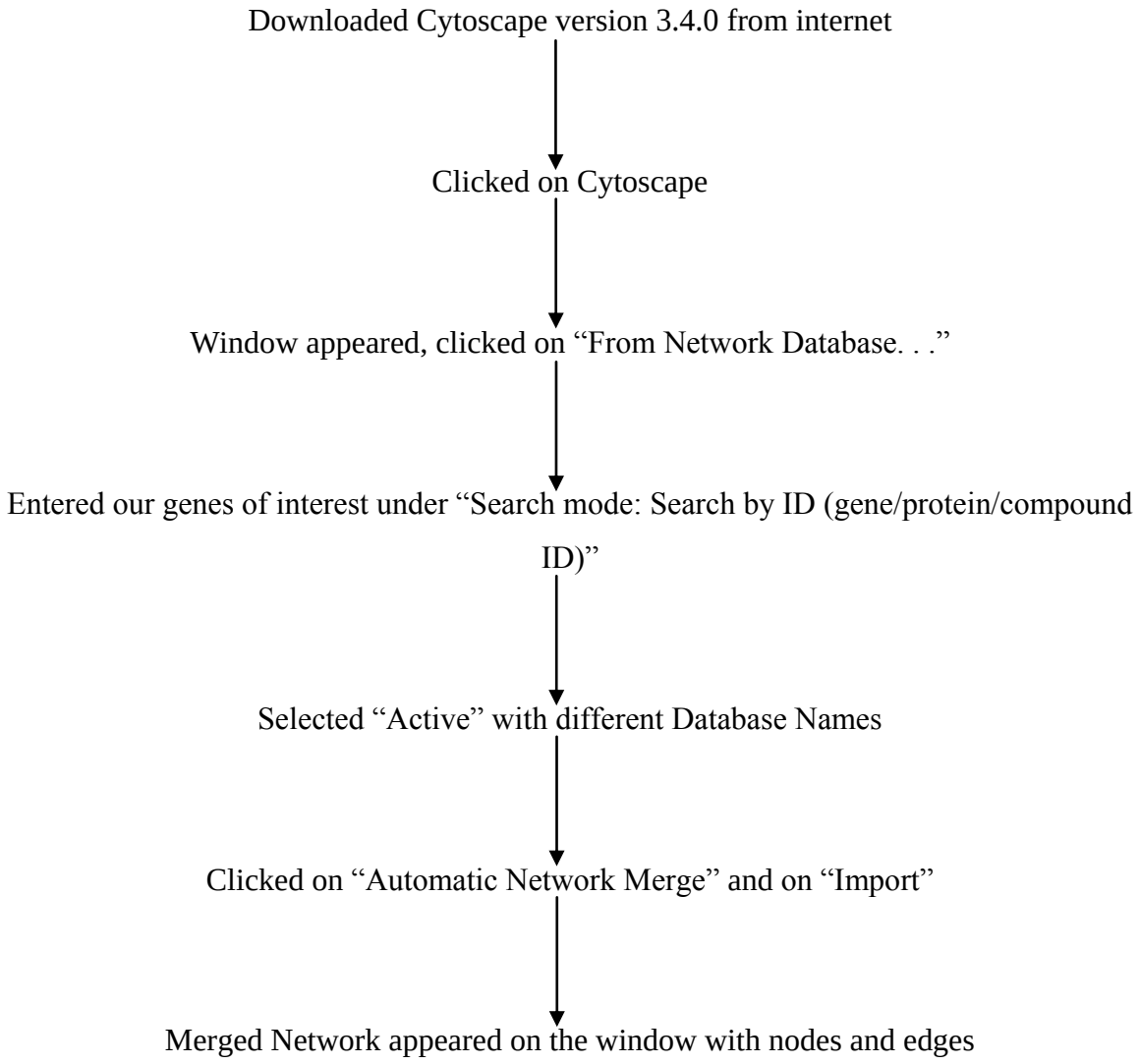
Flowchart:

The steps followed



Cytoscape:

Steps performed



Results

LasI-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.68	1123115	1123143	Citrate transporter
11	978	0.20	389360	389385	Insertion element IS1 protein InsB
46	680	0.004	44490	44520	Lytic transglycosylase
355	ES15	0.18	1177938	1177979	Biotin sulfoxide reductase
446	SP291	0.19	2664100	2664125	Hypothetical protein
591	ES713	0.20	64937	64962	Lysine transglycosylase
592	ES35	0.004	48279	48309	2,5-didehydrogluconate reductase B
593	G-2151	0.19	258601	258626	Isoleucine-t RNA ligase
594	E764	0.19	391863	391888	Resistance protein
848	HPB5174	0.19	1369663	1369688	ABC transporter ATP-binding protein
923	NBRC102416T	0.005	44276	44306	Lytic transglycosylase
1108	696	0.20	659702	659727	23s ribosomal RNA
1141	701	0.20	662175	662200	Hypothetical protein
1153	716	0.73	73007	73036	Ribonuclease HII

