**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO**

EVELYN AYUMI ONGA

**Perfil transcricional da haloarqueia *Halobacterium salinarum* NRC-1
sob estresse osmótico**

**Ribeirão Preto
2021**

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO**

EVELYN AYUMI ONGA

**Perfil transcricional da haloarqueia *Halobacterium salinarum* NRC-1
sob estresse osmótico**

Versão Corrigida

Dissertação apresentada ao Programa de Pós-Graduação
em Bioquímica da Faculdade de Medicina de Ribeirão
Preto da Universidade de São Paulo para a obtenção do
título de Mestre em Ciências.

Área de Concentração: Bioquímica

Orientadora: Prof^a. Dr^a. Tie Koide

Ribeirão Preto

2021

Autorizo, a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na publicação
Serviço de Biblioteca e Documentação
Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo

Onga, Evelyn Ayumi

Perfil transcricional da haloarqueia *Halobacterium salinarum* NRC-1 sob estresse osmótico

Evelyn Ayumi Onga; orientadora, Tie Koide. Ribeirão Preto - 2021

150 f. :il. 30cm

Dissertação (Mestrado em Ciências) – Programa de Pós-Graduação em Bioquímica,
Universidade de São Paulo, Ribeirão Preto, 2021.

Versão Original

1-RNA-seq; 2- Expressão gênica diferencial; 3- Estresse osmótico; 4- *Halobacterium salinarum*; 5- Archaea.

Nome: ONGA, Evelyn Ayumi

Título: Perfil transcricional da haloarqueia *Halobacterium salinarum* NRC-1 sob estresse osmótico

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo, para obtenção do Título de Mestre em Ciências.

Aprovado em: ____/____/____

Banca Examinadora

Prof(a). Dr (a). _____

Instituição: _____

Julgamento: _____

Assinatura: _____

Prof(a).Dr (a)._____

Instituição: _____

Julgamento: _____

Assinatura: _____

Prof(a). Dr (a). _____

Instituição: _____

Julgamento: _____

Assinatura: _____

Apoio e suporte financeiro

Este projeto foi realizado com o apoio financeiro das seguintes entidades:

Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (Processo: 2015/21038-1)

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES

Fundação de Apoio ao Ensino, Pesquisa e Assistência – FAEPA

Agradecimentos

Agradeço a minha família, meus pais, meu irmão, pela ótima criação, por todo apoio, e por sempre confiarem na minha capacidade. Agradeço ao Mitoshi por todo apoio emocional, carinho e companheirismo ao longo dos anos. Agradeço meus queridos amigos Robson, Yurie, Rafael, Lina, Sete, Rato, Shoiti, Danielle, Raphaela e Adriele pela amizade, conversas e maravilhosa companhia, seja pessoalmente ou *online*. Agradeço a Graciely, minha *housemate*, que além de ter se tornado uma grande amiga, me ajudou a me adaptar a Ribeirão Preto, me apresentou a outras pessoas maravilhosas aqui e fez o tempo de quarentena em casa ser menos louco.

Agradeço aos colegas Junier, Alan e Amr por me ajudarem desde o começo do mestrado e por me apoiarem na realização deste projeto. Agradeço também aos colegas do laboratório, Beatriz, Catarina, Júlia, Rhuan e Vinícius pela animação, ensinamentos e ótima companhia.

Agradeço a Silvia, técnica do nosso laboratório, por ser sempre tão dedicada, por ajudar nos experimentos e por todo apoio, tanto emocional, quanto profissional, que ofereceu a mim.

Agradeço ao Programa de Pós-graduação em Bioquímica, aos secretários do programa e a todos colegas pós-graduandos pelos ótimos momentos passados juntos. Agradeço a CAPES pela bolsa de estudos e pela prorrogação da bolsa durante a pandemia.

Agradeço à Prof^a Tie e Prof^o Ricardo pela orientação, pela oportunidade de aprender e ensinar, e por oferecerem uma ótima estrutura para que a ciência seja feita no laboratório.

Resumo

ONGA, Evelyn Ayumi. **Perfil transcricional da haloarqueia *Halobacterium salinarum* NRC-1 sob estresse osmótico**. 2021, 148 f. Dissertação (Mestrado em Bioquímica) – Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2021.

Os lagos salinos podem sofrer variações no decorrer de um ano devido a mudanças de temperatura, incidência de radiação solar, chuvas e evaporação da água. Essas duas últimas afetam a concentração de sais na água e podem ter implicações na pressão osmótica e também no enovelamento das proteínas de organismos que vivem nesses habitats, como as haloarqueias. *Halobacterium salinarum* NRC-1 é uma arqueia halofílica extrema, cuja concentração ótima de crescimento é 4.3M NaCl. Um mecanismo de adaptação desse microrganismo ao seu habitat consiste na manutenção de altos níveis de cloreto de potássio internos (podendo chegar a 5M), além de apresentar proteoma com pl médio de 4,9, com muitos aminoácidos negativamente carregados em sua superfície. Os aminoácidos negativamente carregados interagem com a água e mantêm a força iônica e camada de solvatação que estabiliza a estrutura das proteínas e mantêm suas funções. O objetivo deste trabalho foi identificar quais genes e mecanismos são utilizados por *H. salinarum* NRC-1 para sua sobrevivência em condições de hipo ou hipersalinidade. O transcrito de amostras em meio de cultura com salinidade ótima (4,3M de NaCl), hiposalinidade (2,6M de NaCl) e hipersalinidade (5,1M de NaCl) foi mapeado utilizando o método de RNAseq e foi feita a comparação com experimentos de microarranjos de DNA e proteoma em condições semelhantes disponíveis na literatura. Foram encontrados 35 genes menos expressos e 31 genes mais expressos em hipersalinidade em relação à condição controle, enquanto em hiposalinidade foram 268 genes menos expressos e 209 mais expressos em relação à condição controle. A análise das funções dos genes mostrou alterações nos processos de tradução, formação do arqueolo, mecanismo de obtenção de energia através da luz, vesículas de gás e metabolismo de glicerol. Foram feitos experimentos que confirmaram a expressão diferencial de genes relacionados à formação do arqueolo (ensaio de motilidade) e metabolismo de glicerol, com uma curva de crescimento suplementada com glicerol. No ensaio de motilidade observou-se que as células em meio semi-sólido não se movimentam em hiposalinidade, diferente do que acontece em meio controle, onde as células são capazes de colonizar toda a superfície de uma placa de petri. Já a curva de crescimento com suplementação de glicerol indicou que esse soluto compatível permite um melhor crescimento em baixa salinidade, enquanto em ótima salinidade causa diminuição no crescimento. Possíveis reguladores de expressão gênica na forma de RNAs antisense foram identificados e descritos.

Palavras-chave: RNA-seq. Expressão gênica diferencial. Estresse osmótico. *Halobacterium salinarum*. Archaea.

Abstract

ONGA, Evelyn Ayumi. **Transcriptional profile of haloarchaea *Halobacterium salinarum* NRC-1 under osmotic stress**. 2021, 148 f. Dissertation (Master in Biochemistry) - Ribeirão Preto Medical School - University of São Paulo, Ribeirão Preto, 2021.

Saline lakes may change over the course of a year due to changes in temperature, incidence of solar radiation, rain and water evaporation. These last two affect the concentration of salts in the water and may have implications for the osmotic pressure and also for protein folding of organisms that live in these habitats, such as the haloarchaea. *Halobacterium salinarum* NRC-1 is an extreme halophilic archaeon, whose optimal growth concentration is 4.3M NaCl. An adaptation mechanism this microorganism uses to survive consists in maintaining high levels of internal potassium chloride (which can reach 5M), in addition to presenting a proteome with an average pI of 4.9, with many negatively charged amino acids on its protein's surface. The negatively charged amino acids interact with water and maintain the ionic strength and solvation layer that stabilizes the structure of proteins and maintain their functions. The objective of this work was to identify which genes and mechanisms are used by *H. salinarum* NRC-1 for its survival in conditions of hypo or hypersalinity. The transcriptome of samples in culture medium with optimal salinity (4.3M NaCl), hyposalinity (2.6M NaCl) and hypersalinity (5.1M NaCl) was mapped using the RNAseq method and a comparison was made with DNA microarrays and proteome data under similar conditions available in the literature. There were 35 down-regulated genes and 31 up-regulated genes in hypersalinity compared to the control condition, while in hyposalinity there were 268 down-regulated genes and 209 up-regulated compared to the control condition. The analysis of the gene functions showed alterations in the translation process, formation of the archellum, in the mechanism of obtaining energy through light, gas vesicles and glycerol metabolism. Experiments were carried out that confirmed the differential expression of genes related to the formation of the archellum (motility assay) and glycerol metabolism, with a growth curve supplemented with glycerol. In the motility assay, it was observed that cells in a semi-solid medium do not migrate in hyposalinity, unlike what happens in the control medium, where cells are able to colonize the entire surface of a petri dish. The growth curve with glycerol supplementation, on the other hand, indicated that this compatible solute allows for better growth at low salinity, while at optimal salinity it causes a decrease in cell growth. Possible regulators of gene expression in the form of antisense RNAs have been identified and described.

Keywords: RNA-seq. Differential gene expression. Osmotic stress. *Halobacterium salinarum*. Archaea.

Lista de figuras

Figura 1: Árvore filogenética proposta por Woese, Kandler e Wheelis em 1990.

Figura 2: Representação da relação entre eucariotos e arqueias nos últimos 40 anos.

Figura 3: Habitat natural, coloração das colônias e micrografias de *H. salinarum*.

Figura 4: Esquema representando as várias camadas de processamento de informação genética conhecidas.

Figura 5: Esquema representando alguns dos tipos de RNA existentes.

Figura 6: Curva de crescimento de *H. salinarum* NRC-1 em meio de cultura ótimo, hiposalino e hipersalino e diferença de densidade óptica dos tratamentos.

Figura 7: Após 24 horas sem agitação as culturas de *H. salinarum* NRC-1 em condição controle e hiposalinidade.

Figura 8: *Scatter plot* comparando o valor de M das duas réplicas para o tratamento de hipersalinidade.

Figura 9: Representação do enriquecimento funcional das funções moleculares diferencialmente expressas em meio hipersalino.

Figura 10: Representação do enriquecimento funcional de processos biológicos diferencialmente expressos em meio hipersalino.

Figura 11: *Scatter plot* comparando o valor de M das duas réplicas para o tratamento de hiposalinidade.

Figura 12: Representação do enriquecimento funcional das funções moleculares diferencialmente expressas em hiposalinidade.

Figura 13: Representação do enriquecimento funcional dos processos biológicos diferencialmente expressos em hiposalinidade.

Figura 14: Curva de crescimento de *H. salinarum* NRC-1 em meio de cultura com 2.6M de NaCl e 4.3 M de NaCl suplementado com 30 mg/ml e 60 mg/ml de glicerol.

Figura 15: Gráfico de dispersão das amostras em hipersalinidade de RNA-seq versus Microarranjo de DNA.

Figura 16: Gráfico de dispersão das amostras em hiposalinidade de RNA-seq versus Microarranjo de DNA.

Figura 17: Gráfico de dispersão das amostras em hipersalinidade de RNA-seq versus proteômica.

Figura 18: Gráfico de dispersão das amostras em hiposalinidade de RNA-seq versus proteômica.

Figura 19: Ensaio de motilidade.

Figura 20: Quantificação e dimensões das vesículas de gás.

Figura 21: Representação do cluster *gvp* e promotores.

Figura 22: Metabolismo de glicerol em haloarqueias.

Figura 23: Operon arqueio e representação do modelo motor das arqueias.

Figura 24: Representação dos principais processos biológicos diferencialmente expressos em hiposalinidade e hipersalinidade.

Figura S1: Cobertura dos reads no gene *kef1* (VNG_RS07995).

Lista de Tabelas

Tabela 1: Números de acesso e nome das amostras e bioprojeto no website do SRA.

Tabela 2: Distribuição dos reads entre amostras analisadas.

Tabela 3: Genes menos expressos em meio hipersalino.

Tabela 4: Genes mais expressos em hipersalinidade.

Tabela 5: Tabela contendo os 35 genes menos expressos em meio hiposalino.

Tabela 6: Tabela contendo os 35 genes mais expressos em hiposalinidade.

Tabela 7: Condições de cada experimento para o meio 4,3M de NaCl.

Tabela 8: Genes comuns aos experimentos de proteômica e microarranjo de DNA em hipersalinidade.

Tabela 9: Genes comuns aos experimentos de proteômica e microarranjo de DNA em hiposalinidade.

Tabela 10: Tabela das regiões antisense encontradas.

Tabela 11: Tabela contendo as 21 regiões intergênicas encontradas.

Tabela S1: Tabela de genes menos expressos em hiposalinidade.

Tabela S2: Tabela de genes mais expressos em hiposalinidade.

Tabela S3: Tabela de replicatas do RNA-seq vs microarranjos de DNA em hipersalinidade.

Tabela S4: Tabela de replicatas do RNA-seq vs proteômica em hipersalinidade.

Tabela S5: Tabela de replicatas do RNA-seq vs microarranjos de DNA em hiposalinidade.

Tabela S6: Tabela de replicatas do RNA-seq vs proteômica em hiposalinidade.

Tabela S7: Tabela contendo links para resultados detalhados do projeto.

Abreviaturas

6-HMDP - 6-hidroximentil-diidropteridina difosfato

asRNA - *antisense RNA*; RNA antisense

CDS - *coding sequence*; sequência codificadora

CM - *complex media*; meio complexo

CRISPR - *clustered regularly interspaced short palindromic repeats*; agrupados de curtas repetições palindrômicas regularmente interespaçadas

csp - *cold shock protein*

D.O. - densidade óptica

DEPC - dietil pirocarbonato

DHAP - diidroxiacetona fosfato

DMSO - dimetil sulfóxido

DNA - ácido desoxirribonucleico

DSB - *double-strand break*; quebra de fita dupla

DUF - *domain of unknown function*; domínio de função desconhecida

GO - *gene ontology*

hsp - *heat shock protein*

iTSS - *internal transcription start sites*; sítio de início de transcrição interno

LUCA - *last universal common ancestor*; último ancestral comum universal

miRNA - microRNA

mRNA - *messenger RNA*; RNA mensageiro

ncRNAs - *non coding RNAs*; RNAs não codificantes

PCA - análise de componentes principais

PCR - *polimerase chain reaction*; reação em cadeia de polimerase

PDB - *Protein Data Bank*

pI - ponto isoelétrico

PIR - *Protein Information Resource*

PRF - *Protein Research Foundation*

RBP - *RNA binding protein*; proteína que se liga a RNA

RNA - ácido ribonucleico

RNA-seq - *RNA sequencing*; sequenciamento de RNA

rRNA - *ribosomal RNA*; RNA ribossomal

SBP - *solute-binding proteins*; proteínas que se ligam a soluto

sRNA - *small RNA*; pequeno RNA

TBP - *TATA binding protein*; proteína que se liga a TATA

THF - tetraidrofolato

TMAO - N-óxido de trimetilamina

tRNA - *transfer RNA*; RNA transportador

TSS - *transcription start site*; sítio de início de transcrição

TSSaRNAs - RNAs associados ao sítio de início de transcrição

UTR - *untranslated region*; região não traduzida

Sumário

1. Introdução	17
1.1. Domínio Archaea	17
1.2. Halobacterium salinarum NRC-1	20
1.3. Diferentes concentrações de sais e seus efeitos nas células	23
1.4. Importância de estudos de transcritômica e proteômica para análise da adaptação à salinidade em H. salinarum NRC-1	25
2. Objetivos	30
2.1. Objetivos específicos	30
3. Materiais e Metodologia	31
3.1. Linhagens e condições de cultivo	31
3.2. Curva de crescimento com glicerol	31
3.3. Extração do RNA	31
3.4. Preparo e sequenciamento das bibliotecas de RNAseq	32
3.5. Análise dos dados de sequenciamento	33
3.6. Identificação de novos transcritos	34
3.7. Integração de dados de transcritômica e proteômica	34
3.8. Ensaio de motilidade	35
3.9. Ensaio de quantificação e dimensão de vesículas de gás	35
3.10. Disponibilidade dos dados	35
4. Resultados	36
4.1. Fenótipo do crescimento em meios com diferentes salinidades	37
4.2. RNAseq: dados e qualidade do sequenciamento	39
4.3. Expressão diferencial em hipersalinidade (5,1 M de NaCl)	40
4.3.1. Genes menos expressos em hipersalinidade	42
4.3.2. Genes mais expressos em hipersalinidade	45
4.3.3. Enriquecimento funcional dos genes diferencialmente expressos em hipersalinidade	49
4.4. Expressão diferencial em meio hiposalino (CM 2,6M de NaCl)	51
4.4.1. Genes menos expressos em hiposalinidade	52
Funções dos genes menos expressos em hiposalinidade	54
4.4.2. Genes mais expressos em hiposalinidade	57
Funções dos genes mais expressos em hiposalinidade	59
4.4.3. Enriquecimento funcional dos genes diferencialmente expressos em hiposalinidade	61
4.5. Suplementação com glicerol interferiu no fenótipo da cultura em hiposalinidade	63

4.6. Comparação com dados de proteômica e transcritômica	66
4.6.1. Comparação do transcritoma mapeado por microarranjos e RNAseq	68
4.6.2 Comparação de dados de RNAseq com proteoma	71
4.7. Motilidade das células é menor em hiposalinidade	73
4.8. Novos transcritos identificados	74
4.9. Quantificação de vesículas de gás	78
5. Discussão	81
5.1. Fenótipo do crescimento da <i>H. salinarum</i> sob estresse osmótico	81
5.2. Reprodutibilidade dos dados de RNAseq e coerência com literatura	83
5.3. Produção de intraRNAs sob estresse osmótico	85
5.4. Enriquecimento funcional dos genes diferencialmente expressos	86
5.5. O proteoma e sua concordância com o RNA-seq	87
5.6. Metabolismo de glicerol está mais expresso em hiposalinidade	88
5.7. Arquelos não se apresentaram funcionais em condições hiposalinas	90
5.8. Alterações no transcritoma sugerem mudanças na maquinaria de tradução	91
6. Considerações finais	93
7. Bibliografia	97
8. Anexos	111

1. Introdução

1.1. Domínio Archaea

Os primeiros microrganismos foram descobertos por Hooke e visualizados por Leeuwenhoek em 1665 e 1683, respectivamente (GEST, 2004). Desde então, foram identificados uma variedade deles, mas muitos dos representantes da vida na Terra permaneceram desconhecidos. Organismos capazes de prosperar em ambientes extremos como altos ou baixos pHs, temperaturas, salinidade e pressão são conhecidos há tempos, porém, somente há cerca de 40 anos foram caracterizados geneticamente (GRIBALDO *et al.*, 2010). Inicialmente, os microrganismos extremófilos foram considerados bactérias por apresentarem características morfológicas semelhantes, como organização celular e outros fenótipos (WOESE; FOX, 1977). Quando o estudo dos extremófilos se aprofundou, no entanto, observou-se que não se tratavam apenas de bactérias e fungos, e sim de um grupo novo de microrganismos composto por muitas espécies que têm como habitat ambientes de extrema salinidade, pH, temperatura ou uma combinação dessas, além de serem as únicas formas de vida capazes de realizar a metanogênese (CAVICCHIOLI, 2011; EME; DOOLITTLE, 2015; “Introduction to the Archaea”,)

As arqueias possuem características dos organismos do domínio Eukarya e Bacteria concomitantemente. Seus mecanismos de transferência de informações são semelhantes aos de eucariotos, a estrutura do cromossomo e plasmídeos são semelhantes aos de bactérias. Ademais, as arqueias possuem característica própria e única: a composição de lipídeos de sua membrana celular é diferente dos outros dois domínios (WOESE; KANDLER; WHEELIS, 1990). Enquanto nos eucariotos e bactérias existe uma ligação éster entre glicerol e ácidos graxos, nas arqueias essa ligação é éter. Isso faz com que as ligações sejam mais estáveis e menos susceptíveis às condições de pH e temperatura extremas nas quais habitam (KOGA, 2012).

Estes microrganismos são encontrados em grande parte do planeta, inclusive dentro do corpo humano (CHIBANI *et al.*, 2020; KIM *et al.*, 2020). Segundo a literatura, até os dias atuais não existe descrição de patogenicidade por esse domínio, mesmo em ambientes como a cavidade oral, pele e microbiota onde podem ser encontrados (CAVICCHIOLI *et al.*, 2003; WIRTH; YOUNG, 2020). No entanto, já foram descritas como espécies oportunistas, encontradas com mais frequência em organismos infectados (DRANCOURT *et al.*, 2020). Um exemplo é a arqueia metanogênica *Methanobrevibacter smithii*, que já foi encontrada em

vaginoses (GRINE *et al.*, 2019b), infecções urinárias (GRINE *et al.*, 2019a), fluidos orais de fumantes (GRINE *et al.*, 2018) e pode ser usada como marcador da presença de doenças inflamatórias no intestino (GHAVAMI *et al.*, 2018). A mesma espécie metanogênica foi identificada no estômago de recém-nascidos, mostrando que a colonização destes microrganismos no estômago humano ocorre logo após o nascimento e sua fonte mais provável é a microbiota intestinal de sua progenitora (GRINE; BOUALAM; DRANCOURT, 2017).

A habilidade de algumas arqueias de sobreviver em diversas condições faz com que haja um grande interesse na bioprospecção de novas enzimas, produtos farmacêuticos, cosméticos, suplementos alimentares e processos biotecnológicos, como proteínas anti-CRISPR e Riboswitches (ALQUÉRES *et al.*, 2007; SWATI*, 2019). Uma das aplicações conhecidas de arqueias é a enzima *Vent*, proveniente da arqueia *Thermococcus litoralis*, uma DNA polimerase empregada em reações de PCR que tem taxa de erro dez vezes menor que a *Taq* polimerase comumente utilizada. Outras arqueias são capazes de suportar processos de branqueamento para a produção de biocombustíveis e muitas proteases, esterases, peptidases e lipases têm interesse para a indústria (SCHIRALDI; GIULIANO; DE ROSA, 2002; ALQUÉRES *et al.*, 2007).

Historicamente, o sistema de classificação que dividiu os organismos em três grupos foi proposto para "*reparar as consequências de construir os sistemas taxonômicos ignorando os possíveis caminhos da evolução microbiana e baseando os estudos na premissa de que a vida é dicotômica (plantas ou animais)*", afirmaram Woese, Kandler e Wheelis em seu artigo descrevendo o domínio Archaea (WOESE; KANDLER; WHEELIS, 1990). Sua crítica remete à maneira como a classificação dos organismos era feita anteriormente - por comparação de características das células como a membrana plasmática e presença de envelope nuclear. Woese *et al.* realizaram análises em níveis moleculares que utilizam o RNA ribossomal (rRNA) como meio de comparação e puderam confirmar a presença do terceiro domínio da vida. De acordo com as análises realizadas, foi possível traçar uma árvore filogenética que, até então, atestou a classificação feita e mostrou que as arqueias estão evolutivamente mais próximas do Domínio Eukarya do que o Bacteria (WOESE; KANDLER; WHEELIS, 1990). A árvore filogenética proposta por Woese *et al.* era composta por dois filos, Crenarchaeota e Euryarchaeota (Figura 1).

Figura 1: Árvore filogenética proposta por Woese, Kandler e Wheelis em 1990. Os números representam grupos de organismos em cada domínio. No domínio Bacteria, estão representados a ordem *Thermotogales* (1); flavobactérias e semelhantes (2); as cianobactérias (3); bactérias púrpuras (4); bactérias gram-positivas (5) e as bactérias verdes não-sulfúricas da classe Chloroflexi (6). No domínio Archaea foram representados dois filos: os Crenarchaeota composto pelo gênero *Pyrodictium* (7) e gênero *Thermoproteus* (8) e o filo Euryarchaeota composto pelos Thermococcales (9), Methanococcales (10), Methanobacteriales (11), Methanomicrobiales (12) e os halófilos extremos (13). No domínio Eucarya são representados os animais (14), os ciliados (15), as plantas verdes (16), os fungos (17), os flagelados (18) e os Microsporidia (19).

Apesar do uso da subunidade menor do rRNA ter revolucionado a classificação dos organismos, somente esse método não esclarece a estrutura da árvore da vida e a extensão da diversidade dos organismos, já que a divergência de sequências pode fazer com que os *primers* comumente usados não amplifiquem o gene e sequências muito diferentes das usuais sejam consideradas artefatos (HUG *et al.*, 2016). Com a evolução da genômica e dos métodos de bioinformática, atualmente é possível fazer a reconstrução de genomas a partir de dados *single cell* e metagenômica. Partindo desse fato, Pohlschroder *et al.* e Hug *et al.* reconstruíram e atualizaram a árvore filogenética proposta por Woese *et al.* (Figura 1) (POHLSCHRODER *et al.*, 2015; HUG *et al.*, 2016), analisando os potenciais metabólicos e uma variedade de sequências importantes para classificação filogenética dos organismos. Com isso, puderam construir uma árvore filogenética mais completa e coerente com os genomas dos organismos. Nessa representação (Figura 2a) pode-se observar que após o LUCA (*last universal common ancestor* - último ancestral comum universal) ocorre a separação em apenas dois domínios, que foram chamados de Bacteria e Arkarya pelos autores. Além disso, esta árvore representa a complexidade da evolução das arqueias, já que nenhum dos filos descritos têm

características em comum, mostrando que, considerando essa visão, talvez seria equivocado agrupá-las em um domínio (WATSON, 2019).

Desde a proposta do terceiro domínio novos microrganismos foram caracterizados e vários grupos de pesquisa publicaram suas versões da Árvore da Vida. Foram adicionados novos filios à organização taxonômica e foram criados os superfilios, compostos por diferentes filios. Atualmente, alguns cientistas baseiam seus estudos na hipótese da presença de apenas dois domínios, Bacteria e a junção de Eukarya e Archaea (Figura 2b). Essa hipótese foi mais aceita pela comunidade após o recente descobrimento das Lokiarchaeas (superfilio Asgard), cujo genoma apresenta genes similares aos encontrados em eucariotos (WATSON, 2019) e levantou debates de que o processo de endossimbiose, que culminou na formação dos eucariotos, teria participação das arqueias (SPANG; CACERES; ETTEMA, 2017). Atualmente, as arqueias são divididas em quatro superfilios: Euryarchaeota, TACK, DPANN e Asgard (EME *et al.*, 2017) e têm 18 filios caracterizados (SCHLOSS *et al.*, 2016).

Figura 2: Representação da relação entre eucariotos e arqueias nos últimos 40 anos. A) Representação em três domínios, Arqueia dividido em dois filios, Crenarchaeota e Euryarchaeota. B) Caracterização de novos filios e descoberta do superfilio Asgard e análises filogenômicas sugerem que os eucariotos se originaram a partir das arqueias Asgard.(EME *et al.*, 2017)

1.2. *Halobacterium salinarum* NRC-1

Halobacterium salinarum é um microrganismo extremófilo do domínio Archaea e superfilio Euryarchaeota (WOESE; KANDLER; WHEELIS, 1990; COKER *et al.*, 2009). Foi isolada de peixes conservados em sal em 1922 e nomeada *Pseudomonas salinaria*. Posteriormente foi apresentada como *Serratia salinaria* e em 1957 recebeu o nome de

Halobacterium salinarum (AHARON OREN, 2005). Seu crescimento é predominante em lagos salinos, áreas de retirada de sal do mar e salinas naturais como o Mar Morto, o *Great Salt Lake* nos Estados Unidos, o Lago Assal em Djibouti e Lago Magadi no Kênia (DASSARMA; COKER; DASSARMA, 2009). No Brasil, já foi identificada em amostras de evaporitos (rochas sedimentares originadas pela evaporação em lagos ou mares rasos) na bacia do Sergipe (VREELAND *et al.*, 2007) e em sedimentos halomarinheiros na Baía de Guanabara, no Rio de Janeiro (CLEMENTINO *et al.*, 2007).

A linhagem utilizada neste trabalho é a *Halobacterium salinarum* NRC-1. Ela possui um cromossomo circular com aproximadamente 2 Mpb e dois replicons, pNRC200 e pNRC100, com 350Kpb e 200Kpb, respectivamente (NG *et al.*, 2000). Apresenta parede celular frágil, o que faz com que tenha várias formas dependendo das condições, sendo o formato de bastonete o mais comum (DYALL-SMITH, 2006). Seu genoma abriga 2.406 regiões codificadoras de genes (PFEIFFER *et al.*, 2019). Seu cultivo é feito em meio de crescimento complexo à 37°C (KAUR *et al.*, 2006) e na natureza a faixa de concentração de NaCl do seu habitat pode variar de 1M à 5,3M (PFEIFFER, 2015; OREN; VENTOSA, 2017). A concentração ótima de crescimento é de 4,3M (BALIGA *et al.*, 2004), o que a torna um microrganismo de fácil cultivo em laboratório e de difícil contaminação das amostras (FACCIOTTI *et al.*, 2010). A coloração da cultura é rosa/vermelha devido às bacterioruberinas presentes e muitos dos lagos salinos no planeta têm coloração avermelhada devido à presença dessas haloarqueias (Figura 3ABC) (DYALL-SMITH, 2006).

Figura 3: Habitat natural, coloração das colônias e micrografias de *H. salinarum*. A) Lagos salinos na Namíbia onde *Halobacterium salinarum* habita. Disponível em: <<http://namibia.ellerstrand.se/scenic-flight-two-cities/>> B) Colônias de *Halobacterium salinarum* (BAXTER, 2018). C) Micrografia eletrônica da célula de *Halobacterium salinarum*, onde é possível observar suas estruturas flagelo-like (arquelos). Barra de escala: 2µm (KINOSITA; NISHIZAKA, 2018) D) Micrografia de *Halobacterium salinarum* NRC-1 onde é possível observar em branco as vesículas de gás produzidas por elas (GUTIERREZ, 2019).

Um dos interesses dos cientistas por esse microrganismo se deve à caracterização de uma proteína presente em sua membrana, a bacteriorodopsina. A bacteriorodopsina funciona como uma bomba de próton movida a luz (SOPPA, 2006) e é expressa em ambientes com pouco oxigênio para gerar energia para a sobrevivência da célula e sua estrutura estável faz com que seja resistente à degradação química, fotoquímica e térmica tendo potencial aplicação como biomateriais e foto ou biossensores (SINGH; SINGH, 2017). Além da fotofosforilação, também utiliza a fosforilação oxidativa, a fosforilação a nível de substrato e respiração usando o dimetil sulfóxido (DMSO) e N-óxido trimetilamina (TMAO) para obtenção de energia para sobrevivência da célula (HARTMANN; SICKINGER; OESTERHELT, 1980; MÜLLER; DASSARMA, 2005; COKER *et al.*, 2007), as últimas duas de forma anaeróbica.

Ademais, a *Halobacterium salinarum* NRC-1 dispõe dos genes *gvp*, que expressam proteínas formadoras de vesículas de gás. As vesículas de gás fazem com que as células

sejam capazes de migrar verticalmente para regiões mais oxigenadas e são formadas por um nano compartimento proteico oco preenchido com oxigênio (WALSBY, 1994) (Figura 3D). Estudos de aplicabilidade das vesículas de gás em diversos processos biotecnológicos têm sido feitos. Entre eles, a apresentação de antígenos e biodisponibilização de vacinas (PFEIFER *et al.*, 2009; SINGH; SINGH, 2017) assim como a implementação dessas proteínas como agentes de contraste para ultrassom e ressonância magnética (LAKSHMANAN *et al.*, 2017; FARHADI *et al.*, 2018). O estudo sobre as haloarqueias é muito importante também na pesquisa sobre a existência de vida em outros planetas (astrobiologia), já que conseguem viver em ambientes tão extremos e têm mecanismos desenvolvidos para lidar com esses estresses (LANDIS, 2001; KISH *et al.*, 2009; LEUKO *et al.*, 2015) .

1.3. Diferentes concentrações de sais e seus efeitos nas células

O habitat natural de *H. salinarum* NRC-1 além de extremo é muito dinâmico (KENNEDY *et al.*, 2001): flutuação da salinidade, radiação e temperatura são condições que a população enfrenta para sobreviver (COKER *et al.*, 2007). A salinidade dos lagos varia sazonalmente e é maior durante o verão, quando a evaporação é alta e os níveis de água são baixos. No inverno e na primavera, o aumento das chuvas aumenta o nível de água nos lagos e diminui as concentrações de sais (CAMPBELL, 1995).