LasR-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.23	3625735	3625764	Phage head tail adapter
11	978	0.82	375559	375585	Maltose
46	680	2.7	57894	57914	Proline--tRNA ligase
355	ES15	0.22	3932544	3932573	Glutathione ABC transporter ATP-binding
446	SP291	0.23	3630687	3630716	Terminase
591	ES713	0.23	16485	16514	Phage Kil
592	ES35	0.77	21935	21974	Integrase
593	G-2151	0.22	12783	12812	Transcriptional regulator
594	E764	0.23	13654	13683	Transcriptional regulator
848	HPB5174	0.23	13683	13712	Hypothetical protein
923	NBRC102416T	0.81	32054	32093	DNA polymerase IV
1108	696	0.24	16293	16322	Hypothetical protein
1141	701	0.23	16406	16435	Phage Kil
1153	716	0.25	16275	16304	Hypothetical protein

PqsH-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.030	1407505	1407527	16s ribosomal RNA
11	978	0.009	49118	49150	23s ribosomal RNA
46	680	0.008	64031	64056	Lysine decarboxylase Ldcc
355	ES15	0.028	1799751	1799773	Sugar ABC transporter
446	SP291	0.10	388373	388402	Restriction endonuclease
591	ES713	0.11	50647	50665	23s ribosomal RNA
592	ES35	0.008	35484	35509	-
593	G-2151	0.10	27665	27683	Sigma-factor binding protein Crl
594	E764	0.008	61316	61360	Cytochrome c554
848	HPB5174	0.029	85139	85161	1-deoxy-D-xylulose-5-phosphate reductoisomerase
923	NBRC102416T	0.009	97076	97101	Cobalamin-binding protein
1108	696	0.009	296153	296197	Murein transglycosylase
1141	701	0.11	58614	58643	Proline—tRNA ligase
1153	716	0.032	2570	2592	Lysosyme

PqsA-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.012	1399288	1399356	Potassium transport protein
11	978	0.14	375011	375079	Galactosidase
46	680	0.039	2714	2738	Lysosyme
355	ES15	0.13	1764403	1764471	D-3-phosphoglycerate dehydrogenase
446	SP291	0.14	1236165	1236195	Anion permease
591	ES713	0.14	3915	3983	Antitermination protein
592	ES35	0.039	45247	45271	SAM-dependent methyltransferase
593	G-2151	0.14	468674	468742	Transcriptional regulator
594	E764	0.040	207682	207706	Suagr efflux transporter
848	HPB5174	0.14	372921	372949	PTS sugar transporter
923	NBRC102416T	0.041	113197	113221	ATP-dependent helicase HrpB
1108	696	0.042	330386	330410	Molecular chaperone OsmY
1141	701	0.14	156329	156357	Hypothetical protein
1153	716	0.012	41322	41390	Ribonuclease H

PqsB-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.022	952098	952122	PTS alpha-glucoside transporter
11	978	0.080	123200	123228	ABC transporter permease
46	680	0.075	5701	5726	Protein ninG
355	ES15	0.074	3311565	3311590	Putrescine ABC transporter permease
446	SP291	0.78	2988598	2988623	Glycyl radical enzyme
591	ES713	0.078	9097	9122	DNA replication protein
592	ES35	0.075	2607	2632	Lysosyme
593	G-2151	0.075	52634	52659	16s ribosomal RNA
594	E764	0.077	728173	728198	Lysophosphalipase L2.
848	HPB5174	0.077	30837	30862	Cytosol nonspecific dipeptidase
923	NBRC102416T	0.078	59947	59972	Proline—tRNA ligase
1108	696	0.081	6766	6794	Protein ninB
1141	701	0.078	973182	973207	Ribonuclease P protein component
1153	716	0.024	276638	276662	Molybdopterin adenylyltransferase

PqsC-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.027	451891	451922	Transcriptional regulator
11	978	0.098	88194	88215	30s ribosomal protein S2
46	680	0.026	15942	15964	Hypothetical protein
355	ES15	0.026	3537515	3537537	Phenylalanine—tRNA ligase
446	SP291	0.027	3210730	3210752	Alpha-trehalase
591	ES713	0.027	107196	107218	Ferrichrome porin FhuA
592	ES35	0.026	166838	166860	Recombinase family protein
593	G-2151	0.026	78022	78051	Chaperone
594	E764	0.027	78227	78256	Chaperone
848	HPB5174	0.027	163347	163369	Hypothetical protein
923	NBRC102416T	0.027	15851	15873	Hypothetical protein
1108	696	0.027	269658	269687	Hypothetical protein
1141	701	0.027	164389	164411	Hypothetical protein
1153	716	0.029	216841	216872	Ribulokinase

PqsD-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.027	1279788	1279817	Gluconate transporter
11	978	0.027	382732	382151	GntR family transcriptional regulator
46	680	0.026	19284	19303	Hypothetical protein
355	ES15	0.025	338352	338407	PTS cellobiose IIC component
446	SP291	0.026	925504	925523	Hypothetical protein
591	ES713	0.027	69306	69335	Acetyl-CoA carboxylase carboxyltransferase
592	ES35	0.025	56198	56217	Methionine ABC trsnsporter permease
593	G-2151	0.026	122123	122142	Polysaccharide deacytlase
594	E764	0.026	123907	123926	Carbonate dehydratase
848	HPB5174	0.026	365552	365576	-
923	NBRC102416T	0.027	80425	80454	-
1108	696	0.028	654757	654781	Amino acid ABC transporter permease
1141	701	0.027	22668	22687	Integrase
1153	716	0.028	10666	10695	Replication protein

PqsE-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.024	2929629	2929666	Formyltetrahydrofolate deformylase
11	978	0.007	73200	73228	Ribonuclease HII
46	680	0.023	29799	29823	Xanthine phosphoribosyltransferase
355	ES15	0.022	3291512	3291549	Hypothetical protein
446	SP291	0.024	2968540	2968577	Hypothetical protein
591	ES713	0.024	29135	29172	Xanthine phosphoribosyltransferase
592	ES35	0.028	11577	11612	Hypothetical protein
593	G-2151	0.023	72677	72714	Ribonuclease HII
594	E764	0.028	45214	45234	SAM-dependent methyltransferase
848	HPB5174	0.028	50883	50918	23s ribosomal RNA
923	NBRC102416T	0.024	10672	10696	DNA replication protein
1108	696	0.30	97645	97665	DNA replication protein
1141	701	0.024	953125	953162	GntR family transcriptional regulator
1153	716	0.025	947547	947584	MFS transporter

PqsR-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	0.0002	1839564	1839612	Exonuclease V subunit alpha
11	978	0.0002	88220	88268	Type I methionyl aminopeptidase
46	680	0.0002	7075	7123	23s ribosomal RNA
355	ES15	0.0002	2191515	2191563	Hypothetical protein
446	SP291	0.0002	1884166	1884214	Siroheme synthase
591	ES713	0.0002	85153	85201	UMP kinase
592	ES35	0.0002	316	364	Antirepressor
593	G-2151	0.0002	79897	79945	Outer memberane protein assembly factor BamA
594	E764	0.0002	1024553	1024601	16s ribosomal RNA
848	HPB5174	0.0002	463069	463117	Transposase
923	NBRC102416T	0.0002	23706	23754	Gamma-glutamyl-phosphate reductase
1108	696	0.0002	141287	141335	Dihydrolipoyl dehydrogenase
1141	701	0.0002	123829	123877	Carbonate dehydratese
1153	716	0.0002	152556	152604	AmpE protein

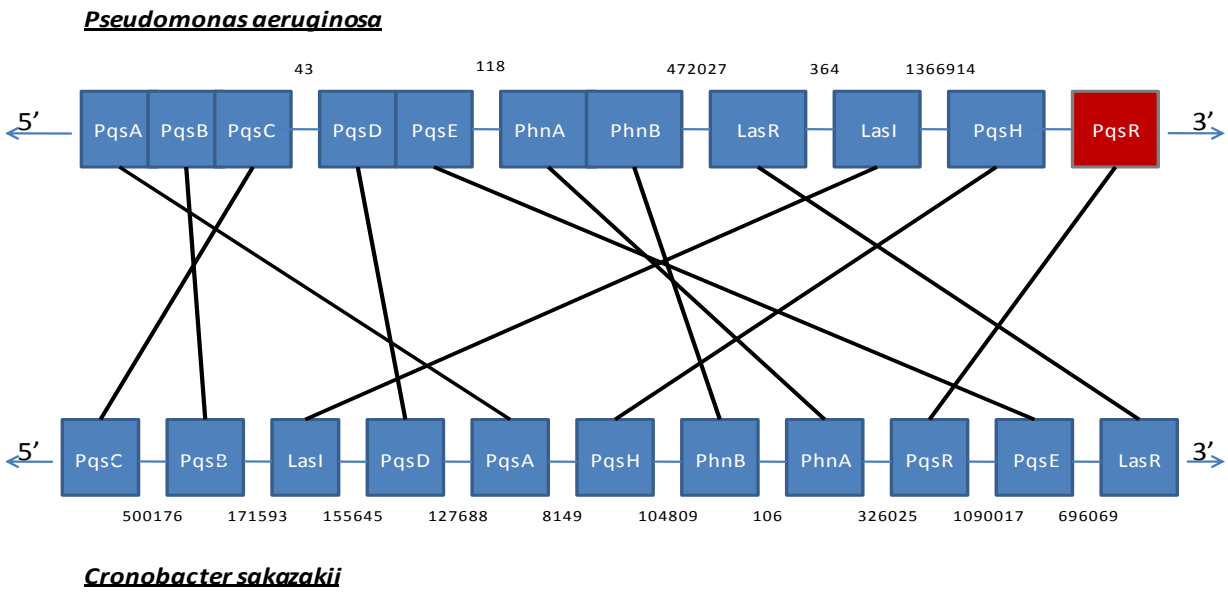
PhnA-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	1e-85	1512703	1513539	Aspartate ammonia-lyase
11	978	4e-72	260045	260730	Isoleucine—tRNA ligase
46	680	5e-83	2235	3071	Hypothetical protein
355	ES15	8e-81	1905509	1906345	Amino acid ABC transporter permease
446	SP291	1e-79	1571613	1582437	2’,3’-cyclic-nucleotide 2’-phosphodiesterase
591	ES713	1e-79	76383	77207	Chaperone
592	ES35	5e-83	16231	17067	Hypothetical protein
593	G-2151	9e-80	355304	356128	Hypothetical protein
594	E764	2e-82	746277	747101	Hydroxymethylbilane synthase
848	HPB5174	1e-84	190400	191236	Peptidoglycan glycosyltransferase
923	NBRC102416T	5e-83	159735	160571	Hypothetical protein
1108	696	2e-82	418842	419666	Hypothetical protein
1141	701	1e-79	399259	400083	ATP dependernt DNA helicase RecQ
1153	716	1e-85	107768	108604	Ferrichrome porin FhuA