Variações na salinidade do meio podem influenciar tanto a estrutura das células, quanto nas funções das biomoléculas nelas presentes. Quando as células são submetidas a meios com alta concentração de sais, a água presente no meio intracelular tende a se mover para o exterior por diferenças na pressão osmótica. O mesmo ocorre quando as células são submetidas a meios com baixa concentração de sais, a água entra na célula e pode até romper a membrana e inviabilizá-la (NELSON, 2013).

Duas estratégias para lidar com o desbalanço osmótico já foram descritas para os microrganismos extremófilos. A primeira é o “*salt-in*”, que envolve o acúmulo de íons inorgânicos no citoplasma na mesma concentração do exterior. A segunda é a estratégia do “*solute compatível*” em que a célula acumula solutos orgânicos em seu citoplasma (MEI *et al.*, 2017). Para sua sobrevivência, *H. salinarum* NRC-1 usa a estratégia “*salt-in*” em que mantém altos níveis de cloreto de potássio (KCl) internos, podendo chegar a 5M (DEOLE *et al.*, 2013), para contrabalancear o soluto extracelular. A presença dos íons cloreto é um fator essencial para resistência do microrganismo à dissecação e a radiação ionizante (KISH *et al.*, 2009).

Utilizando a estratégia “salt-in”, *H. salinarum* transporta para seu citoplasma altas concentrações de sais, o que poderia levar ao efeito *salting-out* nas proteínas citoplasmáticas, caracterizado pela diminuição das interações da camada de solvatação e aumento das interações da água com os íons dos sais. Este processo causa uma perda de solubilidade, estabilidade e precipitação, o que pode gerar também perda de função das proteínas e enzimas presentes (ARAKAWA; TIMASHEFF, 1982; VAUCLARE *et al.*, 2015). Para evitar a perda de função de proteínas e enzimas durante um desbalanço osmótico suas proteínas são adaptadas à alta salinidade, apresentando ponto isoelétrico (pI) médio de 4,9. As cargas negativas encontram-se distribuídas na superfície, que mantém a força iônica e a camada de solvatação que estabiliza monômeros e polímeros, mantendo a solubilidade e funções das proteínas nessas condições (KENNEDY *et al.*, 2001; VAUCLARE *et al.*, 2015, 2020).

Dos estresses ao qual um microrganismo pode ser submetido, a salinidade, por ser tão severa, induz a produção de proteínas relacionadas ao estresse em geral em bactérias (IM *et al.*, 2013). Isso implica que sob estresse salino, genes relacionados ao estresse pelo calor e o oxidativo podem também ser ativados, fazendo com que a célula se proteja simultaneamente dos três fatores através da produção de diferentes proteínas e metabólitos (BESTEN *et al.*, 2009).

O acoplamento das respostas aos estresses se deve à capacidade dos microrganismos de se adaptar às inter-relações temporais que ocorrem há milhares de anos. A evaporação da água no habitat causada pela alta incidência de sol, por exemplo, faz com que haja um aumento na concentração de sais, aumento na exposição à radiação e altas temperaturas. Já a incidência de chuvas faz com que haja uma diminuição da concentração de sais e menores níveis de radiação e temperaturas. Os experimentos em laboratório tentam manter as variáveis das culturas constantes enquanto são aplicados outros tratamentos, mas milhões de anos de evolução fazem com que as respostas acopladas sejam observadas (BALIGA, 2008)

Outros microrganismos são capazes de sobreviver no mesmo ambiente de *H. salinarum* NRC-1, como é o caso da alga *Dunaliella salina*. Para sobreviver à pressão osmótica, essa alga utiliza a estratégia de “soluto compatível” e produz glicerol ou elimina-o para contrabalancear o sal externo. Uma enzima essencial para o metabolismo do glicerol nessa alga é a glicerol-3-fosfato desidrogenase (CHEN; JIANG; WU, 2009; WILLIAMS *et al.*, 2017a). A glicerol-3-fosfato desidrogenase atua na conversão reversível de diidroxiacetona fosfato, que veio da glicólise, em glicerol-3-fosfato que é convertido em glicerol que pode ser mantido dentro da célula ou secretado (CHEN; JIANG; WU, 2009). Quando ocorre a liberação de

glicerol no meio, os outros microrganismos presentes no habitat, entre eles, *H. salinarum* NRC-1, cuja membrana é permeável a glicerol podem utilizar essa fonte de carbono tanto para produção de energia (glicólise e gliconeogênese) e produção de fosfolipídios de membrana já que têm presentes em seu citoplasma as enzimas glicerol quinase e também a glicerol-3-fosfato desidrogenase (WILLIAMS *et al.*, 2017a).

1.4. Importância de estudos de transcritômica e proteômica para análise da adaptação à salinidade em *H. salinarum* NRC-1

Como dita o Dogma Central da Biologia Molecular, as informações de um organismo são armazenadas no DNA, transcritas em RNA mensageiro e então traduzidas em proteínas que irão exercer suas funções no organismo (CRICK, 1970). Diversas formas de regulação das diferentes camadas de processamento de informação são conhecidas atualmente (Figura 4). O processo de regulação da transcrição pode acontecer no início do processo com interação de fatores de transcrição com promotores, organização em operons de procariotos, além do controle da elongação e terminação de um transcrito. A regulação pós-transcricional é feita pelos sRNAs (*small RNAs*), RBPs (*RNA-binding proteins*), riboswitches, e atua na meia-vida e interação dos RNAs com diversas proteínas, influenciando o início da tradução (PICARD *et al.*, 2009; VAN ASSCHE *et al.*, 2015). Já a regulação pós-traducional é caracterizada pela presença de cofatores que doam os grupos utilizados nas modificações e as modificações: metilação, acetilação, fosforilação, glicosilação e ubiquitinação das proteínas, entre outras (NELSON, 2013; MACEK *et al.*, 2019).

Figura 4: Esquema representando as várias camadas de processamento de informação genética conhecidas. Adaptado de (ENGSTROM; PFLEGER, 2017).

Estudos vêm mostrando uma baixa correlação entre a abundância de mRNAs e proteínas, macromoléculas cujos níveis passam por uma regulação fina para atender às necessidades das células (PICARD *et al.*, 2009). Isto mostra, a importância dos processos pós-transcricionais para entendermos melhor o funcionamento dos organismos. Ela pode acontecer de várias maneiras: promovendo a degradação ou processamento do RNA, estabilização de um transcrito, modificação e edição do RNA (BONNEAU *et al.*, 2007; RANDAU *et al.*, 2009). Esses processos envolvem não só maquinaria proteica como as ribonucleases e chaperonas de RNA como também, os riboswitches, as taxas de tradução e degradação do RNA e muitos ncRNAs (ARRAIANO, 1993; PICARD *et al.*, 2009; GOTTESMAN, 2017; HOLMQVIST; VOGEL, 2018). Alguns dos exemplos de ncRNAs amplamente conhecidos são os rRNAs e tRNAs, envolvidos na maquinaria de tradução. O advento de tecnologias para mapear o transcrito têm contribuído para a descoberta de diversos tipos de ncRNAs (ABDELMOHSEN, 2016), assim como seu importante papel na regulação gênica (STAV *et al.*, 2019).

O estudo do transcrito pode ser feito através de microarranjos de DNA e do RNA-seq. São técnicas com princípios diferentes que irão permitir a quantificação dos RNAs contidos numa determinada amostra. A tecnologia do RNA-seq tem permitido a quantificação mais precisa dos níveis relativos de RNAs além da identificação de novos transcritos, inícios de transcrição alternativos e sítios de processamento dos RNAs (LOWE *et al.*, 2017).

Em *H. salinarum*, o perfil transcricional em diferentes concentrações de NaCl (4,3M, 2,9M e 5M de NaCl) foi realizado utilizando a técnica de microarranjos de DNA (*microarray*) (Coker *et al.*, 2007). Em hiposalinidade, 53 genes apresentaram expressão diminuída em relação à concentração padrão (4.3M) e 143 foram mais expressos. Em hiposalinidade, foi possível observar aumento dos níveis relativos de genes relacionados ao transporte de potássio e peptídeos, para manutenção das concentrações do metal e minimizar a perda de aminoácidos ionizados. Os genes relacionados ao transporte de prótons, as subunidades da ATP sintase, glutamina sintetase, xantina/uracila permease, tiorredoxina e transporte de sódio e cobalto foram menos expressos em hiposalinidade. O transcrito mostrou um aumento na expressão de fatores de transcrição e reguladores, além da expressão diminuída do gene *dnaK*, uma chaperona. Genes como *hsp* (heat shock protein), *csp* (cold shock protein), *sod* e *gst*, relacionados a adaptação ao estresses de temperatura e estresses oxidativo e radiação foram mais expressos, reforçando o fato do estresse osmótico desencadear respostas ao estresse diversas. Em hipersalinidade, 29 genes estavam menos expressos e 32 estavam com

expressão aumentada. Observou-se a diminuição na expressão dos mecanismos de transporte de fosfato e enxofre, que podem ser tóxicos em altas concentrações. Os genes cuja expressão estavam aumentadas eram relacionados ao transporte de nitritos e nitratos e malato desidrogenase, carboxipeptidase, carbamoyl fosfato sintase, aspartato carbamoyl transferase e *long chain fatty acid CoA ligase*, relacionados ao metabolismo de lipídeos, turnover de proteínas e biossíntese de pirimidinas. Em resumo, os resultados dos microarranjos de DNA mostraram que o metabolismo das células foi reestruturado para lidar principalmente com o transporte de íons, fosfato e enxofre e manutenção dos níveis iônicos intracelulares (COKER *et al.*, 2007) .

Uma das maneiras de saber o destino de um mRNA é descobrir se a proteína referente a ele está sendo produzida. Para isso podem ser usadas técnicas específicas para identificar cada proteína ou pode ser feito o proteoma, a identificação de todas proteínas sendo produzidas em certo momento da célula (ANDERSON; ANDERSON, 1998; JOO; KIM, 2005).

Em 2009, Leuko *et. al.* publicaram o proteoma de *H. salinarum* NRC-1 em diferentes concentrações de NaCl (4,3M, 2,6M e 5,1M). Neste estudo, 309 proteínas foram diferencialmente expressas e foram associadas às mudanças de concentração de NaCl. Em hiposalinidade, 106 proteínas estavam menos expressas e 62 mais expressas em relação à condição de controle. Em hipersalinidade, 75 proteínas estavam menos expressas e 66 estavam mais expressas em relação à condição de controle (LEUKO *et al.*, 2009). Proteínas não essenciais para a sobrevivência das células nas diferentes concentrações de NaCl foram menos expressas, como as arqueínas - que participam na formação do arqueolo, estrutura semelhante aos flagelos das bactérias - em baixas concentrações de sal. Outras proteínas menos expressas em baixa salinidade foram os transportadores de manganês para evitar a toxicidade do metal, enquanto os transportadores de sódio e potássio não foram diferencialmente expressos em nenhum dos tratamentos. As proteínas que atuam na manutenção da forma e função de outras proteínas, como chaperonas e proteases estavam mais expressas tanto em condições de hiposalinidade quanto hipersalinidade (LEUKO *et al.*, 2009). Em hiposalinidade, sete proteínas envolvidas na oxidação de lipídeos estavam mais expressas, indicando a geração de acetil-CoA para manter o ciclo do ácido tricarboxílico e produzir energia. Em relação ao controle de tradução, a maioria das proteínas ribossomais e fatores de alongamento estava com expressão diminuída enquanto proteínas associadas à biossíntese de nucleotídeos se encontravam mais expressas, sugerindo ativação dos processos de reparo de DNA. Em resumo, o estudo do proteoma mostrou o rearranjo do

metabolismo das células para a maior obtenção de energia, manutenção das proteínas que já estavam formadas e reparo de DNA (LEUKO *et al.*, 2009).

Em *H. salinarum* NRC-1, diversos tipos de RNAs foram identificados por nosso grupo de pesquisa: RNAs intergênicos (Koide *et al.*, 2009), RNAs associados ao início de transcrição (tssRNAs) (Zaramela *et al.*, 2014), RNAs sobrepostos a sequências de inserção (Gomes-Filho *et al.*, 2015), RNAs sobrepostos a sequências codificantes (intraRNAs) (ten-Caten *et al.*, 2018) e RNAs antisense (de Almeida *et al.*, 2019).

Figura 5: Esquema representando alguns dos tipos de RNA existentes. Entre eles os RNAs associados ao início de transcrição (TSSaRNA), os RNAs antisense (asRNA) e os RNAs pequenos (sRNA). Adaptado de CATEN, (2017).

Os RNA antisense (asRNA, Figura 5) são pequenas moléculas que têm função regulatória e reconhecem a sequência de elementos estruturais em mRNAs alvo (SABERI *et al.*, 2016). Esses RNAs podem atuar aumentando a estabilidade de seu mRNA alvo ou aumentar a degradação do alvo pela formação de RNA fita dupla, alvo de RNases (LASA *et al.*, 2011; SABERI *et al.*, 2016; DE ALMEIDA *et al.*, 2019). Nosso grupo de pesquisa identificou 846 asRNAs no transcrito da *H. salinarum* NRC-1 (DE ALMEIDA *et al.*, 2019).

Outro tipo de RNA é o intraRNA, RNA sobreposto a região codificante, com início de transcrição dentro de uma sequência codificadora (CDS, Figura 5) (TEN-CATEN *et al.*, 2018). Os intraRNAs foram identificados através da presença de inícios de transcrição internos (iTSS) nos genes, dentre eles os *kef1*, um antiporter cátion:próton (*VNG_RS07995*). Através do método de Western blotting, foi possível observar que esse intraRNA também era traduzido e gerava isoformas mais curtas das proteínas. Essa diversidade de transcritos produzida mostra a complexidade da informação genômica nos procariotos, podendo produzir tanto diferentes

transcritos com funções variadas quanto diferentes proteínas (TEN-CATEN *et al.*, 2018). Essas descobertas são interessantes pois a proteína Kef1 e sua isoforma poderiam estar relacionadas com a adaptação à salinidade.

Observando os resultados de proteoma e transcriptoma em diferentes salinidades, é possível observar que o aumento de RNA e aumento de proteínas não está correlacionado, indicando uma possível camada de regulação pós-transcricional. Neste trabalho, utilizamos a metodologia de RNA-seq para o estudo do transcriptoma de *H. salinarum* em resposta à salinidade. Dada a importância da adaptação a este estresse nesta arqueia halofílica, a utilização de metodologia de sequenciamento auxilia não só na identificação de genes diferencialmente expressos com maior precisão como também permite a identificação de novos RNAs envolvidos nesta resposta. A comparação sistemática entre os dados publicados e obtidos neste trabalho podem indicar processos afetados pelo estresse osmótico e novos transcritos que podem ter participação importante na adaptação da *H. salinarum* NRC-1 às condições de salinidade às quais é submetida em seu habitat natural.

2. Objetivos

Identificar expressão diferencial e novos transcritos em amostras de *Halobacterium salinarum* NRC-1 crescendo em meios de cultura com diferentes concentrações de cloreto de sódio, que mimetizam as condições ambientais em que o microrganismo normalmente se encontra.

2.1. Objetivos específicos

- Detectar genes diferencialmente expressos em diferentes condições de salinidade na mesma fase de crescimento.
- Identificar novos transcritos importantes para adaptação à salinidade.
- Verificar processos biológicos diferencialmente representados em cada um dos tratamentos.
- Identificar possíveis regulações pós-transcricionais importantes para adaptação à salinidade através da integração dos dados de proteômica e transcritômica.

3. Materiais e Metodologia

3.1. Linhagens e condições de cultivo

Halobacterium salinarum NRC-1 foi cultivada em meio CM (250 g/l NaCl, 20 g/l MgSO₄, 2 g/l KCl, 3 g/l Citrato de Sódio, 10 g/l Peptona Bacteriológica Oxoid)((ROBB; DASSARMA; FLEISCHMANN, 1995) líquido ou em placas de petri (adicionando 1,5% de ágar). Para as curvas de crescimento e amostras para extração de RNA, foram selecionadas colônias de *H. salinarum* NRC-1 em placas de Petri e foi feito o pré-inóculo dessas colônias em tubos de ensaio contendo 5 mL de CM (sob luz e agitação de 160 rpm). Após cinco dias, as culturas foram diluídas em triplicata para densidade óptica (D.O.) de 0,05 (a 600 nm) em meios CM com 4,3 M NaCl (controle), 2,6 M NaCl (hiposalino) e 5,1 M NaCl (hipersalino), assim como descrito por Leuko *et al.* (2009). Iniciou-se então uma curva de crescimento em meio CM à 37°C, sob luz e agitação de 160 rpm, em que a D.O. foi medida aproximadamente a cada 24 horas no Ultrospec 10 Cell density meter (Amersham Biosciences). A curva de crescimento foi feita durante duas semanas. Quando as amostras atingiram a D.O. de 0,5, foram recolhidos 20 ml de cultura e aliquotados em tubos eppendorf de 2 ml e centrifugados a 37°C e 10.000 g por cinco minutos. Após a retirada do sobrenadante, foram adicionados mais 2 ml de cultura a cada tubo e centrifugado novamente, retirando o sobrenadante. Nessa etapa os *pellets* podem ser armazenados no freezer -80°C para uso posterior.

3.2. Curva de crescimento com glicerol

Para a curva de crescimento com o glicerol utilizou-se os dados de Chen *et al.* (2009) que mostram que a *Dunaliella salina* produz em média 30 mg/ml em variadas salinidades e em concentração de 2M de NaCl tem sua produção máxima de aproximadamente 60 mg/ml (CHEN; JIANG; WU, 2009). Os meios CM 4,3 M de NaCl e 2,6M de NaCl foram suplementados com 30 mg/ml e 60 mg/ml de glicerol e foram feitas curvas de crescimento conforme descrito no item 3.1 com medidas de densidade óptica a cada 3 horas. Foram feitas triplicatas de cada condição.

3.3. Extração do RNA

A extração do RNA foi feita utilizando o protocolo adaptado do kit mirVana™ miRNA Isolation kit (Ambion) que consiste na extração orgânica do RNA utilizando fenol-clorofórmio ácido seguida por uma precipitação com etanol e acetato de sódio 3M. Quatro extrações com o fenol-clorofórmio foram feitas para cada amostra e ao final do processo fez-se um PCR

utilizando *primers* que se anelam ao gene da ATPase AAA da família CDC48 (VNG_RS06465, Forward: CCC GAG AAG TTC ACG CGG, Reverse: GCT GCC CTG GCG GCT GTT), mostrando que as amostras estavam livres de contaminação com DNA tornando desnecessário um tratamento com DNase. Foram feitos géis de agarose 1% utilizando o Sybr Safe (Invitrogen) marcador Gene Ruler 1kb (Thermo Fisher Scientific), tampão de amostra 2X (Ambion) e foi submetido a corrente de 100V por aproximadamente 30 minutos com o intuito de verificar a integridade do RNA e observar o resultado do PCR. Também foi utilizado o Bioanalyzer 2100 Expert (Agilent Technologies) utilizando o kit Agilent RNA 6000 Nano para obtenção do número de integridade do RNA (RIN), que deve ser maior que 8 para início do processo de preparo das bibliotecas para o sequenciamento. Foram feitas duplicatas de cada condição (controle, hipersalino e hiposalino) totalizando seis amostras.

3.4. Preparo e sequenciamento das bibliotecas de RNAseq

Para a amostra ser submetida ao RNA-seq, foi depletado o RNA ribossomal utilizando o kit RiboZero (Illumina). O kit RiboZero faz a retirada do rRNA através da ligação deste a *beads* magnéticos revestidos por estreptavidina que se ligam a fragmentos biotinilados complementares ao rRNA e assim é possível sua retirada da amostra. Em seguida iniciou-se o protocolo do Kit Illumina TruSeq Stranded mRNA com 330 ng de RNA para cada amostra. Neste kit o mRNA foi fragmentado usando cátions divalentes em contato com altas temperaturas e a partir dos fragmentos foi produzida a primeira fita de DNA complementar. A partir do cDNA, uma segunda fita de DNA foi produzida, cujo preparo foi feito com dUTPs no lugar de dTTPs, diferente da reação anterior. Foi feito o processo de adenilação das extremidades 3' dos fragmentos para que os adaptadores fossem ligados com sucesso e a reação de ligação dos adaptadores 5' e 3' aos fragmentos aconteceu. Para a identificação de cada biblioteca foram adicionadas sequências índices à cada uma delas para ser possível o reconhecimento dos transcritos, que foram colocados juntos na mesma célula de sequenciamento. Em seguida, foi feito o preparo das amostras para a etapa final do sequenciamento, onde as amostras com diferentes índices foram misturadas em concentrações iguais, diluídas a 13 pM e aplicadas no sequenciador.

O controle de qualidade do DNA obtido das bibliotecas foi feito com o Bioanalyzer 2100 Expert (Agilent Technologies) utilizando o kit Agilent High Sensitivity DNA e Qubit 4 Fluorometer (Thermo Fisher Scientific) com o kit Qubit™ dsDNA HS Assay e o sequenciamento das bibliotecas foram feitos no aparelho MISEq 03081 (Illumina) no Hemocentro da Universidade de São Paulo - Campus Ribeirão Preto. Para o RNA-seq foi

utilizado o kit MiSeq Reagent Kit v2 (Illumina) de 50 ciclos que foram sequenciados de modo pareado (*paired-ended*). Foram sequenciadas 6 bibliotecas, duas réplicas biológicas para cada tratamento.

3.5. Análise dos dados de sequenciamento

Os arquivos fastq gerados pelo sequenciamento foram submetidos à análise *frtc* (From Raw to Coverage), uma pipeline automática criada para fazer a análise de RNA-seq de procariotos (TEN-CATEN *et al.*, 2018). A pipeline incorpora programas que retiram as sequências dos adaptadores e cortam fragmentos de baixa qualidade (Trimmomatic v0.36, BOLGER; LOHSE; USADEL, 2014), o alinhador Hisat2 v2.1.0 (PERTEA *et al.*, 2016), um visualizador de arquivos de alinhamento (Samtools v1.7, LI *et al.*, 2009), MMR default version (KAHLES; BEHR; RÄTSCH, 2016), que seleciona reads com alinhamentos múltiplos e escolhe o melhor local de alinhamento através de um score de alinhamento. Além disso, apresenta os softwares Deeptools v3.1.1 (RAMÍREZ *et al.*, 2016) e Bedtools v2.26.0 (QUINLAN; HALL, 2010), que geram os arquivos de cobertura do sequenciamento.

Para o controle de qualidade e análise de expressão diferencial foi feita uma adaptação do protocolo proposto no trabalho de SILVA *et al.* (2019) cuja pipeline foi disponibilizada no repositório [ccrescentus_RNASeq_analysis](#). Dentre os processos presentes na pipeline estão o controle de qualidade, análise de expressão diferencial e confecção de gráficos e tabelas para a visualização dos dados. O controle da qualidade dos *reads* do sequenciamento foi feito com o pacote Rqc (v 1.18.0) do Bioconductor (SOUZA; CARVALHO; LOPES-CENDES, 2018). Foi feita uma análise de expressão diferencial através de análises de *scatter plot* de cada uma das réplicas dos tratamentos e interseção dos resultados para cada réplica. Foram geradas matrizes de contagem de *reads* utilizando três anotações: a anotação do genoma de referência (NG *et al.*, 2000), a anotação de sequências de inserção (PFEIFFER *et al.*, 2019) e a anotação de regiões antisense (DE ALMEIDA *et al.*, 2019). Ao final da análise obteve-se o valor M para cada um dos genes: $M = \log_2(\#tratamento/\#controle)$, onde o “*#tratamento*” refere-se ao número de *reads*, normalizado pelo total sequenciado, por gene nas amostras cultivadas em hiposalinidade ou hipersalinidade e “*#controle*” refere-se ao número de *reads* por gene das amostras cultivadas em condição controle (CM 4,3M de NaCl). Foram considerados diferencialmente expressos os genes que apresentaram valor de $|M| \geq 1$ em ambas as réplicas.

Usando o software Panther, fez-se o enriquecimento do termos de ontologia genética (MI *et al.*, 2019; ALEXA; RAHNENFUHRER, 2020). Através do website [Gene Ontology](#) foi

gerada uma lista de referência para o organismo em questão e realizado o teste de Fisher para o enriquecimento funcional dos genes diferencialmente expressos. O KeggMapper também foi utilizado para verificar quais vias estariam sendo diferencialmente expressas e a importância dos genes/proteínas para o funcionamento da via (KANEHISA; SATO, 2020).

3.6. Identificação de novos transcritos

A identificação de transcritos foi feita manualmente através da análise dos arquivos de cobertura contendo *reads* que alinharam uma única vez e cuja quantidade foi normalizada. As regiões foram escolhidas se contivessem cobertura de no mínimo 10 *reads* e foram delimitadas de acordo com os *reads* adjacentes. O método utilizado para o preparo das bibliotecas não é específico para identificação da extremidade dos transcritos, portanto as regiões foram delimitadas de acordo com os *reads* próximos e considerando as duplicatas. Foi utilizado o software IGV - *Integrative Genomics Viewer* e foram geradas tabelas de transcritos não anotados em trabalhos prévios do grupo de pesquisa (DE ALMEIDA *et al.*, 2019).

Além desses dados, foi feita uma matriz de contagem para regiões antisense aos genes anotados no genoma de referência de *Halobacterium salinarum* NRC-1 (*Assembly*: ASM680v1). Com o arquivo de anotação em mãos, o sentido da fita foi modificado em todos os genes, fazendo com que as contagens fossem referentes à regiões antisense ao gene. A partir das contagens foi calculado o valor $N = \log_2(\#antisense/\#sense)$, onde o “*#antisense*” refere-se ao número de *reads* antisense, normalizado pelo total sequenciado, por gene nas amostras cultivadas em hiposalinidade ou hipersalinidade e “*#sense*” refere-se ao número de *reads* por gene das amostras cultivadas em hiposalinidade ou hipersalinidade. Destes foram selecionados os transcritos cujo valor nas em ambas as réplicas eram $|N| \geq 1$.

3.7. Integração de dados de transcritômica e proteômica

As tabelas contendo os dados de transcritômica (COKER *et al.*, 2007) e proteômica (LEUKO *et al.*, 2009) retiradas das publicações foram comparadas aos dados obtidos neste trabalho e a visualização dos dados foi produzida no RStudio utilizando os pacotes ggplot2 (v3.3.2) e preprocessCore do Bioconductor (v1.48.0). Houve o contato com os autores das publicações para a obtenção dos dados brutos já que, por serem dados antigos, não estavam depositados em bancos de dados públicos. Quando obteve-se resposta dos autores, também não foi possível a disponibilização dos dados, o que nos levou a usar somente as listas de genes e proteínas presentes nas publicações.

3.8. Ensaio de motilidade

O ensaio de motilidade foi feito como descrito anteriormente (ZARETSKY *et al.*, 2019): 10 µl de cultura líquida com D.O. na fase logarítmica de crescimento e fase estacionária foram inoculados no centro de uma placa de petri com CM semi-sólido 0,3% (m/v) de ágar. As placas foram incubadas por sete dias a 37°C e depois o diâmetro do halo de motilidade foi medido. Quando o halo era deformado, foi feita a média entre o raio menor e o raio maior.

3.9. Ensaio de quantificação e dimensão de vesículas de gás

O ensaio de quantificação de vesículas foi feito usando o protocolo adaptado de preparação de vesículas de gás como agentes de contraste para ultrassom e ressonância magnética (LAKSHMANAN *et al.*, 2017). O isolamento e purificação das vesículas de gás foi feito coletando 600µl das células da superfície de culturas incubadas em temperatura ambiente sem rotação por 48 horas. Após a coleta, adicionou-se 600µl de tampão de lise TMC (10 mM Tris-HCl, 2,5 mM MgCl₂ e 2 mM de CaCl₂) e foi feita a centrifugação a 300g por 4 horas a 4°C. Após a centrifugação, na superfície do líquido do tubo estarão as vesículas de gás e células não lisadas contendo vesículas de gás. Essa camada é visível e pode ser branca ou com coloração rosada, dependendo da quantidade de células não lisadas. As vesículas de gás e células não lisadas foram transferidas para outro tubo e foram adicionados 1,6 mililitros de tampão PBS 1X. A solução foi centrifugada novamente nas mesmas condições e o sobrenadante foi retirado e armazenado em tubo novo. Esta solução foi armazenada sob refrigeração. Foram tomados cuidados para evitar a explosão das vesículas de gás, como o uso de tubos de rosquear e cuidado contra choques mecânicos e congelamento. Para quantificação e morfologia das vesículas de gás, as soluções de vesículas de gás purificadas foram diluídas em água e aplicadas ao equipamento NanoSight NS300, que faz através da análise de imagens e lasers faz a análise de rastreamento de nanopartículas. O experimento foi realizado em colaboração com o Prof. Dr. Fausto Bruno dos Reis Almeida e Dra. Veronica Soares Brauer.

3.10. Disponibilidade dos dados

Os resultados do sequenciamento foram submetidos ao SRA (Sequence Read Archive) sob os códigos: SRX9424053 e SRX9424056 para amostras em salinidade controle, SRX9424052 e SRX9424055 para amostras em hiposalinidade, e SRX9424054 e SRX9424057 para amostras em hipersalinidade (Tabela 1).

Tabela 1: Números de acesso e nome das amostras e bioprojeto no [website do SRA](#).

Accession	Run	Bioproject Accession	Biosample Accession	Nome da amostra
SRX9424052	SRR12971719	PRJNA674319	SAMN16634764	<i>Halobacterium salinarum</i> NRC-1; mid-log phase; 2.6M of NaCl; biological replicate A
SRX9424053	SRR12971718	PRJNA674319	SAMN16634765	<i>Halobacterium salinarum</i> NRC-1; mid-log phase; 4.3M of NaCl; biological replicate A
SRX9424054	SRR12971717	PRJNA674319	SAMN16634766	<i>Halobacterium salinarum</i> NRC-1; mid-log phase; 5.1M of NaCl; biological replicate A
SRX9424055	SRR12971716	PRJNA674319	SAMN16634767	<i>Halobacterium salinarum</i> NRC-1; mid-log phase; 2.6M of NaCl; biological replicate B
SRX9424056	SRR12971715	PRJNA674319	SAMN16634768	<i>Halobacterium salinarum</i> NRC-1; mid-log phase; 4.3M of NaCl; biological replicate B
SRX9424057	SRR12971714	PRJNA674319	SAMN16634769	<i>Halobacterium salinarum</i> NRC-1; mid-log phase; 5.1M of NaCl; biological replicate B

4. Resultados

Nesta seção serão identificados os genes diferencialmente expressos em condições de hipo e hipersalinidade, assim como o perfil de crescimento das culturas de *H. salinarum* NRC-1. Serão apresentados os esforços para a identificação de novos transcritos e para a caracterização funcional dos processos biológicos diferencialmente representados utilizando duas ferramentas: o KeggMapper e o Panther.

Por fim, será feita a comparação entre os dados de transcritômica encontrados neste estudo com trabalhos de transcritômica e proteômica publicados anteriormente e a discussão da relação entre transcritos e proteínas que podem sofrer algum tipo de regulação pós-transcricional.

4.1. Fenótipo do crescimento em meios com diferentes salinidades

Utilizando as mesmas condições descritas por Leuko et. al. (2009), fez-se uma curva de crescimento com a cultura de *H. salinarum* NRC-1 em três salinidades diferentes (Figura 5). Observou-se que o crescimento no meio contendo 5,1 M de NaCl foi semelhante ao do controle (CM 4,3M) enquanto o crescimento em meio hiposalino (2,6M de NaCl) fez com que diminuísse a taxa de crescimento da cultura.