PhnB-Location and function in *Cronobacter sakazakii* genome

ISOLATE ID	ISOLATE	E-VALUE	START	END	PRODUCT IN C.SAKAZAKII
5	658	3e-11	1512336	1512597	C4-dicarboxylate ABC transporter
11	978	1e-09	86298	86378	UMP kinase
46	680	1e-09	25913	25993	Porin OmpC
355	ES15	1e-09	335006	335086	16s rRNA-methyltransferase
446	SP291	1e-09	4316115	4316195	Uracil reductase
591	ES713	1e-09	25664	25744	Porin OmpC
592	ES35	1e-09	28005	28085	Hypothetical protein
593	G-2151	1e-09	25369	25449	Porin OmpC
594	E764	4e-10	25380	25672	Porin OmpC
848	HPB5174	1e-09	25494	25574	Porin OmpC
923	NBRC102416T	1e-09	52561	52641	16s ribosomal RNA
1108	696	4e-10	322256	322548	Purine-nucleoside phosphorylase
1141	701	1e-09	25495	25575	Porin OmpC
1153	716	4e-11	107401	107662	Ferrichrome porin FhuA

Relation between *Cronobacter sakazakii* (Isolate id-5) and *Pseudomonas aeruginosa*



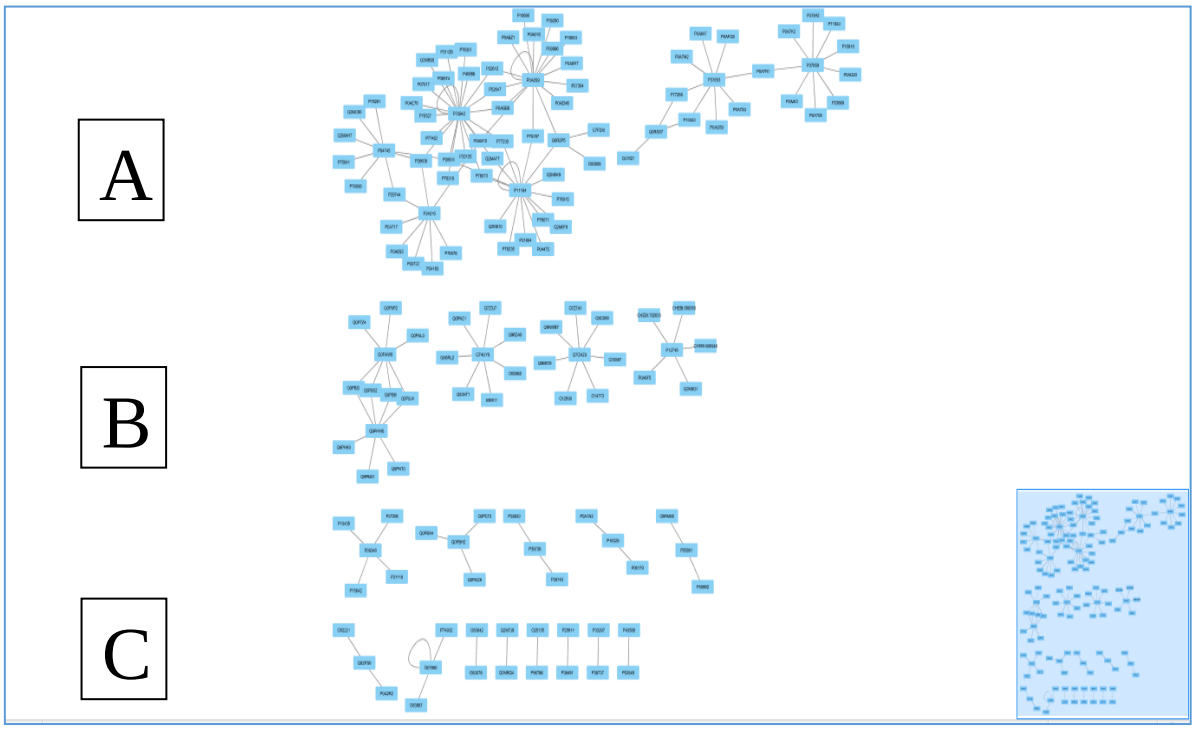