A observação da variação da D.O. das culturas durante 14 dias (Figura 6A) mostrou que nas primeiras 25 horas o comportamento das culturas 2,6M e 5,1M de NaCl são semelhantes. O mesmo pode ser observado quando é feita a diferença entre as densidades ópticas dos tratamentos e controle (Figura 6B). Nas primeiras 48 horas da curva de crescimento, a cultura hipersalina teve um comportamento semelhante ao controle, enquanto a cultura hiposalina apresentou uma diferença negativa, indicando lentidão no crescimento e possível adaptação à condição estressante. Já as fases exponencial (47 a 94 horas) e estacionária (118 a 336 horas) foram distintas para cada uma das condições (Figura 6A). A diferença entre tratamento e controle foram mais negativas na cultura hiposalina, mostrando que as células têm um fenótipo de crescimento diferente em hiposalinidade (Figura 6B). Ainda que a hipersalinidade também seja uma condição estressante às células, a diferença entre o controle foi menor, mesmo que negativa, podendo indicar que a adaptação a maiores concentrações de sal é mais viável, já que menores concentrações podem trazer o rompimento das células (NELSON, 2013).

Figura 6: Curva de crescimento de *H. salinarum* NRC-1 em meio de cultura com 2,6 mol/l de NaCl (CM 2,6), 4,3 mol/l de NaCl (CM 4,3) e 5,1 mol/l de NaCl (CM 5,1). A) O experimento foi feito em triplicata e as faixas presentes representam os desvios em relação à média de cada medida. Intercessão com a linha pontilhada indica o tempo em que foram coletadas alíquotas para obtenção de RNA. B) Representação da diferença entre a densidade óptica de hiposalinidade e controle (pontos azul claro); e hipersalinidade e controle (pontos azul escuro) em função do tempo (horas).

Após a finalização das curvas de crescimento, as culturas foram deixadas na bancada sem agitação e foi possível observar em hiposalinidade uma coloração mais avermelhada e a falta de camada superficial de células (Figura 7). Este fenótipo é indicativo de ausência de vesículas de gás (OREN, 2012). Os possíveis motivos serão abordados no decorrer deste texto.

Figura 7: Culturas de *H. salinarum* NRC-1 após 24 horas sem agitação as culturas de *H. salinarum* NRC-1 em condição controle (A) e hiposalinidade (B).

4.2. RNAseq: dados e qualidade do sequenciamento

Alíquotas das culturas de *H. salinarum* foram retiradas quando atingiram DO 0,5 nos pontos 47 horas (CM 4,3M e CM 5,1M) e 72 horas (CM 2,6M) em duplicata. O RNA foi extraído e uma verificação da presença de DNA foi feita através de amplificação utilizando primers de regiões do DNA. Nenhum DNA foi amplificado nas amostras, indicando a ausência de contaminação por DNA. Para iniciar a preparação das bibliotecas de sequenciamento foi preciso avaliar a integridade do RNA. Essa avaliação foi feita utilizando o sistema Bioanalyzer de eletroforese e todas amostras apresentaram o número de integridade do RNA (RIN) maior que 8, o mínimo exigido pelo protocolo de preparo das bibliotecas. O preparo das bibliotecas foi feito usando o kit Illumina TruSeq Stranded mRNA e a qualidade do DNA produzido foi avaliada usando também o sistema Bioanalyzer. A biblioteca foi quantificada e concentrações iguais das seis amostras foram adicionadas ao sequenciador. Após a etapa de sequenciamento foram gerados mais de 18 milhões de *reads* com 98,7% destes identificados através do *index*, gerando em média 16,45% de *reads* para cada amostra, mostrando que as amostras estavam proporcionalmente representadas no sequenciamento. O sequenciamento *paired-ended* gerou *reads* R1 com pontuação média de 35, enquanto os *reads* R2 tiveram

qualidade média de 20. Após o processo de *trimming*, que descarta fragmentos e *reads* de baixa qualidade ($Q < 30$), a média dos *reads* R1 e R2 se manteve próxima a 34, significando que a probabilidade P de designar um nucleotídeo incorreto à sequência é de até 1 em 1000, já que $P = 10^{-Q/10}$.

Para a análise de expressão diferencial foram selecionados apenas os *reads* que alinharam uma vez ao genoma de referência e os que foram reposicionados pelo MMR, que representaram em média 63,80% dos *reads* totais de cada biblioteca (Tabela 2). As amostras em meio hiposalino, hipersalino e condição ótima de cultivo (controle) também foram chamadas de *low salinity*, *high salinity* e *optimal salinity* na montagem das figuras e no decorrer do texto.

Tabela 2: Distribuição dos *reads* entre amostras analisadas.

Bibliotecas	Replicatas	Número de <i>reads</i>	<i>Reads</i> após o <i>trimming</i>	Número de <i>reads</i> alinhados 1x	% de <i>reads</i> alinhados 1x
NRC-1 CM 2,6M NaCl	A	1.414.865	806.295	619.436	76,82
NRC-1 CM 2,6M NaCl	B	1.459.836	1.024.165	613.800	59,93
NRC-1 CM 4,3M NaCl	A	1.829.755	875.451	468.128	53,47
NRC-1 CM 4,3M NaCl	B	1.595.001	825.486	571.965	69,29
NRC-1 CM 5,1M NaCl	A	1.579.838	871.596	494.061	56,68
NRC-1 CM 5,1M NaCl	B	1.274.678	714.134	476.004	66,65

4.3. Expressão diferencial em hipersalinidade (5,1 M de NaCl)

Como controle de qualidade das réplicas foram feitas análises de agrupamento visualizadas através de *heatmaps* e análise de componentes principais (PCA, não mostrados). Observou-se que uma das réplicas se agrupava com a amostra controle enquanto a outra

tinha características diferentes. Para observar a concordância entre as réplicas dos experimentos, foram criados *scatter plots* comparando o valor de M, ($\log_2$ de H (high salinity)/por O (optimal salinity)) de todos os genes para cada uma das réplicas (Figura 8). As réplicas foram obtidas a partir de RNA extraído de colônias diferentes e submetidas ao mesmo tratamento. Por essa razão, era esperado que os valores de M dos seus genes fossem semelhantes e no gráfico da Figura 8 se organizassem de maneira linear demonstrando que genes mais expressos em uma replicata também estavam mais expressos em outra e o contrário também. No entanto, a tendência linear não é observada no gráfico e observa-se um acúmulo de genes no centro dele. A causa desse acúmulo poderia ser a presença de genes que não tinham reads alinhados em uma das réplicas por se tratarem de colônias diferentes e consequentemente com certa variabilidade genética entre si.

Figura 8: *Scatter plot* comparando o valor de M, que representa o $\log_2$ (H (high salinity) / O (optimal salinity)), das duas réplicas para o tratamento hipersalinidade (5,1M NaCl). Os dados completos estão disponíveis na Tabela S7F.

O critério para definir os genes diferencialmente expressos foi selecionar genes que apresentavam o valor de $|M| \geq 1$ em ambas as réplicas. O valor 1 indica que a expressão dos genes estava duas vezes maior, ou menor, que o controle. Quando os critérios foram aplicados

à lista de genes, foram observados 35 genes cujas réplicas tinham valores de $M \leq -1$ e 31 genes cujas réplicas tinham valores de $M \geq 1$, ou seja, estavam menos expressos e mais expressos, respectivamente (Tabela 3 e Tabela 4).

4.3.1. Genes menos expressos em hipersalinidade

Entre os genes menos expressos em hipersalinidade em relação à condição controle (Tabela 3), estão os que codificam proteínas transportadoras do tipo ABC. Essas proteínas apresentam domínio que se liga a ATP e podem transportar várias substâncias desde prótons e metais até compostos orgânicos maiores como oligonucleotídeos e polissacarídeos (WILKENS, 2015). O genoma de *H. salinarum* possui duas cópias do operon *pst* (*phosphate-specific transport* - transporte específico de fosfato), cada uma contendo cinco genes (FURTWÄNGLER *et al.*, 2010). O operon *pst1* é formado pelos genes *VNG_RS09695*, *VNG_RS09700*, *VNG_RS09705*, *VNG_RS09710* e *VNG_RS09715*. Destes, apenas dois estavam diferencialmente expressos (*VNG_RS09705* e *VNG_RS09715*), contudo os outros genes se encontravam menos expressos mas não atingiram o critério para serem classificados como diferencialmente expressos, pois o número de reads de uma das réplicas era nulo. Já o operon *pst2* é formado pelos genes *VNG_RS01755*, *VNG_RS01760*, *VNG_RS01765*, *VNG_RS01770* e *VNG_RS01775*. Desses, três estavam diferencialmente expressos: dois transportadores de fosfato do tipo ABC e uma permease transportadora de fosfato do tipo ABC, também chamada de *pstA* (*VNG_RS01760*, *VNG_RS01775*, e *VNG_RS01765*). Os outros genes do operon também se encontravam menos expressos e, assim como no operon *pst1*, não atingiram o critério para serem classificados como diferencialmente expressos.

Um transportador de glicerol-3-fosfato (*VNG_RS11500*) estava menos expresso em hipersalinidade. Ele é uma lipoproteína que se encontra ancorada à membrana e faz parte da família de proteínas que se ligam a soluto (*solute-binding proteins* - SBP, UniProtKB - Q9HHQ8). Quando há a depleção de fosfato no ambiente, algumas bactérias são capazes de utilizar o glicerol-3-fosfato como fonte de fosfato ("CDD Conserved Protein Domain Family: PBP2_UgpB").

A proteína de estresse universal (*VNG_RS11195*) é um transportador de aminoácidos catiônicos e, normalmente, se encontra mais expressa quando a célula é exposta a diferentes agentes estressantes (UniProtKB - Q9HHY0). No entanto, o comportamento contrário foi observado quando a *H. salinarum* foi submetida ao crescimento em meio hipersalino. Reguladores transcricionais também foram encontrados menos expressos em hipersalinidade. Foram eles os genes *VNG_RS07385*, *VNG_RS00600* e *VNG_RS12865* que codificam as proteínas: TrmB, uma proteína com motivo HTH e ArsC, respectivamente. O gene *arsC* que

codifica a enzima arsenato redutase, atua também na resistência à arsenato através do transporte deste para o meio extracelular (KUNKA *et al.*, 2020) .

O gene *VNG_RS11575* codifica a enzima fosfatase alcalina. A fosfatase alcalina atua na desfosforilação de resíduos orgânicos tendo como produto o fosfato inorgânico. Este pode ser usado na mobilização de outros compostos orgânicos e formação de moléculas energéticas (MURPHY; TIBBITTS; KANTROWITZ, 1995). Sua baixa expressão, juntamente com a baixa expressão de proteínas transportadoras cujo substrato é fosfato indicam correlação entre os dois processos (RAO; TORRIANI, 1990).

Tabela 3: Genes menos expressos em meio hipersalino. Contém o locus, nome do gene, produto, valor de M que representa o $\log_2 (H (high\ salinity) / O (optimal\ salinity))$ e o desvio padrão dos valores de M das replicatas.

<i>Locus</i>	Nome do gene	Produto	M	sdM
VNG_RS09705	<i>pstA</i>	phosphate ABC transporter permease PstA	-4.9	1.4
VNG_RS11500	VNG_RS11500	ABC transporter substrate-binding protein	-4.6	2.5
VNG_RS11575	VNG_RS11575	alkaline phosphatase	-3.3	2.3
VNG_RS09530	<i>argH</i>	argininosuccinate lyase	-2.9	0.1
VNG_RS09715	VNG_RS09715	substrate-binding domain -containing protein	-2.7	1.3
VNG_RS01775	VNG_RS01775	PstS family phosphateABC transporter substrate-binding protein	-2.7	1.6
VNG_RS10035	VNG_RS10035	caspase family protein	-2.5	1.7
VNG_RS01490	VNG_RS01490	shikimate dehydrogenase	-2.4	0.8
VNG_RS02235	VNG_RS02235	cation:proton antiporter	-2.4	0.2
VNG_RS07060	VNG_RS07060	zinc-binding dehydrogenase	-2.3	0.6
VNG_RS07385	VNG_RS07385	TrmB family transcriptional regulator	-2.3	1.2
VNG_RS01760	VNG_RS01760	phosphate ABC transporter ATP-binding protein	-2.2	1.7
VNG_RS09535	VNG_RS09535	argininosuccinate synthase	-2.1	0.6
VNG_RS12865	VNG_RS12865	ArsC family transcriptional regulator	-2.1	0.8
VNG_RS01765	<i>pstA</i>	phosphate ABC transporter permeasePstA	-2.1	1.3
VNG_RS05140	VNG_RS05140	hypothetical protein	-2.1	0.8
VNG_RS04820	<i>aroA</i>	3-phosphoshikimate 1-carboxyvinyl	-2.0	1.1

		transferase		
VNG_RS09325	VNG_RS09325	DUF5518 domain - containing protein	-2.0	0.5
VNG_RS05240	VNG_RS05240	DUF115 domain - containing protein	-1.9	1.1
VNG_RS06745	VNG_RS06745	HtlC	-1.9	1.1
VNG_RS06960	VNG_RS06960	thermonuclease family protein	-1.9	0.5
VNG_RS07205	VNG_RS07205	ABC transporter substrate-binding protein	-1.9	0.4
VNG_RS09890	VNG_RS09890	hypothetical protein	-1.9	0.2
VNG_RS11195	VNG_RS11195	universal stress protein	-1.9	0.5
VNG_1438a	VNG_1438a	uncharacterized protein	-1.8	0.8
VNG_RS04665	VNG_RS04665	HAMP domain -containing protein	-1.8	0.2
VNG_RS02720	VNG_RS02720	helix-turn-helix transcriptional regulator	-1.7	0.9
VNG_RS08435	VNG_RS08435	AAA family ATPase	-1.6	0.8
VNG_RS02350	VNG_RS02350	VOC family protein	-1.5	0.6
VNG_RS05215	VNG_RS05215	sulfatase-like hydrolase/ transferase	-1.5	0.6
VNG_RS00600	VNG_RS00600	hypothetical protein	-1.4	0.3
VNG_RS01105	VNG_RS01105	CoA pyrophosphatase	-1.3	0.3
VNG_RS02725	VNG_RS02725	ABC transporter ATP-binding protein	-1.2	0.1

Os genes *argH* e *VNG_RS09535* codificam a argininosuccinato liase e argininosuccinato sintase, respectivamente. Eles participam da via de biossíntese de arginina, especificamente na conversão de L-ornitina e carbamoyl fosfato à arginina. O mecanismo da síntese de arginina é composto por três passos: o primeiro passo não tem proteína anotada para *H. salinarum* e outros dois passos são realizados pelas duas enzimas diferencialmente expressas, argininosuccinato sintase e argininosuccinato liase (Uniprot:B0R7V5). A utilização de arginina na *H. salinarum* consiste no transporte de arginina extracelular e conversão em ornitina para produção de energia (FALB *et al.*, 2008).

Outro grupo de genes: *VNG_RS02720*, *VNG_RS02725*, *VNG_RS02730*, *VNG_RS02735* e *VNG_RS02740*, também estava menos expresso. Apenas os dois primeiros genes, que codificam um regulador transcricional contendo domínio hélice-volta-hélice capaz de se ligar ao DNA (WINTJENS; ROOMAN, 1996) e um transportador do tipo ABC com domínio de ligação ao ATP, respectivamente, apareceram na lista de diferencialmente

expressos. Os genes seguintes também apresentaram valores de $M \geq -1$ para pelo menos uma das réplicas. São eles duas permeases (*VNG_RS02730* e *VNG_RS02735*) e um gene contendo um domínio de função desconhecida (DUF123) ligado a um domínio de atividade de endonuclease (*VNG_RS02740* - UniProtKB - Q9HRH5).

Outros dois genes contendo domínio de função desconhecida (DUFs) estavam menos expressos em hipersalinidade. O gene *VNG_RS09325* apresenta o domínio DUF5518, cuja característica é a presença de múltiplas regiões transmembrana (UniProtKB - Q9HMU6). Já o gene *VNG_RS05240* tem em sua estrutura o domínio DUF115. Esse domínio é encontrado em arqueias e está presente na via de biosíntese de tetrahydrofolato (THF), mais precisamente na produção de 6-hidroximetil-diidropteridina difosfato (6-HMDP), que por sua vez, faz parte da via de biossíntese de cofatores (UniProtKB - Q9HQ39).

Um dos conjuntos de genes menos expressos é o composto por *VNG_RS04665*, *VNG_RS06745*, *VNG_RS06960* e *VNG_RS07205*. Esses genes codificam proteínas com atividade de histidina quinase, uma ATPase que atua na fosforilação de histidina, termo endonuclease e um transportador ABC, respectivamente. Estes quatro genes estão relacionados à transdução de sinais e quimiotaxia e atuam na adaptação da população ao ambiente, motilidade e a termoendonuclease (*VNG_RS06960*), em especial, tem em sua estrutura o domínio DUF4350, encontrado em enzimas que atuam de desubiquitinação de proteínas.

Dois dos genes que codificam enzimas da via de biossíntese de corismato estavam menos expressas: a shikimato desidrogenase (*VNG_RS01490*) e a 3-fosfo shikimato 1-carboxivinil transferase (*VNG_RS04820*). Na via de biossíntese dos aminoácidos aromáticos é possível observar que os genes menos expressos atuam nas últimas reações de formação do corismato, o que mostra que a síntese de corismato não era necessária para a população no momento em que foram obtidas as amostras para o RNA-seq.

4.3.2. Genes mais expressos em hipersalinidade

Alguns genes com função de reguladores transcricionais foram identificados como mais expressos em hipersalinidade (Tabela 4), entre eles, uma proteína da família TrmB encontrada menos expressa anteriormente, o gene *VNG_RS13010*. Os outros genes mais expressos são fatores de início da transcrição e indicam a transcrição de diferentes genes como resposta ao estresse osmótico. A tradução de alguns genes também parece ter sido afetada já que várias proteínas ribossomais, que atuam na estabilização do ribossomo e se ligam ao rRNA (KORBEINIKOVA; GARBER; GONGADZE, 2012; PILLA; BAHADUR, 2019), são mais expressas. Também é possível observar o aumento da expressão de uma proteína que atua na modificação pós-transcricional do tRNA, a *4-demethyl wyosinesynthase* TYW1

(VNG_RS08250), que é importante na interação códon-anticódon e evita erros na tradução (UniProtKB - Q9HNF3).

Tabela 4: Genes mais expressos em hipersalinidade. Contém o locus, nome do gene, produto, valor de M que representa o $\log_2$ (H (high salinity) / O (optimal salinity)) e o desvio padrão dos valores de M das replicatas.

Locus	Nome do gene	Produto	M	sdM
VNG_RS09970	VNG_RS09970	ABC transporter substrate-binding protein	3.0	2.1
VNG_0613a	VNG_0613a	uncharacterized protein	2.2	0.5
VNG_RS01015	VNG_RS01015	DUF5059 domain -containing protein	2.1	1.1
VNG_RS03080	VNG_RS03080	30S ribosomal protein S3ae	2.1	1.5
VNG_RS12755	VNG_RS12755	hypothetical protein	2.1	1.1
VNG_RS03095	VNG_RS03095	30S ribosomal protein S15	2.0	1.4
VNG_RS09420	VNG_RS09420	DUF2073 domain -containing protein	1.9	0.0
VNG_RS13010	VNG_RS13010	helix-turn-helix domain -containing protein	1.7	0.7
VNG_RS01245	VNG_RS01245	type I 3-dehydroquinate dehydratase	1.6	0.3
VNG_RS12595	VNG_RS12595	hypothetical protein	1.6	0.1
VNG_RS03395	VNG_RS03395	transcription initiation factor IIB	1.5	0.5
VNG_0875C	VNG_0875C	M50 family metalloprotease	1.5	0.1
VNG_RS02135	VNG_RS02135	proteasome assembly chaperone family protein	1.4	0.1
VNG_RS12975	VNG_RS12975	type 2 isopentenyl -diphosphateDelta-isomerase	1.4	0.0
VNG_RS03675	VNG_RS03675	helix-turn-helix transcriptional regulator	1.4	0.2
VNG_RS08250	VNG_RS08250	4-demethyl wyosine synthaseTYW1	1.4	0.1
VNG_RS11725	VNG_RS11725	hypothetical protein	1.4	0.4
VNG_RS12750	VNG_RS12750	hypothetical protein	1.4	0.2

VNG_RS01170	VNG_RS01170	class I SAM-dependent methyl transferase	1.3	0.3
NC_001869.1:58502..59904:plus NA ISH8B-IS4_ssgr_ISH8	ISH8B-IS4_ssgr_ISH8	NA	1.3	0.4
NC_002608.1:79702..81104:plus NA ISH8E-IS4_ssgr_ISH8	ISH8B-IS4_ssgr_ISH8	NA	1.3	0.4
VNG_RS05585	VNG_RS05585	30S ribosomal protein S17e	1.3	0.1
VNG_RS00290	VNG_RS00290	hypothetical protein	1.3	0.2
NC_002607.1:1608078..1609480:plus NA ISH8B-IS4_ssgr_ISH8	ISH8B-IS4_ssgr_ISH8	NA	1.2	0.3
NC_002608.1:30609..32010:minus NA ISH8A-IS4_ssgr_ISH8	ISH8A-IS4_ssgr_ISH8	NA	1.2	0.3
VNG_RS13225	VNG_RS13225	hypothetical protein	1.2	0.1
VNG_RS02880	VNG_RS02880	transcription initiation factor IIB	1.2	0.2
VNG_RS07520	<i>cofG</i>	7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase subunit CofG	1.2	0.1
VNG_RS01845	VNG_RS01845	2-oxoacid:ferredoxin oxidoreductase subunit beta	1.1	0.1
VNG_RS00860	VNG_RS00860	hypothetical protein	1.1	0.1

Um transportador da família ABC (*VNG_RS09970*) foi encontrado mais expresso. Este transportador apresenta um domínio que se liga a ferro e cobalamina (vitamina B12) que é importante para a via de metabolismo de aminoácidos e biossíntese de THF, o que mantém a homeostase celular (ROWLEY; KENDALL, 2019). A importação de vitamina B12, biossíntese de THF e a metilação, estão interligados. O gene *VNG_RS01170*, codifica uma metiltransferase dependente de S-adenosil metionina, que é produzida através de cobalamina (vitamina B12) e THF. A enzima metionina sintase catalisa a reação de formação de metionina

através da doação do grupo metil da metilcobalamina para a homocisteína. A metionina recebe o grupo adenosil de uma molécula de ATP e pode atuar como substrato para a metiltransferase como S-adenosil metionina. Já o N5-metil THF atua como regenerador da metilcobalamina, doando seu grupo metil à cobalamina resultante da reação com a homocisteína (DE CRÉCY-LAGARD *et al.*, 2012; BHAGAVAN; HA, 2015; TJONG; DIMRI; MOHIUDDIN, 2021).

O gene *VNG_RS07520*, que codifica proteínas da via de biossíntese da coenzima F0 estava mais expresso. A coenzima F0 é cromófora e atua na produção de metano em organismos metanogênicos e no reparo de DNA dependente de luz pelas fotoliasas (GLAS *et al.*, 2009). Como o aumento da salinidade do meio no habitat significa maior evaporação da água e maior exposição solar, este resultado pode confirmar que a população celular discerniu o meio e se adaptou a ele.

Foram encontrados mais expressos genes de sequências de inserção, que podem causar variabilidade genética através da interrupção de genes e modificação de expressão (WRIGHT *et al.*, 2017). Além disso, houve a maior expressão dos genes *VNG_0875C* e *VNG_RS0213*. Ambos genes têm como produto uma metaloprotease e a PAC2, uma chaperona de montagem do proteassoma. Ambas proteínas atuam no processo de proteólise e podem indicar que em hipersalinidade o turnover das proteínas e seu período de meia vida sejam afetados.

Os genes mais expressos também tiveram o domínio de função desconhecida (DUFs) representados. O gene *VNG_RS01015* apresenta o domínio DUF5059, que se liga a cobre e pode atuar no processo de transferência de elétrons como uma proteína semelhante ao citocromo (UniProtKB - P82857). Já o gene *VNG_RS09420* tem em sua estrutura o domínio DUF2073. Esse domínio é encontrado em arqueias e não têm função conhecida (Pfam - PF09846).

O gene *VNG_RS01845* foi mais expresso e faz parte do ciclo do ácido tricarboxílico redutivo, que tem o processamento não enzimático e é cotado como metabolismo utilizado pelos primeiros seres autotróficos (KITADAI; KAMEYA; FUJISHIMA, 2017). O gene *VNG_RS01245*, presente na via de biossíntese de corismato estava mais expresso. Foi visto anteriormente que a hipersalinidade fez com que fosse diminuída a síntese de corismato, no entanto a maior expressão desse gene *VNG_RS01245* mostra que a via de corismato foi reajustada para a biossíntese de sideróforos, substâncias que têm papel importante na aquisição de ferro pelo microrganismos (KADI; CHALLIS, 2009). Por último, o gene *VNG_RS12975* tem sua expressão aumentada, indicando que a reação reversível de conversão do isopentenil difosfato a dimetilalil difosfato que pode culminar na biossíntese de isopreno, síntese de terpenoides, carotenoides, N-glicanos e ubiquinona (KEGG hal00900).

4.3.3. Enriquecimento funcional dos genes diferencialmente expressos em hipersalinidade

Para que fosse possível identificar funções e vias metabólicas diferencialmente expressas em hiposalinidade, foi utilizado o Panther, um software de enriquecimento funcional a partir dos termos de ontologia referente aos genes (GO) (MI *et al.*, 2019). Do total de 76 genes diferencialmente expressos, 51 estavam anotados no banco de dados utilizado pelo software. Os genes cuja classificação não foi possível são genes adicionados na última anotação do genoma de *H. salinarum* NRC-1 proposta por Pfeiffer *et. al* (2019) ou sequências de inserção anotados.

Quando analisamos as funções moleculares (Figura 9) os valores do *ratio* foram obtidos através da divisão entre o número de genes encontrados com o GO na lista de genes diferencialmente expressos e o número de genes anotados no genoma de *H. salinarum* com o mesmo GO. Os termos GO que apresentam o *ratio* = 1 indicam que 100% dos genes que apresentam um GO estavam diferencialmente expressos nesta condição.

Dos genes anotados em *H. salinarum* cujos termos GO são relacionados à atividade antiporter de cátions, prótons e íons metálicos, 100% se encontravam diferencialmente expressos. As enzimas argininosuccinato sintase e liase estavam menos expressas, o que contribuiu para o enriquecimento funcional desse processo biológico na Figura 10. Cerca de 25% dos genes de proteínas de classe *TATA binding protein* (TBP) e fatores de início de transcrição estavam enriquecidos no tratamento de hipersalinidade e em menor porcentagem é possível observar genes relacionados ao transporte transmembrana de substâncias.

Figura 9: Representação do enriquecimento funcional das funções moleculares diferencialmente expressas em meio hipersalino. Classificação pela função molecular mostrando os 20 menores p-valores. O p-valor é o resultado de um teste de Fisher entre o número de genes contendo o termo GO no genoma de referência e o número de genes da lista de genes diferencialmente expressos contendo o mesmo GO. O ratio se refere à divisão entre número de genes diferencialmente expressos contendo um termo GO e o número de genes do genoma de referência com o mesmo termo GO.

Nos resultados de enriquecimento de processos biológicos (Figura 10) várias vias de biossíntese de biomoléculas estavam diferencialmente expressas. Cerca de 70% e 60% dos genes relacionados ao transporte transmembrana de fosfato e biossíntese de corismato estavam diferencialmente expressos, respectivamente. Ambos processos se encontravam menos expressos em hipersalinidade portanto, observou-se que houve diminuição de síntese de aminoácidos aromáticos e de metabólitos secundários como flavonoides e alcaloides (HERRMANN, 1995) relacionados à biossíntese do corismato e a via do xikimato (shikimato). Termos GO relacionados à transcrição e interação proteína-DNA também se encontravam diferencialmente expressos.

Figura 10: Representação do enriquecimento funcional de processos biológicos diferencialmente expressos em meio hipersalino. Classificação pelos processos biológicos mostrando os 20 menores p-valores. O p-valor é o resultado de um teste de Fisher entre o número de genes contendo o termo GO no genoma de referência e o número de genes da lista de genes diferencialmente expressos contendo o mesmo GO. O ratio se refere à divisão entre número de genes diferencialmente expressos contendo um termo GO e o número de genes do genoma de referência com o mesmo termo GO.

4.4. Expressão diferencial em meio hiposalino (CM 2,6M de NaCl)

Para verificar a concordância entre o número de reads de cada gene nas duas replicatas do meio hiposalino foi feito um *scatter plot* (Figura 11). Nele é possível observar uma tendência linear dos dados, indicando concordância entre as réplicas, mas ainda assim, há um acúmulo de pontos na parte central do gráfico, menos acentuado que o observado nos dados de hipersalinidade (Figura 8). Na análise de expressão diferencial em meio hiposalino obteve-se 268 genes com expressão significativamente diminuída e 209 genes com expressão aumentada.

Figura 11: *Scatter plot* comparando o valor de M das duas réplicas para o tratamento hiposalinidade, cuja concentração de NaCl no meio de cultura foi de 2,6M. Link para dados Tabela S1.

4.4.1. Genes menos expressos em hiposalinidade

Na Tabela 5 é possível observar uma lista contendo 35 dos 268 genes que estavam menos expressos na condição de hiposalinidade. Dentre eles estavam o gene *acs* com valor M de -5,5, que indica a expressão diminuída aproximadamente 10 vezes, quando comparado a expressão no meio controle.

Tabela 5: Tabela contendo os 35 genes menos expressos em meio hiposalino. Contém o locus, nome do gene, produto, valor de M que representa o $\log_2 (L (low\ salinity) / O (optimal\ salinity))$ e o desvio padrão dos valores de M das replicatas. Dados completos na Tabela S1.

<i>Locus</i>	Nome do gene	Produto	M	sdM
VNG_RS03880	<i>acs</i>	acetate--CoA ligase	-5.5	0.5
VNG_RS03255	VNG_RS03255	4Fe-4S dicluster domain-containing protein	-5.1	1.0

VNG_RS03260	<i>nrfD</i>	polysulfide reductase NrfD	-5.0	0.0
VNG_RS04635	VNG_RS04635	radical SAM protein	-5.0	1.0
VNG_RS03265	VNG_RS03265	molecular chaperone	-4.9	1.0
VNG_RS03870	VNG_RS03870	hypothetical protein	-4.7	0.1
VNG_RS05380	VNG_RS05380	hypothetical protein	-4.7	0.1
VNG_RS05715	VNG_RS05715	bacteriorhodopsin	-4.6	0.9
VNG_RS08545	VNG_RS08545	hypothetical protein	-4.1	0.0
VNG_RS08550	VNG_RS08550	b(o/a)3-type cytochrome c oxidase subunit 1	-4.1	0.1
VNG_RS10625	<i>gvpA</i>	gas vesicle structural protein GvpA	-4.1	0.4
VNG_RS12320	<i>gvpA</i>	gas vesicle structural protein GvpA	-4.1	0.4
VNG_RS03250	VNG_RS03250	molybdopterin-dependent oxidoreductase	-3.9	0.3
VNG_RS09205	VNG_RS09205	twin-arginine translocation signal domain -containing protein	-3.9	1.1
VNG_RS07650	VNG_RS07650	hypothetical protein	-3.8	1.6
VNG_RS03875	VNG_RS03875	helix-turn-helix domain -containing protein	-3.7	0.0
VNG_RS05675	VNG_RS05675	phytoene/squalene synthase family protein	-3.7	0.0
VNG_RS03245	VNG_RS03245	hypothetical protein	-3.7	2.0
VNG_RS02585	VNG_RS02585	GNAT family N-acetyl transferase	-3.7	0.2
VNG_RS14250	VNG_RS14250	hypothetical protein	-3.6	1.7
VNG_RS12740	VNG_RS12740	hypothetical protein	-3.6	0.1
VNG_RS00745	VNG_RS00745	halorhodopsin	-3.5	0.4
VNG_1459H	VNG_1459H	small CPxCG-related zinc finger protein	-3.4	0.3
VNG_RS06090	VNG_RS06090	magnesium chelatase subunit D family protein	-3.4	1.4
VNG_RS09245	VNG_RS09245	ABC transporter ATP-binding protein	-3.4	0.3
VNG_RS02580	VNG_RS02580	methyl malonyl -CoA mutase family protein	-3.4	2.0

VNG_RS08280	VNG_RS08280	beta-carotene 15,15'-dioxygenase,Brp/Blh family	-3.4	2.4
VNG_RS03240	VNG_RS03240	helix-turn-helix domain -containing protein	-3.4	0.4
VNG_RS14385	VNG_RS14385	halocyanin domain -containing protein	-3.4	0.5
VNG_RS11425	VNG_RS11425	SHOCT domain -containing protein	-3.4	0.1
VNG_RS08555	VNG_RS08555	cytochrome c oxidase subunit II	-3.3	0.1
VNG_RS09240	VNG_RS09240	ABC transporter permease	-3.2	0.4
VNG_RS04345	VNG_RS04345	SAM-dependent methyl transferase	-3.1	0.3
VNG_RS09235	VNG_RS09235	ABC transporter permease	-3.1	0.3
VNG_RS09435	VNG_RS09435	hypothetical protein	-3.1	0.1

4.4.1.1. Funções dos genes menos expressos em hiposalinidade

A tabela completa, contendo os 268 genes menos expressos em hiposalinidade, pode ser encontrada nos materiais suplementares (Tabela S1). Dentre os genes menos expressos em hiposalinidade, se encontram genes que codificam as subunidades de enzimas relacionadas ao estresse oxidativo, como a dimetil sulfoxido redutase (DMSO redutase). A DMSO redutase é formada através da transcrição do operon *dmsREABCD* que dá origem às suas três subunidades (*dmsABC* - VNG_RS03250, VNG_RS03255 e VNG_RS03260), o gene *dmsR* (VNG_RS03240), um possível fator de transcrição e *dmsD* (VNG_RS03265), uma chaperona molecular (MÜLLER; DASSARMA, 2005). A presença de DMSO redutase no genoma de organismos está relacionada à capacidade de crescer em anaerobiose, pois catalisa a reação em que o dimetil sulfoxido é oceptor final de elétrons na fosforilação oxidativa em anaerobiose (KAPPLER; SCHÄFER, 2014). Em bactérias, como a *Rhodobacter sphaeroides*, já foi demonstrada uma relação entre a transcrição da DMSO redutase com a presença de oxigênio no meio através de interações com fatores de transcrição (MOUNCEY; CHOUDHARY; KAPLAN, 1997; ZEILSTRA-RYALLS; KAPLAN, 1998).

Outros genes relacionados à respiração celular estavam menos expressos, entre eles várias cópias das subunidades de citocromo ubiquinol oxidase (VNG_RS10730, VNG_RS10735, VNG_RS12125, VNG_RS12425, VNG_RS12430 e VNG_RS13065), também chamada de complexo III da cadeia transportadora de elétrons e duas subunidades da

citocromo c oxidase (VNG_RS08550 e VNG_RS08555), que atua na última reação da cadeia de transportadora de elétrons (MATTAR; ENGELHARD, 1997; ABRAMSON *et al.*, 2000). Assim como observado em hipersalinidade, o operon *psf* estava menos expresso. Observa-se então que tanto os genes relacionados com a fosforilação oxidativa quanto os transportadores de fosfato inorgânico, que iriam atuar juntos na formação de ATP estão menos expressos. Isto indica uma possível regulação do metabolismo para obtenção de energia de outra maneira.

Uma variedade de genes menos expressos podem estar ligados à obtenção de energia, entre eles, os genes da família das rodopsinas. A bacteriorodopsina (VNG_RS05715) é uma bomba de prótons dependente de luz que atua bombeando prótons para fora da célula, gerando um potencial eletroquímico. Já as halorodopsinas (VNG_RS00745) agem de forma semelhante, bombeando ânions para dentro da célula, que também mantém o potencial eletroquímico necessário para produção de ATP (NELSON, 2013; SONG; GUNNER, 2014). As rodopsinas sensoriais (VNG_RS06435 e VNG_RS06440) atuam na fototaxia e transdução de sinais para na ativação do arqueio através da fosforilação e também estavam menos expressas (SASAKI; SPUDICH, 2008). Sabendo da comunicação entre as rodopsinas e o arqueio, observou-se que as arqueinas B1, B2 e B3 (VNG_RS03735, VNG_RS03740 e VNG_RS03745) e o gene *arlCE* (VNG_RS03715), que formam a estrutura do arqueio, estavam menos expressas. Outros genes do operon arqueio se encontravam menos expressos, mas somente os citados atingiram o valor de M estabelecido para genes diferencialmente expressos.

O que permite às rodopsinas responderem à luz é sua ligação covalente ao cromóforo retinal (HASSON; GAI; ANFINRUD, 1996; SONG; GUNNER, 2014). O retinal é produzido através de várias reações dentre elas a conversão de geranyl difosfato a fitoeno realizada pela enzima fitoeno sintase (VNG_RS05675), licopeno a beta-caroteno pela enzima licopeno beta ciclase (VNG_RS08285) e beta-caroteno a retinal pela enzima beta-caroteno dioxigenase (VNG_RS08280 - KEGG hal01240), as três menos expressas em hiposalinidade. Observa-se então que dois dos mecanismos de obtenção de energia das células de *H. salinarum* se encontram menos expressos, a respiração celular e a fotossíntese.

Grande parte dos genes das vesículas de gás foram encontrados menos expressos. O operon *gvp* é composto por 14 genes e está presente em duplicata no genoma de *H. salinarum* NRC-1 em cada um dos dois minicromossomos. Todos os genes das duas cópias do operon estavam menos expressos, mas somente sete genes em cada uma das cópias atingiram o

valor de M para serem considerados diferencialmente expressos. Os genes *gvpA* e *gvpC*, cujas proteínas formam a estrutura das vesículas de gás, estavam entre eles.

Assim como observado em hipersalinidade, o operon *pst* estava menos expresso. Além desses, estavam diferencialmente expressos transportadores do tipo ABC com possíveis substratos dipeptídeos e oligopeptídeos (*VNG_RS09235*, *VNG_RS09240*, *VNG_RS09245*, *VNG_RS09250*, *VNG_RS09185*, *VNG_RS09190*), cobalto (*VNG_RS06330* e *VNG_RS06335*) e um transportador de potássio (*VNG_RS11310*).

Genes relacionados a maquinaria de reparo e recombinação também foram menos expressos, entre eles uma fotoliase (*VNG_RS06795*), que atua no reparo de danos por radiação, e vários outros genes contendo o domínio hélice-volta-hélice normalmente identificado em proteínas que se ligam ao DNA e regulam a transcrição, como o gene *VNG_RS01035*, um fator de início de transcrição e o gene *VNG_RS01050*, que atua na regulação negativa do metabolismo de ácidos fenólicos. O gene *VNG_RS09565* é um regulador que atua na transcrição dos genes da família Ars de resistência à arsenato (WANG *et al.*, 2004), que também tinha um representante menos expresso em hipersalinidade.

Observou-se que os genes *VNG_RS04635*, *VNG_RS04345* e *VNG_RS00760* estavam menos expressos. Esses genes codificam uma proteína Radical S-adenosil-L-metionina (SAM) e duas enzimas metiltransferase dependente de SAM. Essas proteínas atuam na catálise de diversas reações, como metilações, isomerizações, desaminações e reações redox de radicais livres em todas as classes de biomoléculas (PUEHRINGER; METLITZKY; SCHWARZENBACHER, 2008; GANA *et al.*, 2013). A superóxido dismutase (*VNG_RS05200*), que também atua na proteção contra espécies reativas de oxigênio (YOUNUS, 2018) também estava menos expressa.

Em hipersalinidade, os genes relacionados à biossíntese de cobalamina estavam mais expressos. Quando submetida à hiposalinidade a população celular apresentou diminuição da expressão dos genes da síntese *de novo* desse cofator. Os genes *VNG_RS06120*, *VNG_RS06125*, *VNG_RS06130*, dentre outros; transportadores de cobalto (*VNG_RS06330* e *VNG_RS06335*); e enzimas que ligam a cobalamina como cofator (*VNG_RS01255*, *VNG_RS02580* e *VNG_RS08930*) estavam menos expressos.

Genes de proteínas que atuam no processo de divisão celular e o gene da proteína que forma a camada S, a S-layer, também se encontravam menos expressos. A divisão celular está representada aqui por genes que codificam proteínas semelhantes a tubulina *fstZ2*, *fstZ4* e

fstZ5 (VNG_RS00790, VNG_RS01070 e VNG_RS11450), o gene *minD1*, um inibidor de divisão celular (VNG_RS03685), dois genes de proteínas de segregação e condensação de cromossomos (VNG_RS1345 e VNG_RS1350) e o gene da proteína de replicação A e outra associada (*rpap3*), que mantém o DNA estável durante a replicação (VNG_RS08400 e VNG_RS08285). Já a glicoproteína da *S-layer* (VNG_RS10505) estava menos expressa assim como a glicosiltransferase que atua na N-glicosilação dela (VNG_RS04095) (JAEKEN; CARCHON, 2009).

4.4.2. Genes mais expressos em hiposalinidade

Na Tabela 6 é possível observar uma lista contendo 35 dos 209 genes que estavam mais expressos na condição de hiposalinidade. Dentre eles estavam o gene VNG_RS11470, que codifica uma proteína de transporte de metais do tipo ABC, com valor M de 4, que indica a expressão aumentada aproximadamente 8 vezes, quando comparado à expressão no meio controle.

Tabela 6: Tabela contendo os 35 genes mais expressos em hiposalinidade. Contém o locus, nome do gene, produto, valor de M que representa o $\log_2 (L \text{ (low salinity)} / O \text{ (optimal salinity)})$ e o desvio padrão dos valores de M das replicatas. Dados completos encontram-se na Tabela S2.

Locus	Nome do gene	Produto	M	sdM
VNG_RS11470	VNG_RS11470	metalABC transporter substrate-binding protein	4.4	0.6
VNG_1642H	VNG_1642H	NA	3.7	1.3
VNG_RS11585	queC	7-cyano-7-deazaguanine synthase QueC	3.7	1.1
VNG_RS09085	VNG_RS09085	GNAT family N-acetyl transferase	3.6	0.8
VNG_RS11010	VNG_RS11010	hypothetical protein	3.6	2.1
VNG_RS03620	VNG_RS03620	LLM class flavin-dependent oxidoreductase	3.5	1.2
VNG_RS03640	VNG_RS03640	nitroreductase family protein	3.4	0.5
VNG_RS04655	VNG_RS04655	superoxide dismutase	3.4	0.2
VNG_RS05350	VNG_RS05350	hypothetical protein	3.4	2.2
VNG_RS12050	VNG_RS12050	site-specific integrase	3.2	0.5
VNG_RS02190	VNG_RS02190	DJ-1/Pfpl family protein	3	0.9

VNG_RS11590	VNG_RS11590	7-carboxy-7-deazaguaninesynthaseQueE	3	0.5
VNG_RS02760	VNG_RS02760	MarR family transcriptional regulator	3	1.1
VNG_RS05265	VNG_RS05265	EamA family transporter	3	1.3
VNG_1238C	VNG_1238C	NA	2.9	1.8
VNG_RS07635	glpA	anaerobic glycerol-3-phosphate dehydrogenase subunit A	2.9	0.1
VNG_RS06370	VNG_RS06370	adenosyl cobalamin-dependent ribonucleoside-diphosphate reductase	2.9	0.2
VNG_RS05750	VNG_RS05750	hypothetical protein	2.9	1.3
VNG_RS05635	VNG_RS05635	5'-nucleotidaseC-terminal domain -containing protein	2.8	0.4
VNG_RS08190	VNG_RS08190	glutathioneS- transferaseN-terminal domain -containing protein	2.8	0.9
VNG_RS08445	VNG_RS08445	iron-containing alcohol dehydrogenase	2.8	1.3
VNG_RS04300	VNG_RS04300	tripartite tricarboxylate transporter permease	2.7	1.1
VNG_RS09480	VNG_RS09480	O-acetyl homoserine amino carboxypropyl transferase/cysteine synthase	2.7	0.2
VNG_RS04230	VNG_RS04230	UPF0058 family protein	2.7	0.6
VNG_RS05755	VNG_RS05755	RND family transporter	2.7	1.4
VNG_RS07670	trmY	tRNA(pseudouridine(54)-N(1))-methyl transferase TrmY	2.7	1.1
VNG_RS07645	VNG_RS07645	anaerobic glycerol-3-phosphate dehydrogenase subunit C	2.6	1.1
VNG_RS11220	VNG_RS11220	universal stress protein	2.6	0.2
VNG_RS11645	VNG_RS11645	hypothetical protein	2.6	2.2
VNG_0057H	VNG_as00215_75	NA	2.6	0.5
VNG_RS14195	VNG_RS14195	hypothetical protein	2.6	0.7

VNG_RS05490	VNG_RS05490	acyl-CoA thioesterase	2.6	0.6
VNG_RS01225	VNG_RS01225	fructose-bisphosphate aldolase	2.5	1.8
VNG_RS07640	glpB	glycerol-3-phosphate dehydrogenase subunit GlpB	2.5	0.8

4.4.2.1. Funções dos genes mais expressos em hiposalinidade

A tabela completa, contendo os 209 genes mais expressos em hiposalinidade, pode ser encontrada nos materiais suplementares (Tabela S2). Dentre estes se encontram genes que codificam enzimas topoisomerase, endonucleases e excinuclease. A enzima homóloga a topoisomerase tipo I, *VNG_RS08620*, é uma enzima que atua na manutenção dos superenovelamento durante a replicação e transcrição do DNA (FORTERRE; GADELLE, 2009). As endonucleases *VNG_RS02330* e *VNG_RS08780*, codificam respectivamente a endonuclease III, a junção de Holliday resolvase HJC, e têm função reparo e excisão de bases nitrogenadas e resolve a estrutura heteroduplex das junções de Holliday para a continuação da replicação (MICHEL; SINHA; LEACH, 2018; HUANG *et al.*, 2019). Já a ribonuclease Z (*VNG_RS08725*) participa do processamento da extremidade 3' e maturação dos tRNAs (Uniprot:A0A510NA20). Uma proteína homóloga a PilT ATPase tipo I (*VNG_RS10235*) atua como componente tóxico no sistema toxina-antitoxina (Uniprot:Q9HMB8) e está mais expressa em hiposalinidade. Isso possivelmente indica que a salinidade poderia acarretar no mau funcionamento das proteínas da maquinaria de transcrição e replicação e seria necessário um aumento na expressão dos genes que atuam nas excisões e reparo de nucleotídeos.

Uma variedade de genes que atuam no reparo de dano por radiação no DNA estavam mais expressos. Entre elas, a *uvrA* (*VNG_RS10310*), envolvida no reconhecimento de danos, e *radA* (*VNG_RS09665*) envolvida na recombinação homóloga com finalidade de reparo de quebras da fita dupla (*double-strand break* - DSB) (WHITE; ALLERS, 2018). Já os genes *VNG_RS10875* e *VNG_RS05070*, codificam uma proteína com atividade de peptidase capaz de se ligar ao DNA fita simples e atua na resposta ao dano no DNA (Uniprot:A0A510NA04) e uma tRNA sulfotransferase, que adiciona enxofre ao tRNA e funciona como sensor UV, respectivamente. Outro gene relacionado ao reparo do material genético é o *VNG_RS04905*, uma helicase da família DEAD/DEAH-box, que atua na separação da fita dupla de DNA, na estrutura secundária do RNA e tem função complementar à topoisomerase encontrada mais expressa.

Alguns dos transportadores do tipo ABC também foram encontrados mais expressos. A partir dessa informação concluímos que biotina, zinco, ferro e potássio estavam sendo transportados para dentro das células. Outras famílias de transportadores tiveram genes mais expressos, os transportadores periplasmático tripartido independente de ATP (TRAP, *VNG_RS04300*) e a superfamília de facilitadores principais (MFS, *VNG_RS06540*), que atuam no transporte de uma variedade de substratos em diferentes contextos, desde ácido siálico para formar a estrutura de glicoproteínas, até o transporte de solutos compatíveis em resposta ao estresse osmótico em *Halomonas elongata* (ROSA *et al.*, 2018). Além disso, um canal de íons cloreto (*VNG_RS06005*) e um transportador de arsenato (*VNG_RS12855*) para fora da célula estavam mais expressos. Em hipersalinidade vimos a diminuição da expressão de uma proteína de estresse universal e em hiposalinidade outra proteína com a mesma função (*VNG_RS11220*) se encontrava mais expressa (UniProtKB - Q9HHX5).

A transcrição e tradução também foram afetadas pela diminuição da salinidade e proteínas que se ligam ao TATA box (*VNG_RS11040* e *VNG_RS12735*) estavam mais expressas. Fatores de início de transcrição (*VNG_RS04180*), regulador de transcrição dos transportadores de trealose e maltose (*VNG_RS07310* e *VNG_RS12070*) e proteínas contendo domínio HTH, que atuam como reguladores da transcrição (*VNG_RS09445*) estavam mais expressos. Além disso, ativadores da transcrição de bacterio-opsina (*VNG_RS13030* e *VNG_RS12090*), proteína associada ao retinol da família das rodopsinas (BETLACH; SHAND; LEONG, 1989) estavam mais expressos. Na tradução, dois genes que atuam na modificação de tRNA estavam mais expressos: o *VNG_RS07670* que atua na pseudouridinação e consequente estabilização do pareamento códon-anticódon, e o gene *VNG_RS09675*, que atua na metilação dependente de S-adenosil metionina (Uniprot:Q9HMM2 e A0A510N833). Um gene que atua na ligação entre tRNA e fenilalanina (*VNG_RS09775*) também estava mais expresso. Proteínas ribossomais que ajudam na estabilização das subunidades maior e menor do ribossomo (*VNG_RS04320* e *VNG_RS09840*), um fator de elongação da tradução (*VNG_RS04575*) e uma DNA polimerase (*VNG_RS11805*) também estavam mais expressos. Isto pode indicar que o estresse salino afeta a regulação diferencial de alguns dos genes que atuam nas maquinarias de transcrição e tradução.

Outros genes encontrados mais expressos são os presentes no metabolismo de folatos (*VNG_RS11585*, *VNG_RS11590* e *VNG_RS11595*) de aminoácidos (*VNG_RS07030* e *VNG_RS07035*) e glicerol-3 fosfato (*VNG_RS07635*, *VNG_RS07640* e *VNG_RS07645*). No metabolismo de folatos, três enzimas que atuam consecutivamente para a formação de

7-aminometil-7-carbaguanina. Essa substância é uma das precursoras do nucleosídeo Q, a queuosina, que é incorporada ao tRNA e afeta a conversão de fenilalanina a tirosina e pode atuar também na regulação de transcrição por codon-bias, a preferência por alguns códons do organismo (KOHLMEIER, 2015). As duas enzimas mais expressas do metabolismo de aminoácidos são as duas subunidades da carbamoil-fosfato sintase, que atuam na síntese de carbamoil-fosfato. O carbamoil fosfato, por sua vez, pode ser direcionado a duas vias metabólicas: o metabolismo de pirimidinas e a biossíntese de arginina (KEGG:hal00250). Já os genes relacionados ao glicerol-3-fosfato são as três subunidades que formam a enzima glicerol-3-fosfato desidrogenase, que atua na reação entre uma quinona e o glicerol-3-fosfato resultando em um quinol e dihidroxiacetona fosfato. A dihidroxiacetona fosfato pode ser convertida a gliceraldeído-3-fosfato pela triose-fosfato isomerase e participar da via glicolítica, no metabolismo de aminoácidos e no metabolismo de glicolipídeos que vão compor a membrana celular (KEGG:hal00564,hal01230).

4.4.3. Enriquecimento funcional dos genes diferencialmente expressos em hiposalinidade

Do total de 477 genes diferencialmente expressos, 402 estavam anotados no banco de dados utilizado pelo software. No enriquecimento das funções moleculares (Figura 12) poucas funções tiveram significância menor que 5% ($p\text{-valor} < 0,05$). Dentre elas, 100% dos genes anotados com funções de catabolismo do glicerol-3-fosfato estavam diferencialmente expressos, e também 100% dos genes anotados como opsinas que absorvem e respondem à radiação. Em porcentagem menor, estavam moléculas cujas funções são se ligar ao RNA ou exercer sobre ele atividade catalítica.

Figura 12: Representação do enriquecimento funcional das funções moleculares diferencialmente expressas em hiposalinidade. Classificação pela função molecular. O p-valor é o resultado de um teste de Fisher entre o número de genes contendo o termo GO no genoma de referência e o número de genes da lista de genes diferencialmente expressos contendo o mesmo GO. O ratio se refere à divisão entre número de genes diferencialmente expressos contendo um termo GO e o número de genes do genoma de referência com o mesmo termo GO.

Já a representação de processos biológicos enriquecidos em hiposalinidade (Figura 13) mostrou que 100% dos genes anotados estão relacionados a mecanismos de quimiostase, como resposta ao estímulo de luz e detecção de estímulos externos. Em menor porcentagem, 60% dos genes anotados relacionados à motilidade celular, formação de arqueolo e reações químicas e vias envolvendo terpenóides estavam diferencialmente expressos em hiposalinidade. Aproximadamente 55% dos genes relacionados à resposta à radiação estavam diferencialmente expressos nas mesmas condições.

Figura 13: Representação do enriquecimento funcional dos processos biológicos diferencialmente expressos em hiposalinidade. Classificação pelos processos biológicos. O p-valor é o resultado de um teste de Fisher entre o número de genes contendo o termo GO no genoma de referência e o número de genes da lista de genes diferencialmente expressos contendo o mesmo GO. O ratio se refere à divisão entre número de genes diferencialmente expressos contendo um termo GO e o número de genes do genoma de referência com o mesmo termo GO.

4.5. Suplementação com glicerol interferiu no fenótipo da cultura em hiposalinidade

A análise de expressão diferencial (Tabela 3) revelou que os genes *glpA*, *glpB* e *glpC* (*VNG_RS07635*, *VNG_RS07640* e *VNG_RS07645*, respectivamente), que codificam para as três subunidades da enzima glicerol-3-fosfato desidrogenase, estavam mais expressos em hiposalinidade com média de $M = 2,66$. Sabe-se que o habitat da *H. salinarum* também é lar para a alga *Dunaliella salina*, que produz glicerol, um soluto compatível para lidar com o desbalanço osmótico que ocorre por variações das condições do habitat (ORELLANA *et al.*, 2013; WILLIAMS *et al.*, 2017a).

A co-cultura dos dois microrganismos demonstrou que *D. salina* secreta substâncias e nutrientes importantes para a sobrevivência da *H. salinarum* (ORELLANA *et al.*, 2013). Levantamos a hipótese de que a maior expressão dos genes que codificam a glicerol-3 fosfato desidrogenase poderia estar relacionado com a disponibilidade de glicerol no habitat hiposalino como o secretado por *D. salina*, permitindo o aumento da D.O. das culturas nessas condições (comparados à Figura 4) quando suplementadas com glicerol. Para testar essa possibilidade, foi feita uma curva de crescimento adicionando aos meios 30 e 60 mg/ml de glicerol. As concentrações foram inferidas do trabalho de Chen *et al.* (2009) onde demonstrou-se que a quantidade média de glicerol produzido pela *D. salina* era de 30 mg/ml em diversas concentrações de sal do meio e a quantidade máxima produzida era de 60 µg/mL em meio com 2 M de NaCl (CHEN; JIANG; WU, 2009).

Pode-se observar que ao suplementar com o glicerol, seja 30 mg/ml (Figura 14A) ou 60 mg/ml (Figura 14B), as amostras no CM 2,6M apresentaram um fenótipo de crescimento melhor quando comparados ao CM 2,6M sem suplementação. Uma possível explicação para o acontecimento é a maior oferta de carbono pela presença de glicerol e a maior expressão de genes que atuam na conversão de glicerol em energia. Quando comparados ao controle (CM 4,3M), no entanto, ainda têm a densidade óptica menor.

Na figura 14A e C, as amostras em meio CM 4,3M suplementadas com 30 mg/ml começam a ter fenótipo de crescimento maior que o controle sem suplemento, mas após 62 horas o crescimento cai. O mesmo acontece com o meio CM 4,3M + 60 mg/ml de glicerol que tem seu crescimento médio um pouco menor que o controle sem suplementação durante toda a curva, indicando que a suplementação influenciou negativamente em seu fenótipo. Assim como visto na figura 14A, a suplementação em condições padrão (CM 4,3M) parece interferir negativamente no crescimento das células. Neste caso, se as células consomem o glicerol e

passam a não tê-lo como fonte de carbono, era esperado observarmos uma curva bem semelhante à do meio CM 4,3M. Poderia se pensar então que certas concentrações de glicerol combinadas com o tempo de incubação não são interessantes para o fenótipo de crescimento das células em condições controle e que outros mecanismos podem estar regulando o crescimento das células em meio hiposalino.

Figura 14: Curva de crescimento de *H. salinarum* NRC-1. A) em meio de cultura com 2,6M de NaCl e 4,3 M de NaCl suplementado com 30 mg/ml de glicerol. B) Curva de crescimento de *H. salinarum* NRC-1 em meio de cultura com 2,6 mol/l de NaCl (CM 2,6), 4,3 mol/l de NaCl (CM 4,3) e 5,1 mol/l de

NaCl (CM 5,1) suplementado com 60 mg/ml de glicerol. O experimento foi feito em triplicata. C) Representação da diferença entre a densidade óptica dos tratamentos: CM 2,6M (ponto azul), CM 2,6M + 30mg/ml de glicerol (quadrado roxo), CM 2,6 + 60 mg/ml de glicerol (símbolo +), CM 4.3M + 30 mg/ml (símbolo x) e CM 4,3M +60 mg/ml de glicerol (triângulo invertido) - e a densidade óptica do controle de acordo com o tempo (horas).

4.6. Comparação com dados de proteômica e transcritômica

Nos trabalhos de Coker *et al.* (2007) e Leuko *et al.* (2009) foram apresentados resultados de análises sob condições de estresse osmótico usando as metodologias de proteômica e transcritômica através de microarranjos de DNA, respectivamente. Quando analisamos os resultados dos dois trabalhos, não observamos correlação entre a expressão dos genes e produção de proteínas, o que nos levou a verificar a correlação entre os dois experimentos publicados e o RNA-seq realizado neste trabalho. Em posse dos dados parciais de proteômica e transcritômica obtidos diretamente das respectivas publicações, pode ser observado que os meios utilizados para cada um dos experimentos apresentam diferenças (Tabela 7). Além disso, há a diferença entre a rotação e a temperatura de incubação, o que pode ser um fator que causou os resultados distintos quando os três métodos foram confrontados.

Tabela 7: Condições de cada experimento para o meio 4,3M de NaCl.

Variáveis	Leuko et.al. (2009)	RNA-seq (este trabalho)	Coker et. al. (2007)
Meios de Cultura	NaCl - 250g/l	NaCl - 250g/l	NaCl - 250g/l
	MgSO ₄ .7H ₂ O - 20 g/l	MgSO ₄ .7H ₂ O - 20 g/l	MgSO ₄ .7H ₂ O - 20 g/l
	Citrato de sódio - 3 g/l	Citrato de sódio - 3 g/l	Citrato de sódio - 3 g/l
	KCl - 2 g/l	KCl - 2 g/l	KCl - 2 g/l
	Tryptona - 5 g/l	Peptona - 10 g/l	Peptona 10 g/l
	Extrato de levedura - 3 g/l		Traços de metais - 100 µl/l *
Rotação	160 RPM	160 RPM	220 RPM
Temperatura	37°C	37°C	42°C

*(ROBB; DASSARMA; FLEISCHMANN, 1995)

Quando comparamos os resultados dos métodos de microarranjos de DNA e proteômica, foram encontrados apenas 4 genes em comum em hipersalinidade (Tabela 8) e 6 genes em comum em hiposalinidade entre as análises (Tabela 9). Foi possível recuperar apenas os dados com expressão diferencial significativa das publicações de proteômica e microarranjos, que foram comparados com os dados de expressão diferencial dos nossos dados completos de RNA-seq.

Tabela 8: Genes comuns aos experimentos de proteômica e microarranjo de DNA em hipersalinidade.

Proteômica		Microarranjos de DNA		RNA-seq		Anotação dos genes		
M	sdM	M	sdM	M	sdM	Locus	Nome do gene	Produto
0.2	0.2	1.1	0.1	-0.1	0.3	VNG_RS10240	VNG_RS10240	Carboxy peptidase M32
-1.1	0.4	1	0.1	-0.2	1.2	VNG_RS02095	VNG_RS02095	metal-dependent transcriptional regulator
-0.7	0.2	1.5	0.2	0.4	1.4	VNG_RS07030	<i>carB</i>	carbamoyl -phosphate synthase large subunit
-0.7	0.3	1.2	0.5	-0.6	0.2	VNG_RS07035	<i>carA</i>	Glutamine-hydrolyzing carbamoyl -phosphate synthase small subunit

No trabalho de Leuko *et al.* (2009) foi utilizado o $\log_{10}$ da razão entre quantidade de proteínas no tratamento e controle, onde valores maiores de zero significam que os genes estão mais expressos e valores menores que zero indicam genes menos expressos. Já no trabalho de Coker *et al.* (2007) foram encontrados valores de fold change e o resultado é dado pelo $\log_2$ (fold change), significando que valores maiores que 1,3 representam genes mais expressos e valores menores que -1,3 representam genes com expressão diminuída. Tanto os resultados de de Leuko *et al.* (2009) quanto os de Coker *et al.* (2007) foram convertidos para o valor M, para ser possível a comparação dos dados. Para os resultados do RNA-seq, valores acima de 1 e abaixo de -1 são considerados significativamente mais expressos e menos expressos que a condição controle, respectivamente.

Tabela 9: Genes comuns aos experimentos de proteômica e microarranjo de DNA em hiposalinidade.

Proteômica		Microarranjos de DNA		RNA-seq		Anotação dos genes		
------------	--	----------------------	--	---------	--	--------------------	--	--

M	sdM	M	sdM	M	sdM	Locus	Nome do gene	Produto
0.7	0.3	-1.1	0.2	0.0	0.1	VNG_RS10240	VNG_RS10240	Carboxy peptidase M32
-0.6	-0.2	-0.4	0.2	0.1	0.2	VNG_RS01920	<i>dnaK</i>	Molecular chaperone DnaK
-0.7	0.5	-0.6	0.1	0.3	1.2	VNG_RS06570	VNG_RS06570	50S ribosomal protein L2
-0.9	0.5	-0.5	0.2	0.3	1.4	VNG_RS06580	VNG_RS06580	50S ribosomal protein L22
-1.2	0.4	-0.6	0.2	-1.1	0.7	VNG_RS08295	VNG_RS08295	ATP synthase subunit A
-0.4	0.1	-0.8	0.3	-0.7	0.9	VNG_RS08320	VNG_RS08320	V-type ATP synthase subunit I

O número de genes representados ao mesmo tempo nos três experimentos foi muito baixo, devido principalmente à indisponibilidade dos dados completos das publicações e portanto, impossibilidade de re-análise dos dados brutos. Ainda assim, esperava-se encontrar concordância entre a maioria dos dados de microarranjos de DNA e RNA-seq por se tratarem do mesmo experimento de quantificação de transcritos e por avaliarem o estresse osmótico, mesmo que em condições diferentes. Sabemos que essa mesma afirmação pode não ser sempre verdadeira no caso da proteômica versus RNA-seq, já que existe a possibilidade de regulação pós-transcricional dos transcritos.

4.6.1. Comparação do transcrito mapeado por microarranjos e RNAseq

Os dados de RNA-seq obtidos neste trabalho foram comparados aos dados de microarranjo de DNA e foram encontrados 17 genes em comum nos experimentos em hipersalinidade e 37 genes em hiposalinidade. A seleção de genes em comum aos experimentos de microarranjos de DNA e RNA-seq evidenciou as diferenças entre as réplicas 1 e 2. Nas figuras seguintes é possível observar que os gráficos não representam uma média das réplicas e sim cada uma delas e assim pudemos observar a variabilidade do transcrito da *H. salinarum*. As réplicas são biológicas, cada uma de colônias diferentes e apesar do comportamento não ser o esperado em relação ao experimento de Coker *et al.* (2007), a diferença entre as réplicas é esperada por se tratarem de populações diferentes.

Figura 15: Gráfico de dispersão das amostras em hipersalinidade de RNA-seq versus Microarranjo de DNA. São representados os valores de M ($\log_2(\text{high salinity/optimal salinity})$) em hipersalinidade de duas réplicas do experimento e o correspondente no experimento de microarranjos de DNA. Link para dados TabelaS7G.