Pseudomonas aeruginosa

GENE	START	END	INTRA GENE	INTER GENE
PqsA	10,78,462	10,80,015	1,553	-6
PqsB	10,80,009	10,80,860	851	-7
PqsC	10,80,853	10,81,899	1,046	43
PqsD	10,81,942	10,82,955	1,013	-6
PqsE	10,82,949	10,83,854	905	118
PhnA	10,83,972	10,85,564	1,592	-22
PhnB	10,85,542	10,86,144	602	4,72,027
LasR	15,58,171	15,58,890	719	364
LasI	15,59,254	15,59,859	605	13,66,914
PqsH	29,26,773	29,27,921	1,148	-
PqsR	-	-	-	-

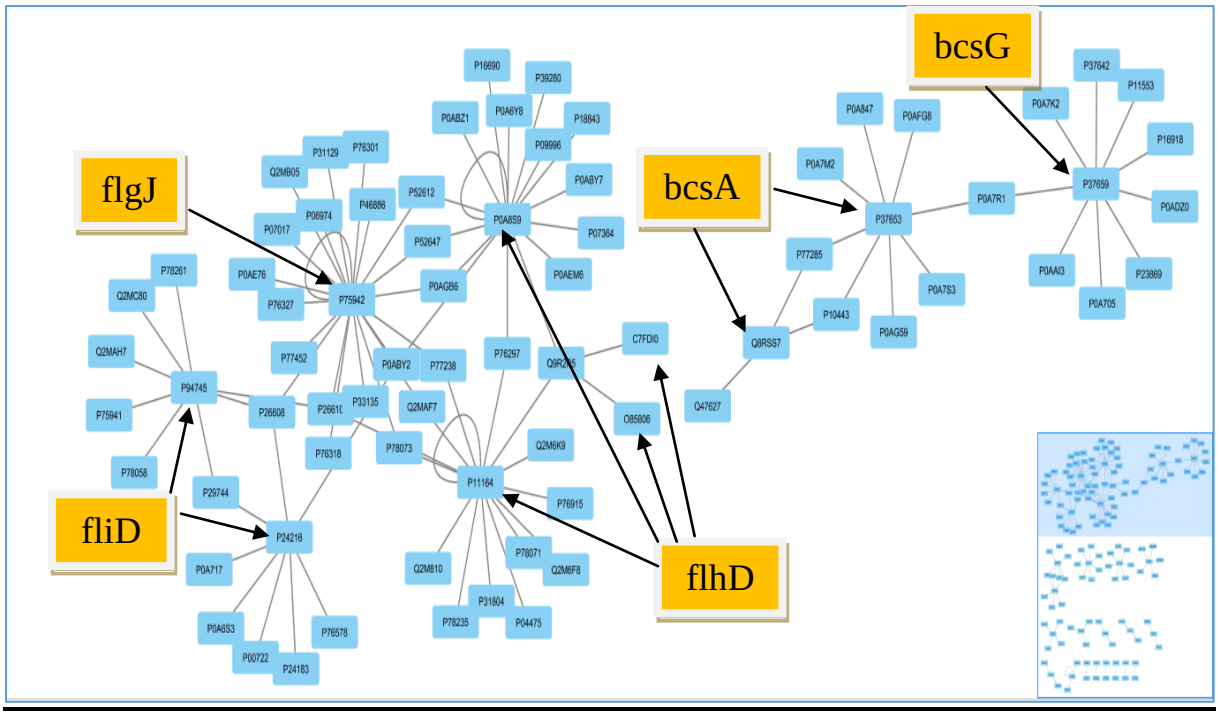
Cronobacter sakazakii

GENE	START	END	INTRA GENE	INTER GENE
PqsC	4,51,891	4,51,922	31	5,00,176
PqsB	9,52,098	9,52,122	24	1,71,593
LasI	11,23,715	11,24,143	428	1,55,645
PqsD	12,79,788	12,79,817	29	1,27,688
PqsA	13,99,288	13,99,356	68	8,149
PqsH	14,07,505	14,07,527	22	1,04,809
PhnB	15,12,336	15,12,597	261	106
PhnA	15,12,703	15,13,539	836	3,26,025
PqsR	18,39,564	18,39,612	48	10,90,017
PqsE	29,29,629	29,29,666	37	6,96,069
LasR	36,25,735	36,25,764	29	-