O critério de seleção dos genes diferencialmente expressos foi selecionar valores $|M| \geq 1$. Para comparação entre os experimentos, o critério usado foi $|M| \geq 0$, já que após a conversão dos critérios usados no experimento de microarranjos de DNA para o valor de M ($\log_2(\text{low ou high salinity/optimal salinity})$) muitos não atingiram o nosso critério para diferencialmente expressos ($|M| \geq 1$), mas de acordo com Coker et. al. (2007) todos genes presentes na publicação foram diferencialmente expressos.

Figura 16: Gráfico de dispersão das amostras em hiposalinidade de RNA-seq versus Microarranjo de DNA. São representados os valores de M ($\log_2(\text{low salinity/optimal salinity})$) em hiposalinidade de duas réplicas do experimento e o correspondente no experimento de microarranjos de DNA. Link para dados Tabela S7G.

Em hipersalinidade, dos 17 genes genes em comum, apenas os genes *VNG_RS09260* (malato desidrogenase) e *VNG_RS09310* (ABC transporter permease) estavam mais expressos nos dois experimentos (RNA-seq e microarranjos de DNA) e nenhum gene foi encontrado menos expresso (Figura 15 e Tabelas S3). Em hiposalinidade, foram encontrados 37 genes em comum (Figura 16 e Tabela S5), entre eles uma cold-shock protein (*VNG_RS00395*), uma proteína de quimiotaxia (*VNG_RS03955*), a superóxido dismutase (*sod1*, *VNG_RS04655*) e uma proteína que se liga ao TATA-box (*VNG_RS12735*) estavam mais expressas em ambos experimentos. Estavam menos expressos 14 genes nos dois experimentos (Figura 16), entre eles várias subunidades de ATP sintase (*VNG_RS08295*, *VNG_RS08300*, *VNG_RS08305*, *VNG_RS08310*, *VNG_RS08315* e *VNG_RS08320*), transportadores do tipo ABC (*VNG_RS09190*, *VNG_RS09195* e *VNG_RS09200*) e um antiporter arginina/ornitina (*VNG_RS11620*).

Por se tratar de dois experimentos para quantificação relativa de transcritos, esperava-se que os genes se agrupassem na diagonal correspondente aos quadrantes inferior esquerdo e superior direito. No entanto isso não foi observado, provavelmente devido às diferentes condições utilizadas nos experimentos (Tabela 7) e a quantidade de dados

disponíveis para o experimento de microarranjo de DNA. É possível observar nas figuras 16 e 17 que enquanto alguns representantes da réplica 1 tinham valores positivos, os mesmos na réplica 2 se encontravam negativos e vice e versa. Esse comportamento também não era esperado na análise dos dados, mas sabe-se que o experimento em si já tem suas desvantagens como a falta de especificidade por usar sondas curtas, o que faz com que seja necessária a validação dos experimentos usando outros métodos, o que não foi feito no experimento de Coker et al. (2007), ou o uso de muitas réplicas técnicas e biológicas (SILVA-JUNIOR; BELARMINO; BENKO-ISEPPON, 2018).

4.6.2 Comparação de dados de RNAseq com proteoma

Baseado no dogma central, sabe-se que existe certa correspondência quando comparamos os transcritos e proteínas. No entanto, vários experimentos visando este fato com mesmas células e condições semelhantes falharam em demonstrar essa alta correlação (GHAZALPOUR *et al.*, 2011; DU *et al.*, 2019). Isso pode acontecer devido a vários motivos, como a eficiência da tradução, mutações na sequência de Shine-Dalgarno, a estrutura tridimensional do RNA, a preferência por códons específicos na tradução dos aminoácidos, diferença de expressão em diferentes fases do ciclo de vida celular, tempo de meia vida do RNA, regulação pós-transcricional e erros experimentais (HAIDER; PAL, 2013).

Foram encontrados 178 genes em comum em hipersalinidade (Tabela S4) e muitos deles não eram diferencialmente expressos, na região central do gráfico (Figura 17). Estavam mais expressos em ambas os experimentos 23 genes, entre eles proteínas ribossomais (*VNG_RS04420* e *VNG_RS10425*), a *radA* (*VNG_RS09665*), uma subunidade da ATP sintase (*VNG_RS08295*), um regulador de transcrição (*VNG_RS03480*) e um fator de início de tradução (*VNG_RS06850*). Estavam menos expressos 8 genes, entre eles a arginino succinato liase (*VNG_RS09530*), a arginino succinato sintase (*VNG_RS09535*) e a ornitina carbamoyl transferase (*VNG_RS11625*).

Figura 17: Gráfico de dispersão das amostras em hipersalinidade de RNA-seq versus proteômica. São representados os valores de M ($\log_2(\text{high salinity/optimal salinity})$) em hipersalinidade de duas réplicas do experimento e o correspondente no experimento de proteômica. Link para dados TabelaS7G.

Em hiposalinidade, foram encontrados 142 genes em comum entre os experimentos de RNA-seq e proteômica (Figura 17 e Tabela S6). Entre eles, 29 genes estavam mais expressos em ambos experimentos, incluindo uma subunidade da gliceraldeído-3-fosfato desidrogenase (*VNG_RS00375*), uma subunidade da glicerol-3-fosfato desidrogenase (*VNG_RS07635*), a triptofano-tRNA ligase (*VNG_RS08610*), carbamato quinase e arginina deiminase (*VNG_RS11630* e *VNG_RS11635*). Outros 29 genes estavam menos expressos em ambos experimentos, entre eles algumas arqueínas (*VNG_RS03735*, *VNG_RS03740*, *VNG_RS03745* e *VNG_RS03940*), proteínas hipotéticas (*VNG_RS00860* e *VNG_RS12750*), gvpH e gvpO (*VNG_RS10600* e *VNG_RS12335*).

O experimento de proteômica abrangeu mais genes e o interessante de comparar este com o RNA-seq é que podem ser formuladas hipóteses em relação a regulação gênica. Por exemplo, 27 genes em condição de hipersalinidade (Tabela S4) e 21 genes em hiposalinidade (Tabela S6) foram encontrados mais expressos enquanto suas respectivas proteínas estavam menos expressas, o que poderia indicar possível regulação pós transcricional desses genes. Dos 27 genes em hipersalinidade, 16 eram proteínas ribossomais, indicando que pode haver a regulação da transcrição de genes de proteínas ribossomais para lidar com o estresse

osmótico, assim como ocorre durante as fases do ciclo celular de *H. salinarum* NRC-1 (LÓPEZ GARCÍA DE LOMANA et al., 2020). Foram encontrados também, 8 genes menos expressos em hipersalinidade (Tabelas S4) e 10 genes em hiposalinidade (Tabela S6) cujas proteínas apresentavam-se mais expressas. Estes representam o tipo de regulação que poderia ocorrer pela ação de RNAs antisense, que se pareiam ao mRNA podendo desestabilizar sua estrutura e impedir a tradução desse (XU; ZHANG; ZHANG, 2018).

Figura 18: Gráfico de dispersão das amostras em hiposalinidade de RNA-seq versus proteômica. São representados os valores de M ($\log_2(\text{low salinity/optimal salinity})$) em hiposalinidade de duas réplicas do experimento e o correspondente no experimento de proteômica. Link para dados TabelaS7G.

4.7. Motilidade das células é menor em hiposalinidade

Foi observado anteriormente que em hiposalinidade houve diminuição da expressão de genes que codificam as proteínas para a formação do arqueolo e também genes que transdutores de sinais relacionados ao arqueolo (Tabela 5). Para confirmar a menor expressão das proteínas relacionadas aos arqueolos foi feito o ensaio de motilidade no meio hiposalino e meio controle em meio semi-sólido, onde foi possível inferir visualmente a motilidade celular. O ensaio mostrou que em meio semi-sólido hiposalino a colônia de células cresce apenas ao

redor do local onde foi inoculada sem apresentar motilidade, ao contrário do que ocorre no meio semi-sólido com salinidade ótima, em que se espalha pela superfície da placa de petri (Figura 19).

Figura 19: Ensaio de motilidade. A) Placas de petri do crescimento em meio hiposalino e meio controle. Linha preta representa o diâmetro do halo de crescimento B) Gráfico representando a média de diâmetros e replicatas para cada amostra.

4.8. Novos transcritos identificados

Analisando manualmente os arquivos de cobertura, foram encontradas 53 regiões contendo transcritos não mapeados anteriormente pelo grupo. Destas, 24 regiões foram identificadas antisense a genes anotados e 29 foram identificadas em regiões intergênicas. A análise utilizando a matriz de contagem de *reads* resultou em 19 regiões cujas contagens antisense ao gene encontraram altas, ou seja, o módulo de seu valor $M (\log_2 (reads \text{ antisense}/reads \text{ sense}))$ maior ou igual a 1. Para verificar se realmente tratam-se de novos transcritos, as regiões foram submetidas ao alinhamento pelo programa Blast, nas modalidades Blastn e Blastx. Blastn para verificar se a sequência de nucleotídeos já foi encontrada em outros organismos e Blastx, que traduz nas três janelas de leitura e compara com proteínas, para ter certeza de que não codificam proteínas já caracterizadas. Na confirmação da existência das sequências em outros genomas utilizando o Blastn, observou-se que os alinhamentos aconteciam para muitas espécies do gênero *Halobacterium*, o que confirma a presença das regiões em outros genomas, mas não sua transcrição ou tradução.

Os nucleotídeos foram alinhados contra o banco de dados contendo todas regiões codificantes do GenBank traduzidas, dados do *Protein Data Bank* (PDB), dados do SwissProt,

dados do *Protein Information Resource* (PIR) e *Protein Research Foundation* (PRF). Foram selecionadas ao final do alinhamento as regiões que não alinharam com nenhuma proteína do banco no sentido correto das sequências, que totalizou 41 das 53 regiões encontradas manualmente - 21 dessas referentes a transcritos antisense. Adicionados às 19 regiões encontradas por meio de matriz de contagem de *reads*, têm-se no total 60 regiões de interesse que podem ser classificadas como: 1) regiões antisense a genes anotados e potencialmente não codificadoras (Tabela 10), e 2) regiões intergênicas que possivelmente são regiões 3' UTR dos genes adjacentes, como descrito na literatura (Tabela 11) (DE ALMEIDA et al., 2019).

Tabela 10: Tabela das regiões antisense encontradas contendo o nome dado a elas, *locus* de referência para as que são antisense, produto do gene antisense e o método da curadoria: manual ou por matriz de contagem de *reads*.

Nome	Locus de referência	Produto	Curadoria
VNG_n002_1	VNG_RS02915	hypothetical protein	manual
VNG_n003_1	VNG_RS08415	hypothetical protein	manual
VNG_n004_1	VNG_RS04880	hypothetical protein	manual
VNG_n005_1	VNG_RS04925	2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone 5'-phosphate reductase	manual
VNG_n006_1	VNG_RS04985	cyclic pyranopterin monophosphate synthase MoaC	manual
VNG_n008_1	VNG_RS05075	hypothetical protein	manual
VNG_n010_1	VNG_RS05170	tRNA-Glu	manual
VNG_n012_1	VNG_RS05360	sensor histidine kinase	manual
VNG_n015_1	VNG_RS05435	phosphopantetheine adenylyltransferase	manual
VNG_n016_1	VNG_RS05480	peptide-methionine (R)-S-oxide reductase MsrB	manual
VNG_n017_1	VNG_RS05550	amphi-Trp domain-containing protein	manual
VNG_n023_1	VNG_RS05500	DUF63 family protein	manual
VNG_n025_1	VNG_RS04950	translation initiation factor IF-2 subunit beta	manual
VNG_n044_1	VNG_RS13540	hypothetical protein	manual
VNG_n047_1	VNG_RS11430	antibiotic biosynthesis monooxygenase	manual
VNG_n050_1	VNG_RS11445	amidohydrolase family protein	manual

VNG_n052_1	VNG_RS11455	sodium-dependent transporter	manual
VNG_n053_1	VNG_RS11490	AAA family ATPase	manual
VNG_n056_1	VNG_RS11515	ABC transporter ATP-binding protein	manual
VNG_n058_1	VNG_RS11615	arginine--tRNA ligase	manual
VNG_n069_1	VNG_RS00335	molybdopterin-synthase adenylyltransferase MoeB	contagem
VNG_n070_1	VNG_RS00595	DUF302 domain-containing protein	contagem
VNG_n071_1	VNG_RS01390	CPBP family intramembrane metalloprotease	contagem
VNG_n072_1	VNG_RS03335	helix-turn-helix domain-containing protein	contagem
VNG_n073_1	VNG_RS03455	MBL fold metallo-hydrolase	contagem
VNG_n074_1	VNG_RS04065	OsmC family protein	contagem
VNG_n075_1	VNG_RS04095	S-layer glycoprotein N-glycosyltransferase AglJ	contagem
VNG_n076_1	VNG_RS04300	tripartite tricarboxylate transporter permease	contagem
VNG_n077_1	VNG_RS04645	halocyanin domain-containing protein	contagem
VNG_n078_1	VNG_RS05225	N-acyl homoserine lactonase family protein	contagem
VNG_n079_1	VNG_RS06360	hypothetical protein	contagem
VNG_n080_1	VNG_RS07120	dipeptide epimerase	contagem
VNG_n081_1	VNG_RS07235	type II methionyl aminopeptidase	contagem
VNG_n082_1	VNG_RS08155	hypothetical protein	contagem
VNG_n083_1	VNG_RS09545	2'-5' RNA ligase family protein	contagem
VNG_n084_1	VNG_RS09725	hypothetical protein	contagem
VNG_n085_1	VNG_RS09950	nicotinate phosphoribosyltransferase	contagem
VNG_n086_1	VNG_RS10195	NADPH-dependent F420 reductase	contagem
VNG_n087_1	VNG_RS11430	antibiotic biosynthesis monooxygenase	contagem

Algumas das regiões encontradas estavam antisense a genes diferencialmente expressos. Os genes *VNG_RS05480*, *VNG_RS05550* e *VNG_RS04300* - que tem as respectivas as regiões VNG_n016_1, VNG_n017_1 e VNG_n76_1 - encontravam-se mais expressos em hiposalinidade e o gene *VNG_RS12755* (região VNG_n027_1) estava

diferencialmente expresso tanto em hiposalinidade, quanto hipersalinidade (Tabela 10). Já os genes *VNG_RS03335*, *VNG_RS04095* e *VNG_RS05225* (*VNG_n72_1*, *VNG_n75_1* e *VNG_n78_1*) estavam menos expressos em hiposalinidade.

Tabela 11: Tabela contendo as 21 regiões intergênicas encontradas, nome dado a elas, nome do cromossomo, início, fim e o método da curadoria.

Nome	Cromossomo	Início da região	Fim da região	Curadoria
VNG_n021_1	NC_002607.1	1332590	1333225	manual
VNG_n022_1	NC_002607.1	1055263	1055329	manual
VNG_n030_1	NC_001869.1	112902	112969	manual
VNG_n031_1	NC_001869.1	114128	114301	manual
VNG_n032_1	NC_001869.1	115618	115690	manual
VNG_n034_1	NC_001869.1	115000	115109	manual
VNG_n035_1	NC_001869.1	123732	123808	manual
VNG_n036_1	NC_001869.1	127008	127080	manual
VNG_n040_1	NC_002608.1	133364	133479	manual
VNG_n042_1	NC_002608.1	136354	136593	manual
VNG_n045_1	NC_002608.1	192101	192180	manual
VNG_n046_1	NC_002608.1	194453	194518	manual
VNG_n048_1	NC_002608.1	196937	196968	manual
VNG_n049_1	NC_002608.1	197277	197383	manual
VNG_n051_1	NC_002608.1	200644	200702	manual
VNG_n054_1	NC_002608.1	212775	212881	manual
VNG_n060_1	NC_002608.1	244442	244801	manual
VNG_n063_1	NC_002608.1	295895	295989	manual
VNG_n066_1	NC_002608.1	324394	324527	manual
VNG_n067_1	NC_002608.1	332229	332696	manual
VNG_n068_1	NC_002608.1	324399	324604	manual

Algumas dessas regiões antisense e intergênicas encontradas (Tabelas 10 e 11), tiveram seu número de *reads* no tratamento de hiposalinidade ou hipersalinidade comparados ao número de *reads* no controle. Obteve-se o valor M ($\log_2(\text{low or high salinity/optimal salinity})$) e utilizou-se os mesmos critérios para encontrar genes diferencialmente expressos (conforme descrito no item 3.4). Foram encontradas mais expressas em hiposalinidade três regiões classificadas com antisense (VNG_n010_1, VNG_n023 e VNG_n053_1) e duas regiões intergênicas (VNG_n060_1 e VNG_n063_1). Estavam menos expressas três regiões antisense (VNG_n002_1, VNG_n004_1 e VNG_n047_1) e cinco regiões intergênicas (VNG_n030_1, VNG_n031, VNG_n032_1, VNG_n034_1 e VNG_n049_1). Das regiões antisense, nenhuma região sense se encontrava diferencialmente expressa. Já sobre as regiões intergênicas, as regiões seguiram o perfil de cobertura dos genes mais próximos na mesma fita, ou seja, quando estavam menos expressos, as regiões também estavam menos expressas. Uma exceção foi a região VNG_n60_1 que estava mais expressa entre os genes *arcB* e *arcD* (VNG_RS11625 e VNG_RS11620), enquanto na mesma condição (hiposalinidade) o gene *arcD* estava menos expresso.

Após observar os resultados, faz-se necessário um estudo mais aprofundado com cada uma das regiões encontradas, como clonagem e superexpressão deles para entendermos melhor sua função na célula. O fato é que os transcritos antisense podem estar afetando a transcrição dos genes tanto de forma positiva quanto negativa, já que foi visto o aumento e diminuição da expressão de genes relacionados a eles (SABERI *et al.*, 2016).

4.9. Quantificação de vesículas de gás

Observou-se que em hiposalinidade as células não flutuavam (Figura 7) e que os genes que codificam as vesículas de gás estavam menos expressos em hiposalinidade. Com a intenção de final de quantificar vesículas de gás em meio hiposalino em experimentos futuros, testou-se um protocolo de purificação de vesículas de gás e a utilização do método de análise de rastreamento de nanopartículas (*Nanoparticle tracking analysis* - NTA) na quantificação dessas. As amostras utilizadas no teste foram duplicatas de vesículas de gás isoladas em meio ótimo (CM 4,3M), e uma amostra em meio hipersalino (CM 5,1M, Tabela S1 H). As amostras em meio hiposalino infelizmente não foram testadas pois não apresentavam camada visível de vesículas de gás, o que acreditou-se indicar um erro no protocolo de purificação, mas que poderia ser o resultado esperado, já que foi observada a diminuição da expressão dos genes relacionados à produção das vesículas.

Em relação à quantificação, observou-se que as amostras em meio ótimo apresentaram concentração média de 7,88e11 partículas/ml enquanto a amostra em meio hipersalino tinha

1,37e12 partículas/ml, quase duas vezes mais partículas que o meio ótimo. Essa informação, no entanto, não é precisa pela falta de replicatas biológicas para a amostra em meio hipersalino.

Quanto ao tamanho das partículas (Figura 20), observou-se que a média de tamanho para as vesículas das três amostras era de aproximadamente 171,3 nm. O equipamento também apresenta ao final da análise um gráfico contendo a variação de concentração de vesículas em diversos tamanhos. Através desse gráfico foi possível observar populações de diversos tamanhos, e enquanto as maiores concentrações de vesículas de gás no meio ótimo tinham tamanho entre 140 e 157 nm, a amostra em meio hipersalino apresentava grande população de partículas com o tamanho de 186 nm. Um problema do dimensionamento de vesículas é que elas são fusiformes e o laser atravessa as partículas em apenas um sentido, que pode ser erroneamente interpretado como maior ou menor por se tratar de um feixe em sentido aleatório nas partículas.

Figura 20: Dimensões das vesículas de gás. Gráficos representando a variação da concentração das vesículas de gás (partículas/ml) em função do tamanho (nm) nas amostras isoladas de células em meio em salinidade ótima (A e B) e em meio hipersalino (C). A área vermelha representa o desvio padrão de triplicatas técnicas e a linha preta representa a média entre as triplicatas.

5. Discussão

5.1. Fenótipo do crescimento da *H. salinarum* sob estresse osmótico

Durante a curva de crescimento utilizando concentrações diferentes de sal, foi observado que em meio hiposalino as culturas não flutuavam sem agitação (Figura 5). A variação brusca de 4,3M para 2,6M pode ser preocupante, já que o aumento de água intracelular para manter o balanço osmótico pode causar a lise das células e desnaturação das proteínas por repulsão dos resíduos de aminoácidos negativamente carregados (MATARREDONA *et al.*, 2020). De acordo com a literatura (BOLLSCHWEILER *et al.*, 2017; VAUCLARE *et al.*, 2020), foi testada a viabilidade das células tanto por contagem quanto por citometria de fluxo nas faixas de 0,5 a 4,5M e 0,5 a 2,5M, respectivamente. Os resultados daqueles testes em ambos experimentos mostraram que em concentração de NaCl próxima a 2,5-2,6M, em torno de 80% das células se encontravam intactas e 20% rompidas, fração que se mantém assim até 4,3M.

Devido à capacidade das vesículas de gás de proporcionar a flutuação das células numa coluna de água, hipotetizou-se que havia uma menor expressão das proteínas relacionadas à elas e que a disponibilidade de oxigênio para a formação das vesículas era baixa. Experimentos que analisaram o knockout de algumas das proteínas estruturais das vesículas de gás demonstraram que as células não flutuavam em cultura (TASHIRO *et al.*, 2016), assim como observado em nosso experimento. A partir das análises dos dados de RNAseq foi possível verificar que alguns dos genes do cluster *gvp* estavam menos expressos (Tabela 3). O cluster *gvp* é formado por 14 genes (representados na figura 21) que atuam na estruturação e também como parte das vesículas de gás. É encontrado nos dois minicromossomos e geram vesículas com cerca de 250 nm de largura e entre 400 e 1500 nm de comprimento (OFFNER *et al.*, 1998). São encontrados nesse cluster 12 asRNAs que poderiam estar atuando na regulação da expressão desses genes, mas nenhum foi diferencialmente expresso nos nossos experimentos.

Figura 21: A) Representação do cluster *gvp* e promotores. Adaptado de Pfeifer *et al.* (2012) B) Imagem retirada do IGV com a representação dos genes e em vermelho, RNAs antisense associados a eles.

Na halófila moderada *Haloferax mediterranei* foi comprovado que a formação das vesículas de gás é influenciada pela concentração de cloreto de sódio (ENGLERT; HORNE; PFEIFER, 1990). Este microrganismo tem seu crescimento ótimo em concentrações de sal de 20% no meio ou 3,4 M de NaCl (CUI *et al.*, 2017) e observou-se que concentrações abaixo de 17% ou 2,9 M de NaCl faziam com que as células produzissem raras ou nenhuma vesícula.

Comparada à *Halobacterium salinarum*, cujo crescimento ocorre em meio com 25% de sal ou 4,3M de sal, o tratamento com meio hiposalino (15% ou 2,6M de NaCl) também gerou o mesmo resultado, com menor produção das vesículas de gás comprovada pela menor expressão dos genes do cluster *gvp*. Em hipersalinidade o fenótipo de menor expressão das vesículas de gás não foi observado.

De acordo com a literatura, um dos motivos que causam a diminuição da expressão das vesículas de gás é a expressão do gene *gvpD*, que atua como repressor da transcrição do resto do cluster. A presença de oxigênio também parece ser limitante para a produção de vesículas. Sabe-se que em meio anaeróbico as vesículas de gás não são expressas e até mesmo mutantes sem o repressor GvpD têm a produção de vesículas de gás diminuídas (OREN, 2012). O gene *gvpD* se encontrava menos expresso em hiposalinidade, assim como todos os outros genes do cluster *gvp* (alguns menos expressos, mas não diferencialmente expressos), indicando que a resposta à salinidade parece ser semelhante à resposta ao estresse por falta de oxigênio. Isso faz sentido quando falamos de hipersalinidade, já que a solubilidade do oxigênio é diminuída em soluções saturadas de sal (OREN, 2012), mas em hiposalinidade a explicação para a diminuição da expressão do cluster *gvp* não está clara. O experimentos de proteômica encontrou também o gene *gvpH* menos expresso em hiposalinidade e *gvpO* mais expresso em hipersalinidade, mas esses resultados não foram

discutidos com profundidade nesta publicação, nem nas publicações mais recentes abrangendo a resposta de *H. salinarum* NRC-1 aos estresses ambientais (HACKLEY; SCHMID, 2019; VAUCLARE et al., 2020).

O fator de transcrição *tbpD* (*TATA-binding protein D*) também pode estar relacionado à regulação do cluster *gvp* em *H. salinarum* já que o aumento de sua transcrição afeta negativamente a transcrição genes *gvpFGHIJKLM* e o knockout desse fator de transcrição além de resultar em aumento da transcrição dos genes *gvpFGHIJKL* e *M*, causa o aumento da expressão de *gvpDEACN*, e *O* (YAO; FACCIOTTI, 2011). Em hiposalinidade foi observado que houve a menor expressão de *tbpD* em apenas uma das réplicas, o que impediu de ser classificado como diferencialmente expresso. Diferente do descrito anteriormente, nossos dados em hiposalinidade houve a diminuição da expressão dos genes *gvpDACNOGIJ* e *K*. Enquanto alguns organismos possuem apenas um gene codifica as *TATA-binding proteins*, *H. salinarum* apresenta em seu genoma seis genes *tbp* (*tbpA* a *tbpF*), que acredita-se estar relacionado com o início de transcrição em diferentes promotores ou em diferentes condições de crescimento (TEUFEL; PFEIFER, 2010). Nesse sentido, observou-se a maior expressão de duas cópias do gene *tbpC* (*TATA-binding protein C*), o que sugere que esse fator de transcrição pode estar atuando na resposta celular à hiposalinidade.

A menor expressão de vesículas de gás possivelmente fez com que fosse observada uma mudança de coloração da cultura em meio hiposalino. Este fenômeno já havia sido reportado anteriormente para a espécie *Haloarcula marismortui* RR12 sob estresse hipersalino. No estudo, foi utilizado o método de micrografia eletrônica de varredura, possibilitando medir a absorbância a 490nm, que visa detectar a produção de pigmentos como carotenóides e bacteriorrubrina que conferem a coloração avermelhada às células (THOMBRE et al., 2016) e observou-se um aumento na concentração desses pigmentos para aquela espécie em hipersalinidade. e também nos experimentos de deleção de genes do cluster *gvp* (DASSARMA et al., 2013). A bacteriorodopsina, que tem um papel na biossíntese de bacteriorrubrina (PECK et al., 2017), foi encontrada diferencialmente expressa nos dados de RNA-seq e proteoma. No entanto no RNA-seq estava menos expressa, enquanto no proteoma estava mais expressa (LEUKO et al., 2009).

5.2. Reprodutibilidade dos dados de RNAseq e coerência com literatura

Foi visto que a diferença entre as replicatas do experimento de RNA-seq afetou bastante na seleção de genes diferencialmente expressos (Figuras 8 e 11). Essas diferenças entre replicatas confirmam a grande plasticidade do genoma de *H. salinarum* NRC-1, cujo

material genético sofre várias mutações pela presença de sequências de inserção e a frequência em que isso ocorre é extremamente alta, chegando a 1% (PFEIFER; WEIDINGER; GOEBEL, 1981; DASSARMA, 1989). O maior número de genes diferencialmente expressos em hiposalinidade refletiu a maior concordância entre as réplicas desse tratamento. Já em hipersalinidade o número de genes diferencialmente expressos foi baixo em relação a hiposalinidade, fazendo-se possível notar que o critério para determinar os genes diferencialmente expressos foi muito rigoroso. Uma maneira simples de contornar esse resultado seria utilizar a média entre os valores M das replicatas, mas assim estaríamos desconsiderando totalmente a plasticidade do genoma e frequência de mutações que ocorrem nesse organismo. Outra forma de diminuir as diferenças entre réplicas e melhorar a reprodutibilidade do experimento seria aumentar o número de amostras sequenciadas para cada tratamento. Dada a disponibilidade de kits de sequenciamento, utilizamos apenas duplicatas para cada condição.

Na integração dos dados obtidos com dados de proteômica e microarranjos de DNA (Figuras 15, 16, 17 e 18) além da diferença entre réplicas, muitos dados estavam incoerentes com o obtido (RNA-seq e microarranjos de DNA). Essa incoerência entre os dados é possivelmente relacionada às diferentes condições de cultivo de cada experimento, como a temperatura de incubação, rotação da incubadora e até a concentração de NaCl no caso dos microarranjos de DNA. A definição de diferencialmente expresso também foi diferente para cada um dos experimentos: para o proteoma foi valores acima ou abaixo de zero do $\log_{10}$ da fração entre quantificação de peptídeos no tratamento e no controle; para os microarranjos de DNA foi valores de $|fold\ change| \geq 1,3$ da intensidade de sinais do tratamento e controle; e para os nossos dados foi o valor $|M| \geq 1$, que representa o $\log_2$ da fração entre quantificação de *reads* no tratamento e no controle. No entanto, para facilitar a comparação entre os experimentos, todos foram convertidos para o valor M, o que fez com que proteínas ou genes que eram considerados em seu experimento original como diferencialmente expressos, pudessem não estar quando convertidos ao valor de M.

Sabe-se que a regulação da expressão gênica está profundamente relacionada com o ciclo celular (LUNDGREN; BERNANDER, 2007; DARNELL et al., 2020), o que influencia na abundância de transcritos e proteínas. Pelas curvas de crescimento apresentadas nas publicações, observou-se que houve uma diferença no crescimento das culturas e que a coleta de amostras para os ensaios de proteômica e microarranjos de DNA ocorreram em tempos diferentes, mesmo que visando a mesma fase do ciclo celular (fase exponencial). Isto indica que ao comparar os experimentos podem ter ocorrido também variações de fases de crescimento.

5.3. Produção de intraRNAs sob estresse osmótico

Nos estudos anteriores do nosso grupo de pesquisa foi demonstrado que existem sítios de início de transcrição internos a regiões codificantes (iTSS) em *H. salinarum* NRC-1 e que a partir desses iTSSs são gerados transcritos que por sua vez são traduzidos em proteínas. Foram encontrados mais de 200 contendo iTSS, mas apenas alguns tiveram a tradução das intraproteínas validadas (TEN-CATEN et al., 2018). Entre elas estão intraproteínas identificadas por *western blot* referentes aos genes *kef1* (VNG_RS07995), *asc3* (VNG_RS05220) e *orc4* (VNG_RS11810).

A Acs3 é uma acil-CoA sintetase que atua na biossíntese e degradação de ácidos graxos. A Orc4 é uma proteína que auxilia na replicação do DNA através da formação do complexo de reconhecimento da origem de replicação (ORC) (AUSIANNIKAVA et al., 2018). A Kef1 é uma proteína de membrana transportadora de prótons em um sistema antiporter que possui domínios para ligação com Na⁺ e K⁺. Entre os três genes, nenhum foi diferencialmente expresso e verificamos o perfil de expressão do intraRNA de *kef1* de acordo com a cobertura dos *reads* (Figura S1), em que a região que dá origem ao intraRNA apresentava quantidade expressivamente maior que o resto do transcrito, mas ainda assim era bem semelhante ao controle.

A presença de outras intraproteínas foi validada utilizando bancos de dados de espectrometria de massas (TEN-CATEN et al., 2018) e verificamos se ocorria a expressão diferencial dos intraRNAs associados aos genes de um transportador ABC (VNG_RS11500), da acetato-coA ligase (VNG_RS03880), de uma oxidoredutase dependente de NAD(P)/FAD (VNG_RS07005), outro transportador ABC com ferro como possível substrato (VNG_RS03600) e uma subunidade da succinato desidrogenase (VNG_RS05105). Dois genes se encontravam diferencialmente expressos em hiposalinidade, o gene da acetato-coA ligase menos expresso e o transportador ABC com ferro como possível substrato mais expresso. Em hipersalinidade, o gene do transportador ABC estava menos expresso. Visualmente nenhum desses genes apresentava região de intraRNA diferencialmente expressa, mas estudos mais aprofundados devem ser aplicados aos genes citados e aos outros 191 genes contendo iTSS para verificar se existe a expressão diferencial dos intraRNAs sob estresse osmótico.