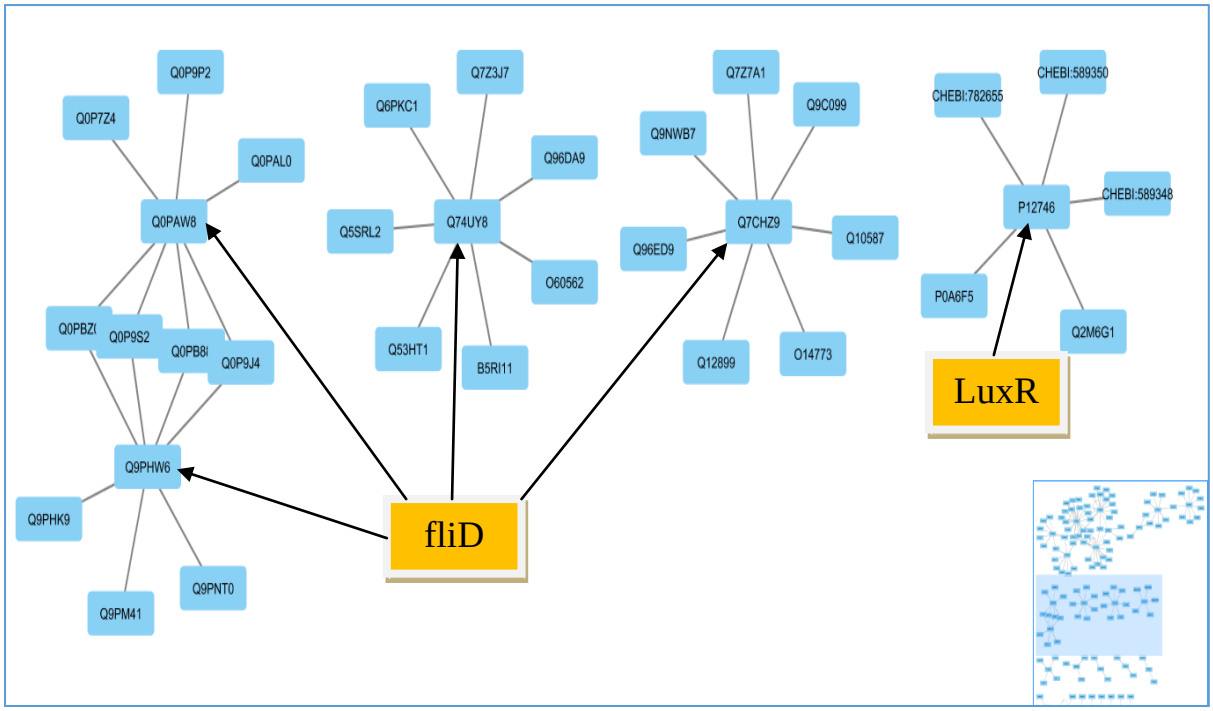
Sequence Similarity Network in Cytoscape (C. *sakazakii*)