5.4. Enriquecimento funcional dos genes diferencialmente expressos

Através do enriquecimento funcional dos genes diferencialmente expressos observou-se que no tratamento de hipersalinidade a resposta das células foi diminuir o transporte transmembrana de ânions e fosfato, diminuir a biossíntese aminoácidos em geral em especial a arginina, indicando que a obtenção de energia é feita de forma aeróbica. Mas também observou-se a diminuição da expressão dos genes codificando enzimas do ciclo do ácido tricarboxílico e ciclo da ureia. Estavam mais expressos fatores de iniciação de transcrição e tradução, o processo de biossíntese de RNA e expressão gênica e biossíntese de macromoléculas. Pudemos observar a expressão diferencial de proteínas de estresse universal, que atuam no transporte de aminoácidos

Em hiposalinidade observou-se que o transporte de ânions e fosfato estavam menos expressos, assim como em hipersalinidade. Além desses, outros mecanismos importantes se encontravam menos expressos, como a biossíntese de carotenóides, motilidade dependente do arqueolo, a biossíntese de cobalamina (vitamina B12) e outras macromoléculas, a detecção e resposta a estímulos de luz, quimiotaxia, a divisão celular e tradução. Estavam mais expressos as respostas ao dano de DNA, transporte transmembrana de cloreto e biossíntese de ácidos orgânicos.

Assim como descrito por Coker et. al (2007), dois genes que atuam como transportadores de potássio foram mais expressos para compensar a perda de K^+ em baixa salinidade, mas também foram encontrados outros possíveis transportadores de potássio menos expressos. Foram encontrados transportadores de sódio menos expressos nos dois tratamentos, diferente do proposto no trabalho de Coker et. al (2007) onde apenas em hiposalinidade estavam menos expressos, o que indicaria que aumento na concentração de sódio intracelular é tolerado pela população. Além disso, transportadores de ferro e fosfato estavam menos expressos nos dois tratamentos, sugerindo uma resposta de menor recrutamento desses íons que podem ser tóxicos às células. Os transportadores de dipeptídeos que se encontravam mais expressos em ambas condições na literatura estavam menos expressos somente em hiposalinidade. Uma possível explicação para esses dados é a diferença entre as duas réplicas crescidas em meio hipersalino, em que vários genes não obtiveram contagem de *reads* e portanto não puderam passar pelos filtros aplicados nos dados.

Acredita-se que as proteína quinases estão envolvidas na resposta ao estresse osmótico em eucariotos e alguns procariotos por atuarem como transdutores de sinais (PIAO et

al., 1999; SHOUMSKAYA *et al.*, 2005; COKER *et al.*, 2007). Em hipersalinidade foi observada uma histidina quinase menos expressa e em hiposalinidade uma quinase com resposta adaptativa que atua na regulação da transcrição estava mais expressa.

Confirmando a afirmação de que o estresse salino pode envolver respostas características de outros estresses, em hiposalinidade *sod1* estava mais expressa enquanto *sod2* estava menos expressa. O fenômeno pode estar relacionado às substâncias que atuam como cofatores para as duas enzimas: *sod1* pode estar ligada a zinco e cobre, enquanto *sod2* se liga a magnésio (MILLER, 2012; NOJIMA *et al.*, 2015). Em hipersalinidade nenhuma superóxido dismutase foi encontrada diferencialmente expressa. Já a glutathione-S-transferase, outra enzima detoxificante, estava mais expressa em hiposalinidade, assim como no estudo de Coker *et al.* (2007). Um gene relacionado à resposta ao estresse por aumento de temperatura se encontrava mais expresso em hiposalinidade, mas nenhum dos genes de resposta ao calor estava diferencialmente expresso nos experimentos.

5.5. O proteoma e sua concordância com o RNA-seq

Alguns genes que não foram identificados nos microarranjos de DNA foram observados no proteoma, como a baixa expressão dos genes relacionados ao arqueolo e alguns genes do operon *gvp*. No proteoma foi observada uma possível maior expressão da bacteriorodopsina em ambas as condições, o que não foi observado no transcrito utilizando o RNA-seq. Ao contrário, o gene da bacteriorodopsina era um dos menos expressos em hiposalinidade no RNA-seq, o que era esperado já que sua produção é induzida em condições anaeróbicas - confirmada no proteoma em hipersalinidade, onde a solubilidade do oxigênio é diminuída mimetizando a condição de anaerobiose (LEUKO *et al.*, 2009; OREN, 2012; PFEIFER, 2015)

Este fenômeno observado pode ser creditado a erros experimentais, às mutações frequentes no genoma de *H. salinarum*, às desvantagens dos métodos de obtenção dos dados e a regulação dependente da abundância de transcritos e das fases de crescimento em que as células se encontram. Essa última, foi estudada em profundidade no trabalho de Lomana *et al.* (2020), onde foi descrita a tradução seletiva de transcritos menos expressos em *H. salinarum*. Neste trabalho, além de demonstrar a variação de transcritos e proteínas referentes ao ribossomo durante as fases celulares, indicam que a composição protéica do ribossomo varia em resposta às condições a que as células são submetidas (LÓPEZ GARCÍA DE LOMANA *et al.*, 2020).

Nos dados do proteoma de Leuko *et al.* (2009) foram encontrados muitas proteínas ribossomais menos expressas em hipo e hipersalinidade, enquanto no transcrito por RNA-seq muitos dos genes se encontravam mais expressos. Um exemplo é a proteína ribossomal 50s L15e, que estava mais expressa no RNA-seq (média de $M=1,3$) e menos expressa no proteoma ($M=-1,8$). Isso poderia indicar que o processo de seleção de transcritos e produção de complexos ribossomais específicos para a condição de estresse osmótico estavam sendo formados. No entanto, para verificarmos especificamente quais proteínas compõem os complexos ribossomais, somente a expressão diferencial do transcrito não é suficiente, sendo necessários mais estudos sobre cada uma dessas proteínas e sua importância para a célula.

5.6. Metabolismo de glicerol está mais expresso em hiposalinidade

Na condição de hiposalinidade, observou-se que os genes que codificam as três subunidades da enzima glicerol-3-fosfato desidrogenase estavam mais expressas, o que nos fez testar o efeito dessa alta expressão no fenótipo de crescimento das culturas. Além dessa enzima, mas duas se encontravam mais expressas: a gliceraldeído-3-fosfato desidrogenase e a fosfoglicerato quinase. No habitat natural, eucariotos halofílicos como algas e fungos também estão presentes. Estes produzem o glicerol para manter seu balanço osmótico e podem liberá-lo quando ocorre o estresse hipotônico ou lise celular (ELEVI BARDAVID; KHRISTO; OREN, 2008). Quando suplementamos as amostras em hiposalinidade com duas concentrações de glicerol, o glicerol conseguiu entrar nas células por difusão e serem utilizados pela maquinaria celular (WILLIAMS *et al.*, 2017).

Figura 22: Metabolismo de glicerol em haloarqueias. Em azul claro é representado o meio extracelular e uma célula de haloarqueia transportando solutos compatíveis do exterior para auxiliar na glicólise, gluconeogênese e síntese de glicerofosfolípidos. Imagem adaptada de Williams et. al.(2017).

As etapas da via de utilização do glicerol (Figura 22) até a diidroxiacetona fosfato não está tão clara com os dados encontrados, pois a glicerol quinase nem a glicerol desidrogenase estavam diferencialmente expressas, mas uma das réplicas tinha a expressão aumentada de glicerol desidrogenase. Portanto, a via no sentido de produção de diidroxiacetona fosfato (DHAP) a partir da glicerol-3-fosfato desidrogenase não deveria acontecer supondo que o objetivo final seja a produção de DHAP. O fato é que, as reações catalisadas pela glicerol quinase, diidroxiacetona quinase e glicerol-3-fosfato desidrogenase culminam na formação de DHAP. A isomerização de DHAP pode acontecer formando o gliceraldeído-3-fosfato que serve de substrato para a gliceraldeído-3-fosfato desidrogenase, formando o 1-3-bisfosfoglicerato que pode ser utilizado na glicólise e via de fixação de carbono. Ademais, a fosfoglicerato quinase tem como substrato o 1-3-bisfosfoglicerato e o produto 3-fosfoglicerato é formado (OUELLETTE; MAKKAY; PAPKE, 2013; WILLIAMS *et al.*, 2017b). O 3-fosfoglicerato serve de esqueleto para

a formação de glicerofosfolídeos (OUELLETTE; RAWN, 2018), formação de piruvato, biotina e outras coenzimas e cofatores, e também a formação de metionina e S-adenosil-metionina, importante como doador de grupo metil para várias reações, incluindo as modificações de tRNA.

5.7. Arquelos não se apresentaram funcionais em condições hiposalinas

As arquelinas são proteínas que formam os filamentos helicais que compõem os arquelos, nome dado às estruturas flagelo-like presentes em algumas arqueias (ALBERS; JARRELL, 2018). A figura 23 representa a organização dos genes do operon. São eles: arquelina A1, A2, B1, B2 e B3, que são as proteínas que compõem a fração extracelular do arquelo, desde o gancho (B3) até o alongamento do arquelo (A1, A2, B1, B2). As regiões codificantes são referentes ao ancoramento da estrutura à membrana (F,G), hidrólise de ATP para movimentação do arquelo (C/D/E, H, I, J) e retirada de peptídeo sinal das arquelinas (K) (ALBERS; JARRELL, 2018; KINOSITA; NISHIZAKA, 2018).

Os resultados da integração dos dados mostraram que os RNAs e proteínas referentes aos produtos arquelina A2, B1, B2, e B3 estavam menos expressos. Para entender melhor o fenótipo observado, fez-se a cultura de *H. salinarum* NRC-1 em meios semi-sólidos à 2,3M de NaCl e à 4,3M de NaCl. Enquanto o controle (CM 4,3M de NaCl) apresentou alta motilidade, a hiposalinidade resultou em baixa motilidade (Figura 23). Os resultados foram os mesmos observados por LOSENSKY *et al.* (2015) quando foi feito o *knockout* do gene *flaI* na estirpe R1 da *H.salinarum* e ZARETSKY *et al.* (2019) com o *knockout* do gene *aglB*, uma glicosil transferase que atua na N-glicosilação das arquelinas.

Figura 23: Operon arqueio e representação do modelo motor das arqueias. Adaptado de (KINOSITA; NISHIZAKA, 2018).

Além da baixa expressão dos genes que formam o arqueio, houve a baixa expressão da proteína da família de rodopsinas VNG_RS06435, que atua como transdutor de sinal para a formação do arqueio (Uniprot:P0DMI3). Um dos RNAs antisense encontrados diferencialmente expressos é justamente antisense ao gene *aglB*, da N-glicosilação. Essa região antisense se encontrava mais expressa, enquanto o gene *aglB* estava menos expresso, mas não diferencialmente expresso. A alta expressão do RNA antisense e baixa expressão do gene podem indicar uma possível regulação da transcrição da glicosil transferase. Estudos mais aprofundados para entender a possível regulação envolvem a superexpressão da região antisense e análise do fenótipo e transcriptoma, que podem ser realizadas no futuro.

5.8. Alterações no transcriptoma sugerem mudanças na maquinaria de tradução

Muitos genes relacionados à tradução estavam diferencialmente expressos em ambos tratamentos, representados pelas proteínas ribossomais, transferases e outras enzimas que atuam na modificação de tRNAs. Observou-se um aumento na expressão de genes relacionados ao reparo de DNA, endonucleases, ribonucleases e biossíntese de bases

nitrogenadas em hiposalinidade. No trabalho de Leuko *et al.* (2009) em ambas as condições a maquinaria de tradução estava menos expressa e em hiposalinidade observaram um aumento da expressão de genes da biossíntese de nucleotídeos que seriam usados no reparo de dano ao DNA.

O tRNA é uma das moléculas que mais sofrem modificações pós-transcricionais e sua função na tradução pode ser melhorada através dessas modificações (CRÉCY-LAGARD; JARROCH, 2021). A enzima capaz de modificar tRNAs mais expressa em hipersalinidade é a 4-demetil wiosina sintase TYW1 (VNG_RS08250). Essa enzima atua na modificação da posição 37 do tRNA ligado à fenilalanina, formando uma wibutosina que auxilia na interação códon-anticódon e evita erros (DE CRÉCY-LAGARD *et al.*, 2010; PERCHE-LETUVÉE *et al.*, 2012; YOUNG; BANDARIAN, 2018, p. 1). tRNAs que não sofreram essa modificação já foram encontrados em células tumorais de mamíferos (WAAS; DE CRÉCY-LAGARD; SCHIMMEL, 2005).

Em hiposalinidade a 4-demetil wiosina sintase também foi mais expressa, juntamente com a tRNA(Phe) (4-demetil wiosine(37) -C(7)) amino carboxi propil transferase, que também atua em um dos vários passos para formação da wibutosina (DE CRÉCY-LAGARD *et al.*, 2010; PERCHE-LETUVÉE *et al.*, 2012). Além desses, os genes tRNA 4-tiouridina (8) sintase (VNG_RS05070) e a tRNA pseudouridina (54)-N(1))-metil transferase (VNG_RS07670), estavam mais expressos. A tRNA 4-tiouridina (8) sintase atua na formação de 4-tiouridina na posição 8 do tRNA e está presente em todos os domínios da vida. A função da 4-tiouridina no tRNA é auxiliar a estabilização da sua estrutura e tem papel importante como sensor de radiação UV e proteção da célula (LIU *et al.*, 2012; NEUMANN *et al.*, 2014). A tRNA pseudouridina (54)-N(1))-metil transferase atua na N-metilação da pseudouridina na posição 54 do tRNA, que faz com que sejam estabilizadas as interações entre os braços T e D do tRNA de bactérias termofílicas (WURM *et al.*, 2012).

Um gene foi menos expresso em hiposalinidade: SAM-dependent methyl transferase (VNG_RS04345). Esse gene apresenta em sua estrutura o domínio TrmO, responsável pela N6-metilação da N6-treonil carbamoil adenosina na posição 37 do tRNA. A N6-treonil carbamoil adenosina responsável pelo reconhecimento de códons que se iniciam com A, ligação dos aminoácidos ao tRNA e manutenção da fase de leitura (MIYAUCHI; KIMURA; SUZUKI, 2013). Não foram observados diferencialmente expressos os genes que atuam na formação da S-adenosil-L-metionina (SAM), doadora do grupo metil para as metiltransferases citadas anteriormente.

Em hipersalinidade estavam mais expressos os genes relacionados a proteínas que se ligam à subunidade 30S do ribossomo enquanto em hiposalinidade genes relacionados a ambas subunidades (30S e 50S) se encontravam mais expressos. Todos os genes diferencialmente expressos observados indicam que há uma mudança na maquinaria de tradução mais expressiva em baixas concentrações de sal com o intuito de melhorar a interação tRNA-ribossomo, a ligação de aminoácidos ao tRNA e evitar erros inerentes a interação códon-anticódon.

6. Considerações finais

O estresse osmótico ao qual a *H. salinarum* NRC-1 é submetida naturalmente em seu habitat pode trazer danos irreversíveis às células. Entre esses danos se encontram a perda de estrutura e função de suas biomoléculas e o rompimento da membrana celular em alguns casos. Durante sua evolução, esse microorganismo conseguiu desenvolver meios para sobreviver ao ambiente tão extremo em que habita e às variações dele.

Quando analisamos o transcrito desse microorganismo sob hiposalinidade e hipersalinidade, observou-se que a baixa concentração de sal fez com que as células aumentassem a expressão de genes relacionados à sua maquinaria de tradução, a resposta a radiação e reparo do DNA, síntese de aminoácidos e de bases nitrogenadas. Além disso genes relacionados a processos considerados não essenciais durante o estresse, como a motilidade através dos arqueos e vesículas de gás, divisão celular, produção de glicoproteínas de membrana, transporte de fosfato e peptídeos para o interior da célula e obtenção de energia através das bacteriorodopsinas estavam menos expressos (Figura 24). Em hipersalinidade, o número menor de genes diferencialmente expressos possivelmente devido ao critério utilizado para a seleção dos genes. Nessa condição, houve um aumento da expressão de proteínas que atuam como reguladores da transcrição e tradução e biossíntese de macromoléculas. Foram menos expressos os processos de transporte de fosfato e outros ânions, biossíntese de arginina e outros aminoácidos e o ciclo da ureia (Figura 24).

Observamos que a análise dos genes diferencialmente expressos considerando cada uma das réplicas nesse caso foi eficiente e confirmou a frequência da variabilidade que pode ocorrer no genomas até mesmo durante o tempo de uma curva de crescimento. O uso de experimentos simples como a curva de crescimento com glicerol, ensaio de motilidade e observação das culturas após a curva de crescimento ajudaram na validação dos genes diferencialmente expressos encontrados.

Estudos englobando o estresse osmótico já haviam sido realizados anteriormente, justamente por serem condições que a população de *H. salinarum* encontram normalmente em seu habitat. Os trabalhos de Coker *et al.* (2007) e Leuko *et. al* (2009), que inspiraram este trabalho, mostraram que a lacuna de tempo e metodologia utilizadas e a disponibilização de dados têm um efeito grande no que conseguimos comparar e integrar dos experimentos. Com a maior resolução que o experimento de RNA-seq proporciona foi possível encontrar mais genes diferencialmente expressos, como os genes do cluster *gvp*, genes da família das rodopsinas e genes relacionados ao metabolismo de glicerol. Com os dados do RNA-seq ainda foi possível a identificação de novos transcritos que podem auxiliar no entendimento do mecanismo de regulação gênica nas haloarqueias. O conjunto de informações obtidas deverá complementar os estudos de regulação gênica global de *H. salinarum* utilizando uma abordagem sistêmica, contribuindo para a construção de modelos globais em condições fisiologicamente relevantes.

Figura 24: Representação dos principais processos biológicos diferencialmente expressos em hiposalinidade e hipersalinidade. Em letras vermelhas estão os processos mais expressos e em letras verdes estão os processos menos expressos.

7. Bibliografia

- ABDELMOHSEN, K. Noncoding RNA. In: SCHWAB, M. (Ed.). **Encyclopedia of Cancer**. Berlin, Heidelberg: Springer, 2016. p. 1–5.
- ABRAMSON, J. et al. The Structure of the Ubiquinol Oxidase from Escherichia Coli and Its Ubiquinone Binding Site. **Nature Structural Biology**, v. 7, n. 10, p. 910–917, out. 2000.
- AHARON OREN. Halophilic microorganisms and their environments. v. 5, n. 2, [s.d.]
- ALBERS, S.-V.; JARRELL, K. F. The Archaeum: An Update on the Unique Archaeal Motility Structure. **Trends in Microbiology**, v. 26, n. 4, p. 351–362, 2018.
- ALEXA, A.; RAHNENFUHRER, J. **topGO: Enrichment Analysis for Gene Ontology**. [s.l.] Bioconductor version: Release (3.11), 2020.
- ALQUÉRES, S. M. C. et al. Exploring the biotechnological applications in the archaeal domain. **Brazilian Journal of Microbiology**, v. 38, n. 3, p. 398–405, set. 2007.
- ANDERSON, N. L.; ANDERSON, N. G. Proteome and Proteomics: New Technologies, New Concepts, and New Words. **ELECTROPHORESIS**, v. 19, n. 11, p. 1853–1861, 1998.
- ARAKAWA, T.; TIMASHEFF, S. N. Preferential interactions of proteins with salts in concentrated solutions. **Biochemistry**, v. 21, n. 25, p. 6545–6552, 7 dez. 1982.
- ARRAIANO, C. M. Post-Transcriptional Control of Gene Expression: Bacterial mRNA Degradation. **World Journal of Microbiology and Biotechnology**, v. 9, n. 4, p. 421–432, 1 jul. 1993.
- AUSIANNIKAVA, D. et al. Evolution of Genome Architecture in Archaea: Spontaneous Generation of a New Chromosome in Haloferax volcanii. **Molecular Biology and Evolution**, v. 35, n. 8, p. 1855–1868, ago. 2018.
- BALIGA, N. S. et al. Systems Level Insights Into the Stress Response to UV Radiation in the Halophilic Archaeon Halobacterium NRC-1. **Genome Research**, v. 14, n. 6, p. 1025–1035, jun. 2004.
- BALIGA, N. S. The Scale of Prediction. **Science**, v. 320, n. 5881, p. 1297–1298, 6 jun. 2008.
- BAXTER, B. Great Salt Lake microbiology: a historical perspective. **International Microbiology**, v. 21, 4 jun. 2018.
- BESTEN, H. M. W. den et al. Phenotypic and Transcriptomic Analyses of Mildly and Severely Salt-Stressed Bacillus Cereus ATCC 14579 Cells. **Applied and Environmental Microbiology**, v. 75, n. 12, p. 4111–4119, 15 jun. 2009.
- BETLACH, M. C.; SHAND, R. F.; LEONG, D. M. Regulation of the Bacterio-Opsin Gene of a Halophilic Archaeobacterium. **Canadian Journal of Microbiology**, v. 35, n. 1, p. 134–140, jan. 1989.

BHAGAVAN, N. V.; HA, C.-E. Chapter 23 - RNA and Protein Synthesis. In: BHAGAVAN, N. V.; HA, C.-E. (Ed.). **Essentials of Medical Biochemistry (Second Edition)**. San Diego: Academic Press, 2015. p. 419–446.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014.

BOLLSCHWEILER, D. et al. Cryo-Electron Microscopy of an Extremely Halophilic Microbe: Technical Aspects. **Extremophiles**, v. 21, n. 2, p. 393–398, 1 mar. 2017.

BONNEAU, R. et al. A Predictive Model for Transcriptional Control of Physiology in a Free Living Cell. **Cell**, v. 131, n. 7, p. 1354–1365, 28 dez. 2007.

CAMPBELL, C. E. Temporal salinity variation in salt evaporation basins in south-eastern Australia. **International Journal of Salt Lake Research**, v. 4, n. 1, p. 45–55, mar. 1995.

CATEN, F. ten. **A transcrição pervasiva na archaea *Halobacterium salinarum* NRC-1 e a identificação de novos transcritos**. 2017. Universidade de São Paulo, 2017. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/95/95131/tde-17042019-143232/>>. Acesso em: 15 out. 2020.

CAVICCHIOLI, R. et al. Pathogenic Archaea: Do They Exist? **BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology**, v. 25, n. 11, p. 1119–1128, nov. 2003.

CAVICCHIOLI, R. Archaea - Timeline of the third domain. **Nature Reviews Microbiology**, v. 9, n. 1, p. 51–61, 2011.

CDD Conserved Protein Domain Family: PBP2_UgpB. Disponível em: <<https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=270451>>. Acesso em: 30 mar. 2021.

CHEN, H.; JIANG, J. G.; WU, G. H. Effects of salinity changes on the growth of *Dunaliella salina* and its isozyme activities of glycerol-3-phosphate dehydrogenase. **Journal of Agricultural and Food Chemistry**, v. 57, n. 14, p. 6178–6182, 2009.

CHIBANI, C. M. et al. **A comprehensive analysis of the global human gut archaeome from a thousand genome catalogue**. [s.l.] **Bioinformatics**, 22 nov. 2020. . Disponível em: <<http://biorxiv.org/lookup/doi/10.1101/2020.11.21.392621>>. Acesso em: 9 jan. 2021.

CLEMENTINO, M. M. et al. Archaeal Diversity in Naturally Occurring and Impacted Environments from a Tropical Region. **Journal of Applied Microbiology**, v. 103, n. 1, p. 141–151, 2007.

COKER, J. A. et al. Transcriptional profiling of the model Archaeon *Halobacterium* sp. NRC-1: responses to changes in salinity and temperature. **Saline systems**, v. 3, p. 6, jul. 2007.

COKER, J. A. et al. Multiple replication origins of *Halobacterium* sp. strain NRC-1: Properties of the conserved *orc7*-dependent *oriC1*. **Journal of Bacteriology**, v. 191, n. 16, p. 5253–5261,

2009.

CRÉCY-LAGARD, V. de; JAROCH, M. Functions of Bacterial tRNA Modifications: From Ubiquity to Diversity. **Trends in Microbiology**, v. 29, n. 1, p. 41–53, 1 jan. 2021.

CRICK, F. Central Dogma of Molecular Biology. **Nature**, v. 227, n. 5258, p. 561–563, ago. 1970.

CUI, Y.-W. et al. Salinity Effect on Production of PHA and EPS by *Haloferax Mediterranei*. **RSC Advances**, v. 7, n. 84, p. 53587–53595, 2017.

DARNELL, C. L. et al. The Ribbon-Helix-Helix Domain Protein CdrS Regulates the Tubulin Homolog FtsZ2 To Control Cell Division in Archaea. **mBio**, v. 11, n. 4, 25 ago. 2020. Disponível em: <<https://mbio.asm.org/content/11/4/e01007-20>>. Acesso em: 14 maio. 2021.

DASSARMA, S. Mechanisms of Genetic Variability in *Halobacterium Halobium*: The Purple Membrane and Gas Vesicle Mutations. **Canadian Journal of Microbiology**, v. 35, n. 1, p. 65–72, jan. 1989.

DASSARMA, S. et al. An improved genetic system for bioengineering buoyant gas vesicle nanoparticles from Haloarchaea. **BMC Biotechnology**, v. 13, n. March 2014, 2013.

DASSARMA, S.; COKER, J. A.; DASSARMA, P. Archaea (overview). In: SCHAECHTER, M. (Ed.). **Encyclopedia of Microbiology (Third Edition)**. Oxford: Academic Press, 2009. p. 1–23.

DE ALMEIDA, J. P. P. et al. The Primary Antisense Transcriptome of *Halobacterium salinarum* NRC-1. **Genes**, v. 10, n. 4, 5 abr. 2019. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6523106/>>. Acesso em: 30 jul. 2020.

DE CRÉCY-LAGARD, V. et al. Biosynthesis of Wyosine Derivatives in tRNA: An Ancient and Highly Diverse Pathway in Archaea. **Molecular Biology and Evolution**, v. 27, n. 9, p. 2062–2077, set. 2010.

DE CRÉCY-LAGARD, V. et al. Comparative Genomics Guided Discovery of Two Missing Archaeal Enzyme Families Involved in the Biosynthesis of the Pterin Moiety of Tetrahydromethanopterin and Tetrahydrofolate. **ACS chemical biology**, v. 7, n. 11, p. 1807–1816, 16 nov. 2012.

DEOLE, R. et al. An Extremely Halophilic Proteobacterium Combines a Highly Acidic Proteome with a Low Cytoplasmic Potassium Content. **The Journal of Biological Chemistry**, v. 288, n. 1, p. 581–588, 4 jan. 2013.

DRANCOURT, M. et al. Methanobrevibacter Smithii Archaemia in Febrile Patients with Bacteremia, Including Those with Endocarditis. **Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America**, 15 jul. 2020.

DU, Y. et al. Integration of Transcriptomic and Proteomic Data Identifies Biological Functions in Cell Populations from Human Infant Lung. **American Journal of Physiology. Lung Cellular and Molecular Physiology**, v. 317, n. 3, p. L347–L360, 01 2019.

DYALL-SMITH, M. The halohandbook. **Protocols for haloarchaeal genetics**. Edited by: **Dyall-Idots**, n. March, p. 1–144, 2006.

ELEVI BARDAVID, R.; KHRISTO, P.; OREN, A. Interrelationships between Dunaliella and Halophilic Prokaryotes in Saltern Crystallizer Ponds. **Extremophiles**, v. 12, n. 1, p. 5–14, 1 jan. 2008.

EME, L. et al. Archaea and the Origin of Eukaryotes. **Nature Reviews Microbiology**, v. 15, n. 12, p. 711–723, dez. 2017.

EME, L.; DOOLITTLE, W. F. Archaea. **Current Biology**, v. 25, n. 19, p. R851–R855, 5 out. 2015.

ENGLERT, C.; HORNE, M.; PFEIFER, F. Expression of the Major Gas Vesicle Protein Gene in the Halophilic Archaeobacterium *Haloferax Mediterranei* Is Modulated by Salt. **Molecular and General Genetics MGG**, v. 222, n. 2, p. 225–232, 1 jul. 1990.

ENGSTROM, M. D.; PFLEGER, B. F. Transcription Control Engineering and Applications in Synthetic Biology. **Synthetic and Systems Biotechnology**, v. 2, n. 3, p. 176–191, 1 set. 2017.

FACCIOTTI, M. T. et al. Large scale physiological readjustment during growth enables rapid, comprehensive and inexpensive systems analysis. **BMC Systems Biology**, v. 4, p. 1–19, 2010.

FALB, M. et al. Metabolism of halophilic archaea. **Extremophiles**, v. 12, n. 2, p. 177–196, mar. 2008.

FARHADI, A. et al. Recombinantly Expressed Gas Vesicles as Nanoscale Contrast Agents for Ultrasound and Hyperpolarized MRI. **AIChE journal. American Institute of Chemical Engineers**, v. 64, n. 8, p. 2927–2933, ago. 2018.

FORTERRE, P.; GADELLE, D. Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms. **Nucleic Acids Research**, v. 37, n. 3, p. 679–692, fev. 2009.

FURTWÄNGLER, K. et al. Regulation of Phosphate Uptake via Pst Transporters in *Halobacterium Salinarum* R1. **Molecular Microbiology**, v. 76, n. 2, p. 378–392, 2010.

GANA, R. et al. Structural and Functional Studies of S-Adenosyl-L-Methionine Binding Proteins: A Ligand-Centric Approach. **BMC structural biology**, v. 13, p. 6, 25 abr. 2013.

GEST, H. The Discovery of Microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, Fellows of the Royal Society. **Notes and Records of the Royal Society of London**, v. 58, n. 2, p. 187–201, maio 2004.

GHAVAMI, S. B. et al. Alterations of the human gut *Methanobrevibacter smithii* as a biomarker for inflammatory bowel diseases. **Microbial Pathogenesis**, v. 117, p. 285–289, abr. 2018.

GHAZALPOUR, A. et al. Comparative Analysis of Proteome and Transcriptome Variation in Mouse. **PLoS Genetics**, v. 7, n. 6, 9 jun. 2011. Disponível em:

<<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3111477/>>. Acesso em: 23 nov. 2020.

GLAS, A. F. et al. The Archaeal Cofactor F₀ Is a Light-Harvesting Antenna Chromophore in Eukaryotes. **Proceedings of the National Academy of Sciences**, v. 106, n. 28, p. 11540–11545, 14 jul. 2009.

GOTTESMAN, S. Post-Transcriptional Regulation and the Bacterial Response to Stress. **The FASEB Journal**, v. 31, n. S1, p. 22.1–22.1, 2017.

GRIBALDO, S. et al. **The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse?** Nature Publishing Group. [s.l.: s.n.]. Disponível em: <www.nature.com/reviews/micro>.

GRINE, G. et al. Tobacco-Smoking-Related Prevalence of Methanogens in the Oral Fluid Microbiota. **Scientific Reports**, v. 8, n. 1, p. 9197, 15 jun. 2018.

GRINE, G. et al. Co-Culture of Methanobrevibacter Smithii with Enterobacteria during Urinary Infection. **EBioMedicine**, v. 43, p. 333–337, maio 2019a.

GRINE, G. et al. Detection of Methanobrevibacter Smithii in Vaginal Samples Collected from Women Diagnosed with Bacterial Vaginosis. **European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology**, v. 38, n. 9, p. 1643–1649, set. 2019b.

GRINE, G.; BOUALAM, M. A.; DRANCOURT, M. Methanobrevibacter Smithii, a Methanogen Consistently Colonising the Newborn Stomach. **European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology**, v. 36, n. 12, p. 2449–2455, dez. 2017.

GUTIERREZ, J. M. **Regulação pós-transcricional em *Halobacterium salinarum* NRC-1: caracterização do gene VNG_RS05855**. 2019. Universidade de São Paulo, 2019. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/95/95131/tde-03122019-115209/>>. Acesso em: 11 jan. 2021.

HACKLEY, R. K.; SCHMID, A. K. Global Transcriptional Programs in Archaea Share Features with the Eukaryotic Environmental Stress Response. **Journal of Molecular Biology**, RNA polymerase reaches 60. v. 431, n. 20, p. 4147–4166, 20 set. 2019.

HAIDER, S.; PAL, R. Integrated Analysis of Transcriptomic and Proteomic Data. **Current Genomics**, v. 14, n. 2, p. 91–110, abr. 2013.

HARTMANN, R.; SICKINGER, H.-D.; OESTERHELT, D. **Anaerobic growth of halobacteria (halophilism/photosynthesis/bacteriorhodopsin/substrate level phosphorylation)** **Biochemistry**. [s.l.: s.n.].

HASSON, K. C.; GAI, F.; ANFINRUD, P. A. The Photoisomerization of Retinal in Bacteriorhodopsin: Experimental Evidence for a Three-State Model. **Proceedings of the**

National Academy of Sciences, v. 93, n. 26, p. 15124–15129, 24 dez. 1996.

HERRMANN, K. The Shikimate Pathway: Early Steps in the Biosynthesis of Aromatic Compounds. **The Plant Cell**, v. 7, n. 7, p. 907–919, jul. 1995.

HOLMQVIST, E.; VOGEL, J. RNA-Binding Proteins in Bacteria. **Nature Reviews Microbiology**, v. 16, n. 10, p. 601–615, out. 2018.

HUANG, Q. et al. Phosphorylation of the Archaeal Holliday Junction Resolvase Hjc Inhibits Its Catalytic Activity and Facilitates DNA Repair in *Sulfolobus Islandicus* REY15A. **Frontiers in Microbiology**, v. 10, 2019. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmicb.2019.01214/full>>. Acesso em: 17 abr. 2021.

HUG, L. A. et al. A new view of the tree of life. **Nature Microbiology**, v. 1, n. 5, p. 16048, maio 2016.

IM, S. et al. Transcriptome analysis of salt-stressed *Deinococcus radiodurans* and characterization of salt-sensitive mutants. **Research in Microbiology**, v. 164, n. 9, p. 923–932, 2013.

Introduction to the Archaea. Disponível em: <<https://ucmp.berkeley.edu/archaea/archaea.html>>. Acesso em: 30 jul. 2020.

JAEKEN, J.; CARCHON, H. Glycosylation: General Overview. In: SQUIRE, L. R. (Ed.). **Encyclopedia of Neuroscience**. Oxford: Academic Press, 2009. p. 935–942.

JOO, W.-A.; KIM, C.-W. Proteomics of Halophilic archaea. **Journal of Chromatography B**, v. 815, n. 1–2, p. 237–250, fev. 2005.

KADI, N.; CHALLIS, G. L. Chapter 17. Siderophore Biosynthesis a Substrate Specificity Assay for Nonribosomal Peptide Synthetase-Independent Siderophore Synthetases Involving Trapping of Acyl-Adenylate Intermediates with Hydroxylamine. **Methods in Enzymology**, v. 458, p. 431–457, 2009.

KAHLES, A.; BEHR, J.; RÄTSCH, G. MMR: a tool for read multi-mapper resolution. **Bioinformatics**, v. 32, n. 5, p. 770–772, 1 mar. 2016.

KANEHISA, M.; SATO, Y. KEGG Mapper for Inferring Cellular Functions from Protein Sequences. **Protein Science**, v. 29, n. 1, p. 28–35, 2020.

KAPPLER, U.; SCHÄFER, H. Transformations of Dimethylsulfide. In: KRONECK, P. M. H.; TORRES, M. E. S. (Ed.). **The Metal-Driven Biogeochemistry of Gaseous Compounds in the Environment**. Metal Ions in Life Sciences. Dordrecht: Springer Netherlands, 2014. p. 279–313.

KAUR, A. et al. Survival strategies of an archaeal organism to withstand stress from transition metals. **Genome Research**, p. 841–854, 2006.

KENNEDY, S. P. et al. Understanding the adaptation of *Halobacterium* species NRC-1 to its

extreme environment through computational analysis of its genome sequence. **Genome Research**, v. 11, n. 10, p. 1641–1650, out. 2001.

KIM, J. Y. et al. The human gut archaeome: identification of diverse haloarchaea in Korean subjects. **Microbiome**, v. 8, n. 1, p. 114, 4 ago. 2020.

KINOSITA, Y.; NISHIZAKA, T. Cross-kymography analysis to simultaneously quantify the function and morphology of the archaeum. **Biophysics and Physicobiology**, v. 15, p. 121–128, 12 maio 2018.

KISH, A. et al. Salt shield: intracellular salts provide cellular protection against ionizing radiation in the halophilic archaeon, *Halobacterium salinarum* NRC-1. **Environmental Microbiology**, v. 11, n. 5, p. 1066–1078, maio 2009.

KITADAI, N.; KAMEYA, M.; FUJISHIMA, K. Origin of the Reductive Tricarboxylic Acid (rTCA) Cycle-Type CO₂ Fixation: A Perspective. **Life**, v. 7, n. 4, 23 out. 2017. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5745552/>>. Acesso em: 31 mar. 2021.

KOGA, Y. Thermal adaptation of the archaeal and bacterial lipid membranes. **Archaea (Vancouver, B.C.)**, v. 2012, p. 789652, 2012.

KOHLMEIER, M. Chapter 10 - Water-Soluble Vitamins and Nonnutrients. In: KOHLMEIER, M. (Ed.). **Nutrient Metabolism (Second Edition)**. San Diego: Academic Press, 2015. p. 567–671.

KOROBENIKOVA, A. V.; GARBER, M. B.; GONGADZE, G. M. Ribosomal Proteins: Structure, Function, and Evolution. **Biochemistry. Biokhimiia**, v. 77, n. 6, p. 562–574, jun. 2012.

KUNKA, K. S. et al. Acid Experimental Evolution of the Haloarchaeon *Halobacterium* Sp. NRC-1 Selects Mutations Affecting Arginine Transport and Catabolism. **Frontiers in Microbiology**, v. 11, p. 535, 2020.

LAKSHMANAN, A. et al. Preparation of biogenic gas vesicle nanostructures for use as contrast agents for ultrasound and MRI. **Nature Protocols**, v. 12, n. 10, p. 2050–2080, out. 2017.

LANDIS, G. A. Martian Water: Are There Extant Halobacteria on Mars? **Astrobiology**, v. 1, n. 2, p. 161–164, jun. 2001.

LASA, I. et al. Genome-Wide Antisense Transcription Drives mRNA Processing in Bacteria. **Proceedings of the National Academy of Sciences of the United States of America**, v. 108, n. 50, p. 20172–20177, 13 dez. 2011.

LEUKO, S. et al. Global Protein-Level Responses of *Halobacterium salinarum* NRC-1 to Prolonged Changes in External Sodium Chloride Concentrations. **Journal of Proteome Research**, v. 8, n. 5, p. 2218–2225, 1 maio 2009.

LEUKO, S. et al. The Survival and Resistance of *Halobacterium salinarum* NRC-1, *Halococcus hamelinensis*, and *Halococcus morrhuae* to Simulated Outer Space Solar Radiation. **Astrobiology**, v. 15, n. 11, p. 987–997, 1 nov. 2015.

LI, H. et al. The Sequence Alignment/Map Format and SAMtools. **Bioinformatics (Oxford, England)**, v. 25, n. 16, p. 2078–2079, 15 ago. 2009.

LIU, Y. et al. Biosynthesis of 4-Thiouridine in tRNA in the Methanogenic Archaeon *Methanococcus Maripaludis**. **Journal of Biological Chemistry**, v. 287, n. 44, p. 36683–36692, 26 out. 2012.

LÓPEZ GARCÍA DE LOMANA, A. et al. Selective Translation of Low Abundance and Upregulated Transcripts in *Halobacterium Salinarum*. **mSystems**, v. 5, n. 4, p. e00329-20, /msystems/5/4/msys.00329-20.atom, 28 jul. 2020.

LOSENSKY, G. et al. Novel Pili-like Surface Structures of *Halobacterium Salinarum* Strain R1 Are Crucial for Surface Adhesion. **Frontiers in Microbiology**, v. 5, 2015. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmicb.2014.00755/full>>. Acesso em: 13 ago. 2020.

LOWE, R. et al. Transcriptomics technologies. 2017. Disponível em: <<https://doi.org/10.1371/journal.pcbi.1005457.t001>>.

LUNDGREN, M.; BERNANDER, R. Genome-Wide Transcription Map of an Archaeal Cell Cycle. **Proceedings of the National Academy of Sciences**, v. 104, n. 8, p. 2939–2944, 20 fev. 2007.

MACEK, B. et al. Protein Post-Translational Modifications in Bacteria. **Nature Reviews Microbiology**, v. 17, n. 11, p. 651–664, nov. 2019.

MATARREDONA, L. et al. The Role of Stress Proteins in Haloarchaea and Their Adaptive Response to Environmental Shifts. **Biomolecules**, v. 10, n. 10, p. 1390, out. 2020.

MATTAR, S.; ENGELHARD, M. Cytochrome *Ba3* from *Natronobacterium Pharaonis*--an Archaeal Four-Subunit Cytochrome-c-Type Oxidase. **European Journal of Biochemistry**, v. 250, n. 2, p. 332–341, 1 dez. 1997.

MEI, Y. et al. Effects of salinity on the cellular physiological responses of *Natrinema* sp. J7-2. **PLoS ONE**, v. 12, n. 9, 2017. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5604999/>>.

MI, H. et al. Protocol Update for Large-Scale Genome and Gene Function Analysis with the PANTHER Classification System (v.14.0). **Nature Protocols**, v. 14, n. 3, p. 703–721, mar. 2019.

MICHEL, B.; SINHA, A. K.; LEACH, D. R. F. Replication Fork Breakage and Restart in *Escherichia Coli*. **Microbiology and Molecular Biology Reviews**, v. 82, n. 3, 1 set. 2018. Disponível em: <<https://mmbr.asm.org/content/82/3/e00013-18>>. Acesso em: 17 abr. 2021.

MILLER, A.-F. Superoxide dismutases: ancient enzymes and new insights. **FEBS letters**, v. 586, n. 5, p. 585–595, 9 mar. 2012.

MIYAUCHI, K.; KIMURA, S.; SUZUKI, T. A Cyclic Form of N⁶-Threonylcarbamoyladenosine as a Widely Distributed tRNA Hypermodification. **Nature Chemical Biology**, v. 9, n. 2, p. 105–111, fev. 2013.

MOUNCEY, N. J.; CHOUDHARY, M.; KAPLAN, S. Characterization of genes encoding dimethyl sulfoxide reductase of *Rhodobacter sphaeroides* 2.4.1T: an essential metabolic gene function encoded on chromosome II. **Journal of Bacteriology**, v. 179, n. 24, p. 7617–7624, dez. 1997.

MÜLLER, J. A.; DASSARMA, S. Genomic Analysis of Anaerobic Respiration in the Archaeon *Halobacterium* sp. Strain NRC-1: Dimethyl Sulfoxide and Trimethylamine N-Oxide as Terminal Electron Acceptors †. **JOURNAL OF BACTERIOLOGY**, v. 187, n. 5, p. 1659–1667, 2005.

MURPHY, J. E.; TIBBITTS, T. T.; KANTROWITZ, E. R. Mutations at Positions 153 and 328 In *Escherichia Coli* Alkaline Phosphatase Provide Insight Towards the Structure and Function of Mammalian and Yeast Alkaline Phosphatases. **Journal of Molecular Biology**, v. 253, n. 4, p. 604–617, 3 nov. 1995.

NELSON, D. L. (David L., 1942-. **Lehninger principles of biochemistry**. [s.l.] Sixth Edition. New York : W.H. Freeman, 2013., 2013.

NEUMANN, P. et al. Crystal structure of a 4-thiouridine synthetase–RNA complex reveals specificity of tRNA U8 modification. **Nucleic Acids Research**, v. 42, n. 10, p. 6673–6685, 2 jun. 2014.

NG, W. V. et al. Genome sequence of *Halobacterium* species NRC-1. **Proceedings of the National Academy of Sciences**, v. 97, n. 22, p. 12176–12181, 2000.

NOJIMA, Y. et al. Superoxide Dismutases, SOD1 and SOD2, Play a Distinct Role in the Fat Body during Pupation in Silkworm *Bombyx mori*. **PLoS ONE**, v. 10, n. 2, 25 fev. 2015. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4340916/>>. Acesso em: 24 abr. 2021.

OFFNER, S. et al. Structural characteristics of halobacterial gas vesicles. n. 144, p. 1331–1342, 1998.

ORELLANA, M. V. et al. A Role for Programmed Cell Death in the Microbial Loop. **PLoS ONE**, v. 8, n. 5, 8 maio 2013. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3648572/>>. Acesso em: 2 set. 2020.

OREN, A. The Function of Gas Vesicles in Halophilic Archaea and Bacteria: Theories and Experimental Evidence. **Life : Open Access Journal**, v. 3, n. 1, p. 1–20, 27 dez. 2012.

OREN, A.; VENTOSA, A. *Halobacterium*. In: **Bergey's Manual of Systematics of Archaea and Bacteria**. [s.l.] American Cancer Society, 2017. p. 1–12.

OUELLETTE, M.; MAKAY, A. M.; PAPKE, R. T. Dihydroxyacetone metabolism in *Haloferax volcanii*. **Frontiers in Microbiology**, v. 4, 16 dez. 2013. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3863723/>>. Acesso em: 25 abr. 2021.

OUELLETTE, R. J.; RAWN, J. D. 31 - Lipids and Biological Membranes. In: OUELLETTE, R. J.; RAWN, J. D. (Ed.). **Organic Chemistry (Second Edition)**. [s.l.] Academic Press, 2018. p.

1001–1032.

PECK, R. F. et al. Opsin-Mediated Inhibition of Bacterioruberin Synthesis in Halophilic Archaea. **Journal of Bacteriology**, v. 199, n. 21, 1 nov. 2017.

PERCHE-LETUVÉE, P. et al. 4-Demethylwyosine Synthase from *Pyrococcus abyssi* Is a Radical-S-adenosyl-L-methionine Enzyme with an Additional [4Fe-4S]₂ Cluster That Interacts with the Pyruvate Co-substrate. **The Journal of Biological Chemistry**, v. 287, n. 49, p. 41174–41185, 30 nov. 2012.

PERTEA, M. et al. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. 2016. Disponível em: <<http://ccb.jhu.edu/software.shtml>>.

PFEIFER, F. et al. Haloarchaea and the formation of gas vesicles. **Journal of Agricultural and Food Chemistry**, v. 57, n. 1, p. 385–402, mar. 2009.

PFEIFER, F. Distribution, formation and regulation of gas vesicles. 2012. Disponível em: <www.nature.com/reviews/micro>.

PFEIFER, F. Haloarchaea and the formation of gas vesicles. **Life**, v. 5, n. 1, p. 385–402, 2015.

PFEIFER, F.; WEIDINGER, G.; GOEBEL, W. Genetic variability in *Halobacterium halobium*. **Journal of bacteriology**, v. 145, p. 375–81, 1 fev. 1981.

PFEIFFER, F. et al. The Genome Sequence of the *Halobacterium salinarum* Type Strain Is Closely Related to That of Laboratory Strains NRC-1 and R1. 2019. Disponível em: <<https://doi.org/10.1093/bioinformatics/>>.

PIAO, H. L. et al. An Arabidopsis GSK3/Shaggy-Like Gene That Complements Yeast Salt Stress-Sensitive Mutants Is Induced by NaCl and Absciscic Acid. **Plant Physiology**, v. 119, n. 4, p. 1527–1534, 1 abr. 1999.

PICARD, F. et al. Examination of Post-Transcriptional Regulations in Prokaryotes by Integrative Biology. **Comptes Rendus Biologies**, v. 332, n. 11, p. 958–973, nov. 2009.

PILLA, S. P.; BAHADUR, R. P. Residue Conservation Elucidates the Evolution of R-Proteins in Ribosomal Assembly and Function. **International Journal of Biological Macromolecules**, v. 140, p. 323–329, 1 nov. 2019.

POHLSCHRODER, M. et al. The universal tree of life: an update. 2015. Disponível em: <www.frontiersin.org>.

PUEHRINGER, S.; METLITZKY, M.; SCHWARZENBACHER, R. The pyrroloquinoline quinone biosynthesis pathway revisited: A structural approach. **BMC Biochemistry**, v. 9, p. 8, 27 mar. 2008.

QUINLAN, A. R.; HALL, I. M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. **Bioinformatics**, v. 26, n. 6, p. 841–842, 15 mar. 2010.

RAMÍREZ, F. et al. DeepTools2: A next Generation Web Server for Deep-Sequencing Data

Analysis. **Nucleic Acids Research**, v. 44, n. W1, p. W160-165, 08 2016.

RANDAU, L. et al. A Cytidine Deaminase Edits C to U in Transfer RNAs in Archaea. **Science**, v. 324, n. 5927, p. 657–659, 1 maio 2009.

RAO, N. N.; TORRIANI, A. Molecular Aspects of Phosphate Transport in Escherichia Coli. **Molecular Microbiology**, v. 4, n. 7, p. 1083–1090, 1990.

ROBB, F. T.; DASSARMA, S.; FLEISCHMANN, E. M. **Archaea: a laboratory manual - halophiles**. [s.l.] Cold Spring Harbor Laboratory, 1995.

ROSA, L. T. et al. Tripartite ATP-Independent Periplasmic (TRAP) Transporters and Tripartite Tricarboxylate Transporters (TTT): From Uptake to Pathogenicity. **Frontiers in Cellular and Infection Microbiology**, v. 8, 12 fev. 2018. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5812351/>>. Acesso em: 19 abr. 2021.

ROWLEY, C. A.; KENDALL, M. M. To B12 or Not to B12: Five Questions on the Role of Cobalamin in Host-Microbial Interactions. **PLOS Pathogens**, v. 15, n. 1, p. e1007479, 3 jan. 2019.

SABERI, F. et al. Natural antisense RNAs as mRNA regulatory elements in bacteria: a review on function and applications. **Cellular & Molecular Biology Letters**, v. 21, n. 1, p. 6, 28 jul. 2016.

SASAKI, J.; SPUDICH, J. L. Signal Transfer in Haloarchaeal Sensory Rhodopsin– Transducer Complexes†. **Photochemistry and Photobiology**, v. 84, n. 4, p. 863–868, 2008.

SCHIRALDI, C.; GIULIANO, M.; DE ROSA, M. **Perspectives on biotechnological applications of archaea**. [s.l.: s.n.]

SCHLOSS, P. D. et al. Status of the Archaeal and Bacterial Census: An Update. **mBio**, v. 7, n. 3, 6 jul. 2016. Disponível em: <<https://mbio.asm.org/content/7/3/e00201-16>>. Acesso em: 19 mar. 2021.

SHOUMSKAYA, M. A. et al. Identical Hik-Rre Systems Are Involved in Perception and Transduction of Salt Signals and Hyperosmotic Signals but Regulate the Expression of Individual Genes to Different Extents in Synechocystis. **The Journal of Biological Chemistry**, v. 280, n. 22, p. 21531–21538, 3 jun. 2005.

SILVA, L. G. et al. OxyR and the Hydrogen Peroxide Stress Response in Caulobacter Crescentus. **Gene**, v. 700, p. 70–84, 5 jun. 2019.

SILVA-JUNIOR, A. C. da; BELARMINO, L. C.; BENKO-ISEPPON, A. M. DNA microarray plataforma: prospecção tecnológica e demanda científica. **REVISTA GEINTEC-GESTAO INOVACAO E TECNOLOGIAS**, v. 8, n. 2, p. 4369–4380, 19 jun. 2018.

SINGH, A.; SINGH, A. K. Haloarchaea: worth exploring for their biotechnological potential. **Biotechnology Letters**, v. 39, n. 12, p. 1793–1800, 2017.

SONG, Y.; GUNNER, M. R. Halorhodopsin Pumps Cl⁻ and Bacteriorhodopsin Pumps Protons by a Common Mechanism That Uses Conserved Electrostatic Interactions. **Proceedings of the National Academy of Sciences**, v. 111, n. 46, p. 16377–16382, 18 nov. 2014.

SOPPA, J. From genomes to function: Haloarchaea as model organisms. **Microbiology**, v. 152, n. 3, p. 585–590, 2006.

SOUZA, W. de; CARVALHO, B. de S.; LOPES-CENDES, I. Rqc: A Bioconductor Package for Quality Control of High-Throughput Sequencing Data. **Journal of Statistical Software**, v. 87, n. 1, p. 1–14, 12 nov. 2018.

SPANG, A.; CACERES, E. F.; ETTEMA, T. J. G. Genomic Exploration of the Diversity, Ecology, and Evolution of the Archaeal Domain of Life. **Science**, v. 357, n. 6351, 11 ago. 2017. Disponível em: <<https://science.sciencemag.org/content/357/6351/eaaf3883>>. Acesso em: 1 jul. 2020.

STAV, S. et al. Genome-wide discovery of structured noncoding RNAs in bacteria. **BMC Microbiology**, v. 19, n. 1, p. 66, 22 mar. 2019.

SWATI*, A. G. and D. **Riboswitches in Archaea**. 2. Disponível em: <<http://www.eurekaselect.com/171843/article>>. Acesso em: 23 jun. 2020.

TASHIRO, Y. et al. Molecular Genetic and Physical Analysis of Gas Vesicles in Buoyant Enterobacteria. **Environmental Microbiology**, v. 18, n. 4, p. 1264–1276, 2016.

TEN-CATEN, F. et al. Internal RNAs overlapping coding sequences can drive the production of alternative proteins in archaea. **RNA Biology**, v. 15, n. 8, p. 1119–1132, 2018.

TEUFEL, K.; PFEIFER, F. Interaction of Transcription Activator GvpE with TATA-Box-Binding Proteins of *Halobacterium Salinarum*. **Archives of Microbiology**, v. 192, n. 2, p. 143–149, 1 fev. 2010.

THOMBRE, R. S. et al. Biology and Survival of Extremely Halophilic Archaeon Haloarcula Marismortui RR12 Isolated from Mumbai Salterns, India in Response to Salinity Stress. **Scientific Reports**, v. 6, n. 1, p. 25642, 27 maio 2016.

TJONG, E.; DIMRI, M.; MOHIUDDIN, S. S. Biochemistry, Tetrahydrofolate. In: **StatPearls**. Treasure Island (FL): StatPearls Publishing, 2021.

VAN ASSCHE, E. et al. RNA-Binding Proteins Involved in Post-Transcriptional Regulation in Bacteria. **Frontiers in Microbiology**, v. 6, 2015. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmicb.2015.00141/full>>. Acesso em: 4 maio. 2021.

VAUCLARE, P. et al. Molecular Adaptation and Salt Stress Response of *Halobacterium Salinarum* Cells Revealed by Neutron Spectroscopy. **Extremophiles: Life Under Extreme Conditions**, v. 19, n. 6, p. 1099–1107, nov. 2015.

VAUCLARE, P. et al. Surviving Salt Fluctuations: Stress and Recovery in *Halobacterium*

Salinarum , an Extreme Halophilic Archaeon. **Scientific Reports**, v. 10, n. 1, p. 3298, 24 fev. 2020.

VREELAND, R. H. et al. Isolation of Live Cretaceous (121–112 Million Years Old) Halophilic Archaea from Primary Salt Crystals. **Geomicrobiology Journal**, v. 24, n. 3–4, p. 275–282, 23 jul. 2007.

WAAS, W. F.; DE CRÉCY-LAGARD, V.; SCHIMMEL, P. Discovery of a Gene Family Critical to Wyosine Base Formation in a Subset of Phenylalanine-Specific Transfer RNAs. **The Journal of Biological Chemistry**, v. 280, n. 45, p. 37616–37622, 11 nov. 2005.

WALSBY, A. E. Gas vesicles. **Microbiological Reviews**, v. 58, n. 1, p. 94–144, mar. 1994.

WANG, G. et al. Arsenic Resistance in Halobacterium sp. Strain NRC-1 Examined by Using an Improved Gene Knockout System. **Journal of bacteriology**, v. 186, p. 3187–94, 1 jun. 2004.

WATSON, T. The Trickster Microbes That Are Shaking up the Tree of Life. **Nature**, v. 569, n. 7756, p. 322–324, 14 maio 2019.

WHITE, M. F.; ALLERS, T. DNA repair in the archaea—an emerging picture. **FEMS Microbiology Reviews**, v. 42, n. 4, p. 514–526, 1 jul. 2018.

WILKENS, S. Structure and mechanism of ABC transporters. **F1000Prime Reports**, v. 7, 3 fev. 2015. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4338842/>>. Acesso em: 5 nov. 2020.

WILLIAMS, T. J. et al. Glycerol metabolism of haloarchaea. **Environmental Microbiology**, v. 19, n. 3, p. 864–877, 2017a.

WILLIAMS, T. J. et al. **Glycerol metabolism of haloarchaea**. [s.l.] Blackwell Publishing Ltd, 2017b. v. 19

WINTJENS, R.; ROOMAN, M. Structural Classification of HTH DNA-Binding Domains and Protein – DNA Interaction Modes. **Journal of Molecular Biology**, v. 262, n. 2, p. 294–313, 20 set. 1996.

WIRTH, J.; YOUNG, M. The Intriguing World of Archaeal Viruses. **PLOS Pathogens**, v. 16, n. 8, p. e1008574, 13 ago. 2020.

WOESE, C. R.; FOX, G. E. Phylogenetic Structure of the Prokaryotic Domain: The Primary Kingdoms. **Proceedings of the National Academy of Sciences**, v. 74, n. 11, p. 5088–5090, 1 nov. 1977.

WOESE, C. R.; KANDLER, O.; WHEELIS, M. L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. **Proceedings of the National Academy of Sciences**, v. 87, n. 12, p. 4576–4579, 1990.

WRIGHT, M. S. et al. Assessment of Insertion Sequence Mobilization as an Adaptive Response to Oxidative Stress in *Acinetobacter baumannii* Using IS-seq. **Journal of Bacteriology**, v. 199,

n. 9, 11 abr. 2017. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5388817/>>. Acesso em: 10 fev. 2021.

WURM, J. P. et al. Identification of the enzyme responsible for N1-methylation of pseudouridine 54 in archaeal tRNAs. **RNA**, v. 18, n. 3, p. 412–420, mar. 2012.

XU, J.; ZHANG, J.; ZHANG, W. Antisense RNA: the new favorite in genetic research. **Journal of Zhejiang University. Science. B**, v. 19, n. 10, p. 739–749, out. 2018.

YAO, A. I.; FACCIOTTI, M. T. Regulatory Multidimensionality of Gas Vesicle Biogenesis in *Halobacterium salinarum* NRC-1. **Archaea**, v. 2011, 24 out. 2011. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3202098/>>. Acesso em: 11 ago. 2020.

YOUNG, A. P.; BANDARIAN, V. TYW1: A radical SAM enzyme involved in the biosynthesis of wybutosine bases. **Methods in enzymology**, v. 606, p. 119–153, 2018.

YOUNUS, H. Therapeutic potentials of superoxide dismutase. **International Journal of Health Sciences**, v. 12, n. 3, p. 88–93, 2018.

ZARETSKY, M. et al. N-Glycosylation Is Important for *Halobacterium salinarum* Archaeellin Expression, Archaeellum Assembly and Cell Motility. **Frontiers in Microbiology**, v. 10, n. JUN, p. 1367, jun. 2019.

ZEILSTRA-RYALLS, J. H.; KAPLAN, S. Role of the fnrL Gene in Photosystem Gene Expression and Photosynthetic Growth of *Rhodobacter sphaeroides* 2.4.1. **Journal of Bacteriology**, v. 180, n. 6, p. 1496–1503, mar. 1998.

8. Anexos

Tabela S1: Tabela de genes menos expressos em hiposalinidade. Contém o locus, nome do gene, produto, valor de M que representa o $\log_2$ (L (*Low*) / O (*optimal salinity*)) e o desvio padrão dos valores de M das replicatas.