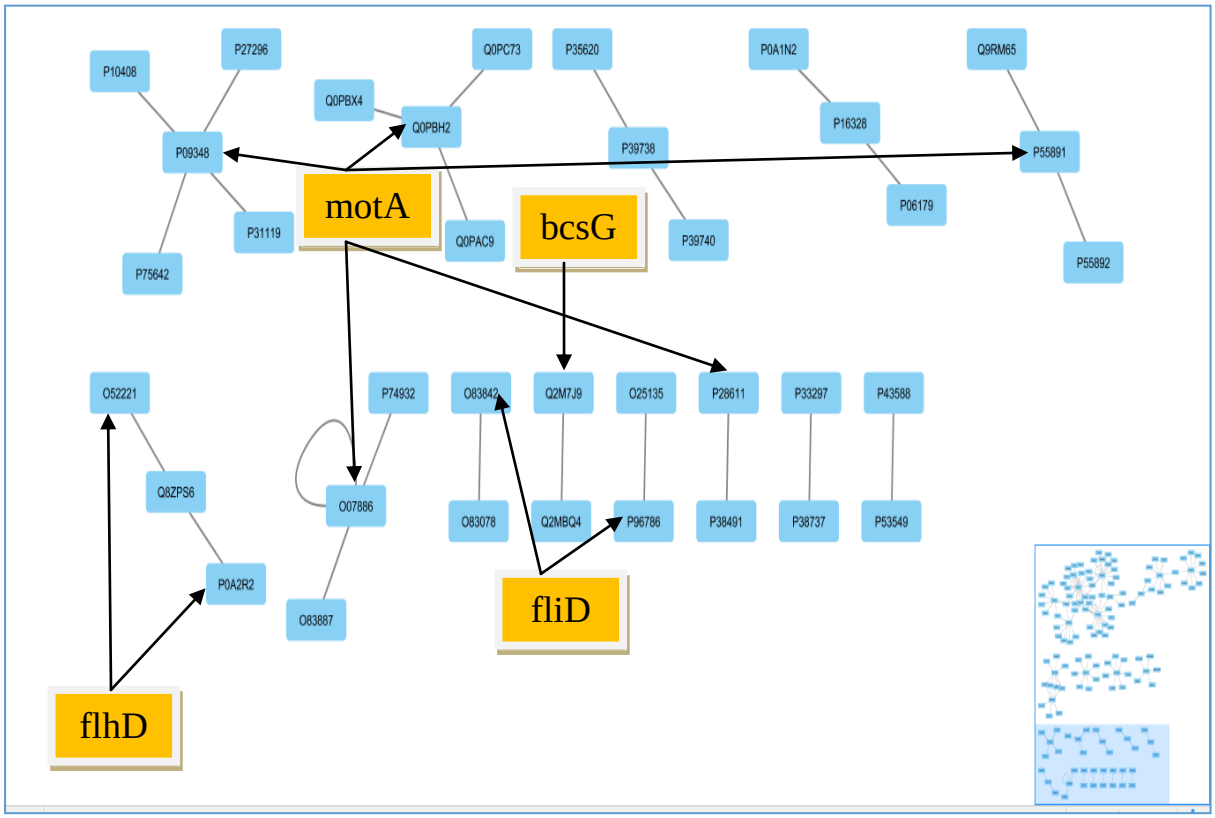
Result A,B and C are shown below



A


Q0P9P2
Q0P7Z4
Q0PAL0
Q0PAW8
Q0PBZ
Q0P9S2
Q0PB8
Q0P9J4
Q9PHW6
Q9PHK9
Q9PM41
Q9PNT0
Q6PKC1
Q7Z3J7
Q96DA9
Q5SRL2
Q74UY8
Q60562
Q53HT1
B5RI11
fliD
Q9NWB7
Q7Z7A1
Q9C099
Q96ED9
Q7CHZ9
Q10587
Q12899
O14773
CHEBI:782655
CHEBI:589350
P12746
CHEBI:589348
P0A6F5
Q2M6G1
LuxR

B


P10408
P27296
P09348
P31119
P75642
Q0PBX4
Q0PC73
Q0PBH2
Q0PAC9
motA
P35620
P39738
P39740
bcsG
P0A1N2
P16328
P06179
Q9RM65
P55891
P55892
O52221
Q8ZPS6
P0A2R2
flhD
P74932
O07886
O83887
O83842
O83078
Q2M7J9
Q2MBQ4
O25135
P96786
fliD
P28611
P38491
P33297
P38737
P43588
P53549

C

Interaction of Biofilm genes in *Cronobacter sakazakii*

Biofilm genes	**Interaction with**	**Significance**
flgJ	fliD, flhD	Flagella formation
fliD	flgJ, flhD	Flagella formation
bcsA	bcsG	Cellulose biosynthesis
luxR	-	Quorum sensing
motA	-	Motility
flhD	flgJ, fliD	Flagella formation
bcsG	bcsA	Cellulose biosynthesis

Summary and Conclusions

The present study investigated the role of different quorum sensing genes in *Cronobacter sakazakii*. Our laboratory has previously reported the expression of long chain AHLs in *Cronobacter sakazakii*. To elucidate the role of different genes involved in quorum sensing of this pathogen, a total of 11 genes involved in quorum sensing in *Pseudomonas aeruginosa* (a model organism) were BLAST in PUBMLST database for *Cronobacter sakazakii*. The BLAST analysis indicated of genome comparison indicated that among 11 genes, except LasR, was not observed in our operonic structure of *Cronobacter sakazakii*. Followed by operonic structure of QS genes, their relation with the biofilm forming genes in *Cronobacter sakazakii* was determined using Cytoscape. The network obtained after Cytoscape analysis indicated that genes regulating the cellulose biosynthesis are bcsG, bcsA, genes regulating flagella formation are flgJ, flhD, and fliD, gene responsible for quorum sensing is luxR and gene responsible for motility is motA.

References

1. Nealson, K. H., and John W. Hastings. "Bacterial bioluminescence: its control and ecological significance." *Microbiological reviews* 43.4 (1979): 496.
2. Fuqua, W. Claiborne, Stephen C. Winans, and E. Peter Greenberg. "Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators." *Journal of bacteriology* 176.2 (1994): 269.
3. Pesci, Everett C., and Barbara H. Iglewski. "The chain of command in Pseudomonas quorum sensing." *Trends in microbiology* 5.4 (1997): 132-134.
4. Engebrecht, JoAnne, and Michael Silverman. "Identification of genes and gene products necessary for bacterial bioluminescence." *Proceedings of the National Academy of Sciences* 81.13 (1984): 4154-4158.
5. Davies, David G., et al. "The involvement of cell-to-cell signals in the development of a bacterial biofilm." *Science* 280.5361 (1998): 295-298.
6. Zhu, Jun, and Stephen C. Winans. "Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells." *Proceedings of the National Academy of Sciences* 96.9 (1999): 4832-4837.
7. Wang, X. D., P. AJ de Boer, and Lawrence I. Rothfield. "A factor that positively regulates cell division by activating transcription of the major cluster of essential cell division genes of Escherichia coli." *The EMBO journal* 10.11 (1991): 3363.
8. Swift, Simon, et al. "Quorum sensing-dependent regulation and blockade of exoprotease production in Aeromonas hydrophila." *Infection and immunity* 67.10 (1999): 5192-5199.
9. Parsek, Matthew R., and E. Peter Greenberg. "Acyl-homoserine lactone quorum sensing in gram-negative bacteria: a signaling mechanism involved in associations with higher organisms." *Proceedings of the National Academy of Sciences* 97.16 (2000): 8789-8793.
10. Rahme, Laurence G., et al. "Common virulence factors for bacterial pathogenicity in plants and animals." *Science* 268.5219 (1995): 1899..
11. Pearson, James P., et al. "A second N-acylhomoserine lactone signal produced by Pseudomonas aeruginosa." *Proceedings of the National Academy of Sciences* 92.5 (1995): 1490-1494.
12. Dawson, Martin H., and Richard HP Sia. "In vitro transformation of pneumococcal types: I. a technique for inducing transformation of pneumococcal types in vitro." *The Journal of experimental medicine* 54.5 (1931): 681.

13. Dunny, Gary M., Byron L. Brown, and Don B. Clewell. "Induced cell aggregation and mating in Streptococcus faecalis: evidence for a bacterial sex pheromone." *Proceedings of the National Academy of Sciences* 75.7 (1978): 3479-3483.
14. Dworkin, Martin, and Dale Kaiser. "Cell interactions in myxobacterial growth and development." *Science* 230.4721 (1985): 18-24.
15. Passador, Luciano, et al. "Expression of Pseudomonas aeruginosa virulence genes requires cell-to-cell communication." *Science* 260.5111 (1993): 1127-1130.
16. Brint, J. Mark, and Dennis E. Ohman. "Synthesis of multiple exoproducts in Pseudomonas aeruginosa is under the control of RhlR-RhlI, another set of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family." *Journal of Bacteriology* 177.24 (1995): 7155-7163.
17. Ochsner, Urs A., and Jakob Reiser. "Autoinducer-mediated regulation of rhamnolipid biosurfactant synthesis in Pseudomonas aeruginosa." *Proceedings of the National Academy of Sciences* 92.14 (1995): 6424-6428.
18. Pesci, Everett C., et al. "Regulation of las and rhl quorum sensing in Pseudomonas aeruginosa." *Journal of bacteriology* 179.10 (1997): 3127-3132.
19. Pesci, Everett C., et al. "Quinolone signaling in the cell-to-cell communication system of Pseudomonas aeruginosa." *Proceedings of the National Academy of Sciences* 96.20 (1999): 11229-11234.
20. Moré, Margret I., L. David Finger, and Joel L. Stryker. "Enzymatic synthesis of a quorum-sensing autoinducer through use of defined substrates." *Science* 272.5268 (1996): 1655.
21. Choi, S. H., and E. P. Greenberg. "The C-terminal region of the Vibrio fischeri LuxR protein contains an inducer-independent lux gene activating domain." *Proceedings of the National Academy of Sciences* 88.24 (1991): 11115-11119.
22. Stevens, Ann M., Katherine M. Dolan, and E. Peter Greenberg. "Synergistic binding of the Vibrio fischeri LuxR transcriptional activator domain and RNA polymerase to the lux promoter region." *Proceedings of the National Academy of Sciences* 91.26 (1994): 12619-12623.
23. Gray, Kendall M., et al. "Interchangeability and specificity of components from the quorum-sensing regulatory systems of Vibrio fischeri and Pseudomonas aeruginosa." *Journal of bacteriology* 176.10 (1994): 3076-3080.

24. Davey, Mary Ellen, and George A. O'toole. "Microbial biofilms: from ecology to molecular genetics." *Microbiology and molecular biology reviews* 64.4 (2000): 847-867.