<i>Locus</i>	Nome do gene	Produto	M	sdM
VNG_RS03880	<i>acs</i>	acetate--CoA ligase	-5.5	0.5
VNG_RS03255	VNG_RS03255	4Fe-4S dicluster domain -containing protein	-5.1	1.0
VNG_RS03260	<i>nrfD</i>	polysulfide reductase NrfD	-5.0	0.0
VNG_RS04635	VNG_RS04635	radical SAM protein	-5.0	1.0
VNG_RS03265	VNG_RS03265	molecular chaperone	-4.9	1.0
VNG_RS03870	VNG_RS03870	hypothetical protein	-4.7	0.1
VNG_RS05380	VNG_RS05380	hypothetical protein	-4.7	0.1
VNG_RS05715	VNG_RS05715	bacteriorhodopsin	-4.6	0.9
VNG_RS08545	VNG_RS08545	hypothetical protein	-4.1	0.0
VNG_RS08550	VNG_RS08550	b(o/a)3-type cytochrome c oxidase subunit 1	-4.1	0.1
VNG_RS10625	<i>gvpA</i>	gas vesicle structural protein GvpA	-4.1	0.4
VNG_RS12320	<i>gvpA</i>	gas vesicle structural protein GvpA	-4.1	0.4
VNG_RS03250	VNG_RS03250	molybdopterin-dependent oxidoreductase	-3.9	0.3
VNG_RS09205	VNG_RS09205	twin-arginine translocation signal domain -containing protein	-3.9	1.1
VNG_RS07650	VNG_RS07650	hypothetical protein	-3.8	1.6

VNG_RS03875	VNG_RS03875	helix-turn-helix domain -containing protein	-3.7	0.0
VNG_RS05675	VNG_RS05675	phytoene/squalene synthase family protein	-3.7	0.0
VNG_RS03245	VNG_RS03245	hypothetical protein	-3.7	2.0
VNG_RS02585	VNG_RS02585	GNAT family N-acetyl transferase	-3.7	0.2
VNG_RS14250	VNG_RS14250	hypothetical protein	-3.6	1.7
VNG_RS12740	VNG_RS12740	hypothetical protein	-3.6	0.1
VNG_RS00745	VNG_RS00745	bacteriorhodopsin	-3.5	0.4
VNG_1459H	VNG_1459H	NA	-3.4	0.3
VNG_RS06090	VNG_RS06090	magnesium chelatase subunit D family protein	-3.4	1.4
VNG_RS09245	VNG_RS09245	ABC transporter ATP-binding protein	-3.4	0.3
VNG_RS02580	VNG_RS02580	methyl malonyl -CoA mutase family protein	-3.4	2.0
VNG_RS08280	VNG_RS08280	beta-carotene 15,15'-dioxygenase,Brp/Blh family	-3.4	2.4
VNG_RS03240	VNG_RS03240	helix-turn-helix domain -containing protein	-3.4	0.4
VNG_RS14385	VNG_RS14385	halocyanin domain -containing protein	-3.4	0.5
VNG_RS11425	VNG_RS11425	SHOCT domain -containing protein	-3.4	0.1
VNG_RS08555	VNG_RS08555	cytochrome c oxidase subunit II	-3.3	0.1
VNG_RS09240	VNG_RS09240	ABC transporter permease	-3.2	0.4
VNG_RS04345	VNG_RS04345	SAM-dependent methyl transferase	-3.1	0.3
VNG_RS09235	VNG_RS09235	ABC transporter permease	-3.1	0.3

VNG_RS09435	VNG_RS09435	hypothetical protein	-3.1	0.1
VNG_RS09440	VNG_RS09440	DoxX family protein	-3.1	0.4
VNG_RS01395	VNG_RS01395	HAMP domain -containing protein	-3.0	0.0
VNG_RS01780	VNG_RS01780	phosphateuptakeregulatorPhoU	-3.0	0.1
VNG_RS09230	VNG_RS09230	twin-arginine translocation signal domain -containing protein	-3.0	0.3
VNG_RS09250	VNG_RS09250	ATP-binding cassette domain -containing protein	-3.0	0.6
VNG_RS03935	VNG_RS03935	flagellin	-3.0	0.1
VNG_RS03940	VNG_RS03940	flagellin	-3.0	0.1
VNG_RS10635	<i>gvpN</i>	gas vesicle protein GvpN	-3.0	0.9
VNG_RS12330	<i>gvpN</i>	gas vesicle protein GvpN	-3.0	0.9
VNG_RS06345	VNG_RS06345	halocyanin	-2.9	2.0
VNG_RS02465	VNG_RS02465	acyl-CoA carboxylase subunit beta	-2.9	0.5
VNG_RS01765	<i>pstA</i>	phosphateABC transporter permeasePstA	-2.8	0.1
VNG_RS05705	VNG_RS05705	beta-carotene 15,15'-dioxygenase,Brp/Blh family	-2.8	0.7
VNG_RS12750	VNG_RS12750	hypothetical protein	-2.8	0.8
VNG_RS01735	VNG_RS01735	HADhydrolase family protein	-2.7	0.7
VNG_RS04060	VNG_RS04060	hypothetical protein	-2.7	0.3
VNG_RS06115	VNG_RS06115	nicotinate-nucleotide--dimethyl benzimidazole phosphoribosyl transferase	-2.7	0.8

VNG_RS11600	VNG_RS11600	dicarboxylate/aminoacid:cation symporter	-2.7	2.3
VNG_RS01700	VNG_RS01700	M20 family metallo-hydrolase	-2.7	0.4
VNG_RS03855	VNG_RS03855	hypothetical protein	-2.7	0.2
VNG_RS03860	VNG_RS03860	hypothetical protein	-2.7	0.5
VNG_RS06100	VNG_RS06100	precorrin-8Xmethyl mutase	-2.7	0.5
VNG_RS06105	VNG_RS06105	cobalt-precorrin-7(C(5))-methyl transferase	-2.7	0.2
VNG_RS13370	VNG_RS13370	ATP-binding protein	-2.7	0.2
VNG_RS10630	<i>gvpC</i>	gas vesicle protein GvpC	-2.7	0.9
VNG_RS01775	VNG_RS01775	PstS family phosphateABC transporter substrate-binding protein	-2.6	0.1
VNG_RS04665	VNG_RS04665	HAMP domain -containing protein	-2.6	0.6
VNG_RS00900	VNG_RS00900	hypothetical protein	-2.6	1.0
VNG_RS06095	<i>cobN</i>	cobalto chelatase subunit CobN	-2.6	0.4
VNG_RS06125	VNG_RS06125	cob(I)yrinicacida,c-diamide adenosyltransferase	-2.6	1.0
VNG_RS12325	<i>gvpC</i>	gas vesicle protein GvpC	-2.6	1.0
VNG_RS01760	VNG_RS01760	phosphateABC transporter ATP-binding protein	-2.6	0.6
VNG_RS05700	VNG_RS05700	helix-turn-helix domain -containing protein	-2.6	0.4
VNG_RS07095	VNG_RS07095	hypothetical protein	-2.6	0.9
VNG_6275a	VNG_6275a	NA	-2.6	0.1
VNG_RS07585	VNG_RS07585	surface glycoprotein	-2.5	0.4

VNG_RS08170	VNG_RS08170	hypothetical protein	-2.5	0.1
VNG_RS11255	VNG_RS11255	Hsp20/alpha crystallin family protein	-2.5	0.7
VNG_RS05140	VNG_RS05140	hypothetical protein	-2.5	0.1
VNG_RS05180	VNG_RS05180	cupin domain -containing protein	-2.5	0.2
VNG_RS01050	VNG_RS01050	helix-turn-helix transcriptional regulator	-2.4	0.6
VNG_RS02485	VNG_RS02485	Glu/Leu/Phe/Val dehydrogenase	-2.4	0.6
VNG_RS06120	VNG_RS06120	cobyrinic acid, c-diamide synthase	-2.4	2.0
VNG_RS01580	VNG_RS01580	NAD(P)-dependent glycerol-1-phosphate dehydrogenase	-2.4	1.2
VNG_RS01755	<i>phoU</i>	phosphate signaling complex protein PhoU	-2.4	0.4
VNG_2490a	VNG_2490a	NA	-2.4	1.1
VNG_RS06240	VNG_RS06240	TIGR00266 family protein	-2.4	0.1
VNG_RS01070	VNG_RS01070	cell division protein	-2.3	0.3
VNG_RS02360	VNG_RS02360	hypothetical protein	-2.3	0.8
VNG_RS02920	VNG_RS02920	universal stress protein	-2.3	0.6
VNG_RS03040	VNG_RS03040	acyl -CoA dehydrogenase family protein	-2.3	0.0
VNG_RS10970	VNG_RS10970	ParA family protein	-2.3	0.8
VNG_RS02720	VNG_RS02720	helix-turn-helix transcriptional regulator	-2.3	1.5
VNG_RS09195	VNG_RS09195	ABC transporter permease	-2.3	1.2
VNG_RS09200	VNG_RS09200	ABC transporter permease	-2.3	1.5

VNG_RS01770	<i>pstC</i>	phosphateABC transporter permeasesubunitPstC	-2.2	0.6
VNG_RS05620	VNG_RS05620	dodecin domain -containing protein	-2.2	0.3
VNG_RS09185	VNG_RS09185	ABC transporter ATP-binding protein	-2.2	1.1
VNG_RS13565	VNG_RS13565	PQQ-like beta-propeller repeat protein	-2.2	0.1
NC_002607.1:1987840..1989697:minus NA ISH34-IS200_IS605	ISH34-IS200_IS605	NA	-2.2	0.3
VNG_RS00800	VNG_RS00800	zinc ribbon domain -containing protein	-2.2	0.8
NC_002608.1:181743..183597:minus NA ISH12-IS200_IS605	ISH12-IS200_IS605	NA	-2.2	0.1
VNG_RS00895	VNG_RS00895	DUF4129 domain -containing protein	-2.2	1.2
VNG_RS11435	VNG_RS11435	hypothetical protein	-2.2	0.2
VNG_1740a	VNG_1740a	NA	-2.1	1.3
VNG_RS13360	VNG_RS13360	hypothetical protein	-2.1	0.1
VNG_RS10350	<i>sepF</i>	cell division protein SepF	-2.1	0.4
VNG_RS10640	<i>gvpO</i>	gas vesicle protein GvpO	-2.1	0.8
VNG_RS12335	<i>gvpO</i>	gas vesicle protein GvpO	-2.1	0.8
NC_002607.1:265582..266984:minus NA ISH8B-IS4_ssgr_ISH8	ISH8B-IS4_ssgr_ISH8	NA	-2.1	1.1

VNG_RS02635	VNG_RS02635	ABC transporter ATP-binding protein	-2.1	0.3
VNG_1781C	<i>uppS2</i>	NA	-2.1	0.3
VNG_RS11100	VNG_RS11100	hypothetical protein	-2.1	0.3
VNG_6149H	VNG_as13755_1 539	NA	-2.1	0.3
VNG_RS00930	VNG_RS00930	DUF192 domain -containing protein	-2.1	0.1
VNG_RS05955	VNG_RS05955	acetyl /propionyl /methyl crotonyl -CoAcarboxyl asesubunitalpha	-2.1	1.1
VNG_RS06765	VNG_RS06765	ZIP family metal transporter	-2.1	0.5
VNG_RS07370	VNG_RS07370	NAD(+)/NADH kinase	-2.1	0.5
VNG_RS08145	VNG_RS08145	citrate(Si)-synthase	-2.1	0.1
VNG_RS09190	VNG_RS09190	ABC transporter ATP-binding protein	-2.1	0.9
VNG_RS11620	<i>arcD</i>	arginine/ornithineantiporterArcD	-2.1	0.1
VNG_RS12020	VNG_RS12020	type IV pilin	-2.1	0.2
NC_002607.1:36 125..38023:plus NA ISH12-IS200_ IS605	ISH12-IS200_IS 605	NA	-2.1	0.4
VNG_RS06130	VNG_RS06130	cobyricacid synthase	-2.0	0.6
VNG_2202H	VNG_2202H	NA	-2.0	1.0
NC_002607.1:20 518..21934:minus NA ISH38-IS200	ISH38-IS200_IS 605_ssgr_IS134 1	NA	-2.0	0.3

_IS605_ssgr_IS1 341				
VNG_RS06800	<i>crtI</i>	phytoene desaturase	-2.0	0.2
VNG_RS05095	<i>purD</i>	phosphoribosyl amine--glycine ligase	-2.0	0.6
VNG_1438a	VNG_1438a	NA	-2.0	0.5
VNG_RS10505	<i>csg</i>	major cell surface glycoprotein	-2.0	0.5
NC_001869.1:16 3485..164887:mi nus NA ISH8B-IS 4_ssgr_ISH8	ISH8B-IS4_ssgr_ ISH8	NA	-2.0	0.8
VNG_RS01355	VNG_RS01355	hypothetical protein	-1.9	1.1
VNG_RS08285	VNG_RS08285	lycopene cyclase domain -containing protein	-1.9	0.4
VNG_RS05690	VNG_RS05690	AAA family ATPase	-1.9	0.0
VNG_RS03325	VNG_RS03325	hypothetical protein	-1.9	0.4
VNG_RS03740	VNG_RS03740	flagellin	-1.9	0.1
VNG_RS06440	VNG_RS06440	bacteriorhodopsin	-1.9	0.1
VNG_RS13260	VNG_RS13260	hypothetical protein	-1.8	0.6
VNG_RS03105	VNG_RS03105	methyl -acceptingchemotaxis protein	-1.8	0.3
VNG_RS03655	VNG_RS03655	ZIP family metal transporter	-1.8	0.6
VNG_RS05135	VNG_RS05135	hypothetical protein	-1.8	0.4
VNG_RS07690	VNG_RS07690	Na ⁺ /H ⁺ -antiporter NhaC family protein	-1.8	0.4
VNG_RS09555	VNG_RS09555	DNAstarvation/stationary phase protection protein	-1.8	1.0

NC_002607.1:1231046..1232448:plus NA ISH8E-IS4_ssgr_ISH8	ISH8E-IS4_ssgr_ISH8	NA	-1.8	0.6
VNG_RS00905	VNG_RS00905	DUF58 domain -containing protein	-1.8	0.8
VNG_RS02110	VNG_RS02110	TAXI family TRAP transporter solute-binding subunit	-1.8	0.8
VNG_RS03745	VNG_RS03745	flagellin	-1.8	0.1
VNG_RS03890	VNG_RS03890	hypothetical protein	-1.8	0.1
VNG_1170a	VNG_1170a	NA	-1.8	0.1
VNG_RS06435	VNG_RS06435	HAMP domain -containing protein	-1.8	0.1
VNG_RS00795	VNG_RS00795	typelltoxin-antitoxinsystemParD family antitoxin	-1.8	0.5
VNG_RS03415	VNG_RS03415	TraB/GumN family protein	-1.8	0.1
VNG_RS06330	<i>cbiQ</i>	cobaltECF transporter TcomponentCbiQ	-1.8	0.2
VNG_RS06420	VNG_RS06420	PGF-CTERMsorting domain -containing protein	-1.8	0.1
VNG_RS08440	VNG_RS08440	CBS domain -containing protein	-1.8	0.4
VNG_RS11660	VNG_RS11660	hypothetical protein	-1.8	0.2
VNG_RS12125	VNG_RS12125	cytochrome ubiquinol oxidase subunit il	-1.8	0.5
NC_002607.1:11454..12876:minus NA ISH37-IS200_IS605_ssgr_IS1341	ISH37-IS200_IS605_ssgr_IS1341	NA	-1.8	0.1

VNG_RS01385	VNG_RS01385	hypothetical protein	-1.7	0.6
VNG_RS01890	VNG_RS01890	AMP-binding protein	-1.7	0.6
VNG_RS03110	VNG_RS03110	BMP family protein	-1.7	0.1
VNG_RS10035	VNG_RS10035	caspase family protein	-1.7	0.7
VNG_6036H	VNG_6036H	NA	-1.7	0.1
VNG_RS00790	<i>ftsZ</i>	cell division protein FtsZ	-1.7	0.4
VNG_RS04660	VNG_RS04660	acyl -CoA dehydrogenase family protein	-1.7	0.3
VNG_RS09565	VNG_RS09565	helix-turn-helix transcriptional regulator	-1.7	0.4
VNG_RS10620	<i>gvpD</i>	gas vesicle protein GvpD	-1.7	0.8
VNG_RS02860	VNG_RS02860	2-dehydropantoate2-reductase	-1.7	0.8
NC_002608.1:79702..81104:plus NA ISH8E-IS4_ssgr_ISH8	ISH8E-IS4_ssgr_ISH8	NA	-1.7	0.8
VNG_RS02790	VNG_RS02790	thioredoxin family protein	-1.7	0.5
VNG_RS03735	VNG_RS03735	flagellin	-1.7	0.1
VNG_RS04095	<i>aglJ</i>	S-layer glycoprotein N-glycosyl transferaseAglJ	-1.7	0.5
VNG_RS04140	VNG_RS04140	hypothetical protein	-1.7	0.5
VNG_RS04900	VNG_RS04900	hypothetical protein	-1.7	0.4
VNG_RS06795	VNG_RS06795	deoxyribodipyrimidinephoto-lyase	-1.7	0.2
VNG_RS07205	VNG_RS07205	ABC transporter substrate-binding protein	-1.7	0.2

VNG_RS08960	VNG_RS08960	1-(5-phosphoribosyl)-5-[(5-phosphoribosyl amino)methyl ideneamino]imidazole-4-carboxamide isomerase	-1.7	0.5
VNG_RS09880	VNG_RS09880	universal stress protein	-1.7	0.2
VNG_RS10125	VNG_RS10125	SIMPL domain -containing protein	-1.7	0.1
VNG_RS11680	VNG_RS11680	hypothetical protein	-1.7	0.2
VNG_RS12290	<i>gvpJ</i>	gas vesicle protein GvpJ	-1.7	0.4
VNG_RS12315	<i>gvpD</i>	gas vesicle protein GvpD	-1.7	0.9
VNG_7030b	VNG_7030b	NA	-1.7	0.1
VNG_RS12430	VNG_RS12430	cytochrome ubiquinol oxidase subunit il	-1.7	0.5
VNG_RS13065	VNG_RS13065	cytochrome ubiquinol oxidase subunit il	-1.7	0.1
VNG_RS00760	VNG_RS00760	classE SAM-dependent methyl transferase	-1.6	0.3
VNG_RS02780	VNG_RS02780	hypothetical protein	-1.6	0.0
VNG_RS05970	VNG_RS05970	universal stress protein	-1.6	0.0
VNG_RS07530	VNG_RS07530	phosphoribosyl amino imidazolesuccinocarboxamide synthase	-1.6	0.1
VNG_RS08930	VNG_RS08930	methyl aspartate mutase subunit E	-1.6	0.0
VNG_RS10230	VNG_RS10230	MarR family transcriptional regulator	-1.6	0.7
VNG_RS10645	VNG_RS10645	ParA family protein	-1.6	0.7
VNG_RS13150	<i>ffs</i>		-1.6	0.6

NC_002608.1:337564..338966:minus NA ISH8B-IS4_ssgr_ISH8	ISH8B-IS4_ssgr_ISH8	NA	-1.6	0.7
VNG_RS00635	VNG_RS00635	NAD(P)/FAD-dependent oxidoreductase	-1.6	0.4
VNG_RS01275	VNG_RS01275	SPFH/Band7/PHB domain protein	-1.6	0.1
VNG_RS02195	VNG_RS02195	purine phosphoribosyl transferase family protein	-1.6	0.4
VNG_RS03560	VNG_RS03560	kynureninase	-1.6	0.6
VNG_RS03970	VNG_RS03970	glycosyl transferase	-1.6	0.8
VNG_RS05225	VNG_RS05225	N-acyl homoserine lactonase family protein	-1.6	0.6
VNG_RS06785	VNG_RS06785	PAS domain -containing protein	-1.6	0.1
VNG_RS00005	<i>mpB</i>		-1.6	0.1
NC_001869.1:79702..81104:plus NA ISH8E-IS4_ssgr_ISH8	ISH8E-IS4_ssgr_ISH8	NA	-1.6	0.6
NC_002607.1:34593..36073:minus NA ISH39-IS200_IS605_ssgr_IS1341	ISH39-IS200_IS605_ssgr_IS1341	NA	-1.6	0.8
VNG_RS10605	<i>gvpG</i>	gas vesicle protein GvpG	-1.6	0.5
VNG_RS03695	VNG_RS03695	type II/IV secretion system ATPase subunit	-1.5	0.0
VNG_RS04830	VNG_RS04830	PAS domain -containing sensor histidine kinase	-1.5	0.1

VNG_1404a	VNG_1404a	NA	-1.5	0.3
VNG_RS06185	VNG_RS06185	hypothetical protein	-1.5	0.0
VNG_RS06335	VNG_RS06335	energy-couplingfactorABC transporter substrate-binding protein	-1.5	0.4
VNG_RS06415	VNG_RS06415	hypothetical protein	-1.5	0.0
VNG_RS07545	VNG_RS07545	hypothetical protein	-1.5	0.3
NC_002607.1:1608078..1609480:plus NA ISH8B-IS4_ssgr_ISH8	ISH8B-IS4_ssgr_ISH8	NA	-1.5	0.6
VNG_RS14165	VNG_RS14165	hypothetical protein	-1.5	0.5
VNG_RS02745	VNG_RS02745	hypothetical protein	-1.5	0.5
VNG_RS00195	VNG_RS00195	glycosyl transferase family 4 protein	-1.5	0.4
VNG_RS00860	VNG_RS00860	hypothetical protein	-1.5	0.6
VNG_RS03335	VNG_RS03335	helix-turn-helix domain -containing protein	-1.5	0.1
VNG_RS03715	VNG_RS03715	hypothetical protein	-1.5	0.1
VNG_RS03910	VNG_RS03910	hypothetical protein	-1.5	0.1
NC_002608.1:136964..138366:plus NA ISH8E-IS4_ssgr_ISH8	ISH8E-IS4_ssgr_ISH8	NA	-1.5	0.6
VNG_RS00945	VNG_RS00945	MBLfold metallo-hydrolase	-1.4	0.0
VNG_RS01035	VNG_RS01035	transcriptioninitiationfactorlIB	-1.4	0.4

VNG_RS02085	VNG_RS02085	HD domain -containing protein	-1.4	0.6
VNG_RS09575	VNG_RS09575	aminopeptidase	-1.4	0.4
VNG_RS10735	VNG_RS10735	cytochrome ubiquinol oxidase subunit il	-1.4	0.4
VNG_RS11715	VNG_RS11715	ParA family protein	-1.4	0.1
VNG_7019	<i>gvpl1a</i>	NA	-1.4	0.6
VNG_RS00735	VNG_RS00735	restriction endonuclease	-1.4	0.2
VNG_RS01285	<i>pyk</i>	pyruvate kinase	-1.4	0.1
VNG_RS01345	VNG_RS01345	segregation/condensation protein A	-1.4	0.2
VNG_RS02080	VNG_RS02080	AIM24 family protein	-1.4	0.4
VNG_RS02785	VNG_RS02785	MLB fold metallo-hydrolase	-1.4	0.5
VNG_RS03700	VNG_RS03700	ATPase	-1.4	0.1
VNG_RS03725	VNG_RS03725	PAS domain -containing protein	-1.4	0.1
VNG_RS09295	VNG_RS09295	tryptophanase	-1.4	0.4
VNG_RS11310	VNG_RS11310	TrkA family potassium uptake protein	-1.4	0.5
VNG_RS02450	VNG_RS02450	hypothetical protein	-1.3	0.0
VNG_1381H	VNG_1381H	NA	-1.3	0.3
VNG_RS08005	VNG_RS08005	MLB fold metallo-hydrolase	-1.3	0.0
VNG_RS01350	<i>smc</i>	chromosome segregation protein SMC	-1.3	0.1
VNG_RS05670	VNG_RS05670	GNAT family N-acetyl transferase	-1.3	0.1

VNG_RS14235	VNG_RS14235	hypothetical protein	-1.3	0.1
VNG_RS09785	VNG_RS09785	non-histonechromosomalMC1 family protein	-1.3	0.1
VNG_RS00150	VNG_RS00150	transcriptional regulator	-1.3	0.1
VNG_RS02200	VNG_RS02200	Na(+)/H(+) antiporter subunit D	-1.3	0.2
VNG_RS03225	VNG_RS03225	typell/IV secretion system ATPase subunit	-1.3	0.1
VNG_RS03685	VNG_RS03685	MinD/ParA family protein	-1.3	0.2
VNG_RS03785	VNG_RS03785	chemotaxis protein CheA	-1.3	0.2
VNG_RS07355	VNG_RS07355	ATP-binding protein	-1.3	0.4
VNG_RS08400	VNG_RS08400	replication protein A	-1.3	0.1
VNG_RS08635	<i>lipA</i>	lipoyl synthase	-1.3	0.1
VNG_RS10590	VNG_RS10590	gas vesicle protein K	-1.3	0.4
VNG_RS05200	VNG_RS05200	superoxide dismutase	-1.2	0.1
VNG_RS12425	VNG_RS12425	cytochrome ubiquinol oxidase subunit I	-1.2	0.1
VNG_RS06230	VNG_RS06230	chemotaxis protein CheC	-1.2	0.3
VNG_RS10365	VNG_RS10365	hypothetical protein	-1.2	0.0
VNG_RS10730	VNG_RS10730	cytochromeubiquinoloxidasesubunitI	-1.2	0.3
VNG_RS11990	VNG_RS11990	hypothetical protein	-1.2	0.3
VNG_RS00155	VNG_RS00155	hypothetical protein	-1.2	0.2
VNG_RS00955	VNG_RS00955	prefoldin subunit beta	-1.2	0.2

VNG_RS01405	VNG_RS01405	hypothetical protein	-1.2	0.1
VNG_RS02225	VNG_RS02225	DUF4040 domain -containing protein	-1.2	0.1
VNG_RS03220	VNG_RS03220	type1 secretion system F family protein	-1.2	0.1
VNG_RS08405	VNG_RS08405	RPA family protein	-1.2	0.2
VNG_RS11450	VNG_RS11450	cell division protein	-1.2	0.2
VNG_RS01255	VNG_RS01255	cobalamin-binding protein	-1.1	0.0
VNG_RS03790	VNG_RS03790	chemotaxis protein CheB	-1.1	0.1
VNG_RS11315	VNG_RS11315	amino acid permease	-1.1	0.0
VNG_RS01615	<i>purB</i>	adenylosuccinate lyase	-1.1	0.1
VNG_RS02690	VNG_RS02690	phytoene/squalene synthase family protein	-1.1	0.1
VNG_RS05930	VNG_RS05930	MgtC/SapB family protein	-1.1	0.1
VNG_RS04825	VNG_RS04825	aminopeptidaseP family protein	-1.0	0.0

Tabela S2: Tabela de genes mais expressos em hiposalinidade. Contém o locus, nome do gene, produto, valor de M que representa o $\log_2 (L \text{ (low salinity)} / O \text{ (optimal salinity)})$ e o desvio padrão dos valores de M das replicatas.

<i>Locus</i>	Nome do gene	Produto	M	sdM
VNG_RS11470	VNG_RS11470	metalABC transporter substrate-binding protein	4.4	0.6
VNG_1642H	VNG_1642H	NA	3.7	1.3
VNG_RS11585	<i>queC</i>	7-cyano-7-deazaguanine synthase QueC	3.7	1.1

VNG_RS09085	VNG_RS09085	GNAT family N-acetyl transferase	3.6	0.8
VNG_RS11010	VNG_RS11010	hypothetical protein	3.6	2.1
VNG_RS03620	VNG_RS03620	LLM class flavin-dependent oxidoreductase	3.5	1.2
VNG_RS03640	VNG_RS03640	nitroreductase family protein	3.4	0.5
VNG_RS04655	VNG_RS04655	superoxide dismutase	3.4	0.2
VNG_RS05350	VNG_RS05350	hypothetical protein	3.4	2.2
VNG_RS12050	VNG_RS12050	site-specific integrase	3.2	0.5
VNG_RS02190	VNG_RS02190	DJ-1/Pfpl family protein	3	0.9
VNG_RS11590	VNG_RS11590	7-carboxy-7-deazaguaninesynthaseQueE	3	0.5
VNG_RS02760	VNG_RS02760	MarR family transcriptional regulator	3	1.1
VNG_RS05265	VNG_RS05265	EamA family transporter	3	1.3
VNG_1238C	VNG_1238C	NA	2.9	1.8
VNG_RS07635	<i>glpA</i>	anaerobic glycerol-3-phosphate dehydrogenase subunit A	2.9	0.1
VNG_RS06370	VNG_RS06370	adenosyl cobalamin-dependent ribonucleoside-diphosphate reductase	2.9	0.2
VNG_RS05750	VNG_RS05750	hypothetical protein	2.9	1.3
VNG_RS05635	VNG_RS05635	5'-nucleotidaseC-terminal domain -containing protein	2.8	0.4
VNG_RS08190	VNG_RS08190	glutathioneS- transferaseN-terminal domain -containing protein	2.8	0.9

VNG_RS08445	VNG_RS08445	iron-containing alcohol dehydrogenase	2.8	1.3
VNG_RS04300	VNG_RS04300	tripartite tricarboxylate transporter permease	2.7	1.1
VNG_RS09480	VNG_RS09480	O-acetyl homoserine amino carboxypropyl transferase/cysteine synthase	2.7	0.2
VNG_RS04230	VNG_RS04230	UPF0058 family protein	2.7	0.6
VNG_RS05755	VNG_RS05755	RND family transporter	2.7	1.4
VNG_RS07670	<i>trmY</i>	tRNA(pseudouridine(54)-N(1))-methyl transferase TrmY	2.7	1.1
VNG_RS07645	VNG_RS07645	anaerobic glycerol-3-phosphate dehydrogenase subunit C	2.6	1.1
VNG_RS11220	VNG_RS11220	universal stress protein	2.6	0.2
VNG_RS11645	VNG_RS11645	hypothetical protein	2.6	2.2
VNG_0057H	VNG_as00215_75	NA	2.6	0.5
VNG_RS14195	VNG_RS14195	hypothetical protein	2.6	0.7
VNG_RS05490	VNG_RS05490	acyl-CoA thioesterase	2.6	0.6
VNG_RS01225	VNG_RS01225	fructose-bisphosphate aldolase	2.5	1.8
VNG_RS07640	<i>glpB</i>	glycerol-3-phosphate dehydrogenase subunit GlpB	2.5	0.8
VNG_7140	VNG_7140	NA	2.5	1.1
VNG_RS12945	VNG_RS12945	type IV toxin-antitoxin systemAbiEi family antitoxin domain -containing protein	2.5	1.6
VNG_RS01530	VNG_RS01530	diphthine--ammonia ligase	2.4	0.5

VNG_1598a	<i>tfx</i>	NA	2.4	1.2
VNG_RS07035	<i>carA</i>	glutamine-hydrolyzing carbamoyl -phosphate synthase small subunit	2.4	1.5
VNG_1139G	VNG_as04165_63 8	NA	2.4	0.2
VNG_1173a	VNG_1173a	NA	2.3	0.5
VNG_RS04855	VNG_RS04855	amino acid permease	2.3	0.5
VNG_RS05550	VNG_RS05550	amphi-Trp domain -containing protein	2.3	0.2
VNG_2185	<i>parA11</i>	NA	2.3	0.5
VNG_RS11465	VNG_RS11465	metal ABC transporter ATP-binding protein	2.3	0.4
VNG_RS04725	<i>hutU</i>	urocanate hydratase	2.3	0.7
VNG_RS08525	VNG_RS08525	HIT family protein	2.3	1
VNG_RS09655	VNG_RS09655	cysteine desulfurase	2.3	1
VNG_RS09665	<i>radA</i>	DNA repair and recombination protein RadA	2.3	0.1
VNG_RS11630	<i>arcC</i>	carbamate kinase	2.3	0.4
VNG_RS12760	VNG_RS12760	hypothetical protein	2.3	1.8
VNG_RS02880	VNG_RS02880	transcriptioninitiationfactorIIIB	2.2	0.5
VNG_RS07810	VNG_RS07810	UMP kinase	2.2	1.6
VNG_RS09675	VNG_RS09675	class I SAM-dependent methyl transferase family protein	2.2	0.9
VNG_RS12735	VNG_RS12735	TATA-box-binding protein	2.2	1.2

VNG_RS05315	VNG_RS05315	ornithine cyclodeaminase family protein	2.2	0.8
VNG_RS02855	VNG_RS02855	hypothetical protein	2.2	1
VNG_RS03635	VNG_RS03635	NADH:flavin oxidoreductase/NADH oxidase	2.2	0.6
VNG_RS07030	<i>carB</i>	carbamoyl -phosphate synthase large subunit	2.2	0.7
VNG_RS11040	VNG_RS11040	TATA-box-binding protein	2.2	0
VNG_RS04225	VNG_RS04225	hypothetical protein	2.1	0.9
VNG_RS06005	VNG_RS06005	chloride channel protein	2.1	0.2
VNG_RS06810	VNG_RS06810	hypothetical protein	2.1	0.1
VNG_RS06830	VNG_RS06830	bacteriorhodopsin	2.1	1.2
VNG_RS12695	VNG_RS12695	DEAD/DEAHboxhelicase family protein	2.1	1.3
VNG_RS04740	<i>hutH</i>	histidine ammonia-lyase	2.1	1
VNG_RS06855	VNG_RS06855	creatininase family protein	2.1	1
VNG_RS07295	<i>nadA</i>	quinolinate synthase NadA	2.1	0.4
VNG_RS12905	VNG_RS12905	hypothetical protein	2.1	0.4
VNG_RS04155	VNG_RS04155	glycosyl transferase	2	0.6
VNG_RS05650	VNG_RS05650	hypothetical protein	2	1.3
VNG_RS05660	VNG_RS05660	DoxX family protein	2	0.6
VNG_RS09070	VNG_RS09070	ABC transporter ATP-binding protein	2	1.2
VNG_RS12855	VNG_RS12855	TRC40/GET3/ArsA family transport-energizing ATPase	2	0.6

VNG_RS00045	VNG_RS00045	sugar transferase	2	1.1
VNG_RS03005	VNG_RS03005	hypothetical protein	2	1.1
VNG_RS10950	VNG_RS10950	hypothetical protein	2	1.1
VNG_0026C	VNG_as00105_20	NA	2	0.2
VNG_RS00375	VNG_RS00375	typell glyceraldehyde-3-phosphate dehydrogenase	2	1
VNG_RS03950	VNG_RS03950	glutaredoxin	2	1.1
VNG_RS07490	VNG_RS07490	hypothetical protein	2	0.6
VNG_2021a	VNG_2021a	NA	2	0.3
VNG_RS10935	VNG_RS10935	phageintegraseSAM-like domain -containing protein	2	0.4
VNG_6393a	VNG_6393a	NA	2	0.7
VNG_RS13675	VNG_RS13675	hypothetical protein	2	0.1
VNG_RS01120	VNG_RS01120	FAD-binding oxidoreductase	1.9	0.1
VNG_RS04795	VNG_RS04795	DUF2064 domain -containing protein	1.9	0.1
VNG_RS05480	<i>msrB</i>	peptide-methionine(R)-S-oxidoreductase MsrB	1.9	0.1
VNG_RS06870	VNG_RS06870	HAD family hydrolase	1.9	0.8
VNG_RS09370	VNG_RS09370	sulfurtransferase	1.9	0.5
VNG_RS12070	VNG_RS12070	helix-turn-helix domain -containing protein	1.9	0.5
VNG_RS02095	VNG_RS02095	metal-dependent transcriptional regulator	1.9	0.1
VNG_7037c	VNG_7037c	NA	1.9	1

VNG_RS03395	VNG_RS03395	transcriptioninitiationfactorIIb	1.9	0.4
VNG_RS03565	VNG_RS03565	PAS domain S-box protein	1.9	0.4
VNG_RS04575	VNG_RS04575	elongationfactor1-beta	1.9	0.4
VNG_RS07135	VNG_RS07135	hypothetical protein	1.9	1.1
VNG_RS12910	VNG_RS12910	DNA-binding protein	1.9	0.4
VNG_RS04790	VNG_RS04790	AAA family ATPase	1.8	0.4
VNG_RS05030	VNG_RS05030	winged helix-turn-helix transcriptional regulator	1.8	0.2
VNG_RS05365	VNG_RS05365	Lrp/AsnC family transcriptional regulator	1.8	0.9
VNG_RS09840	VNG_RS09840	30Sribosomal protein S6e	1.8	0.1
VNG_RS10850	VNG_RS10850	nucleotidyl transferase AbiEii/AbiGii toxin family protein	1.8	0.4
VNG_RS12755	VNG_RS12755	hypothetical protein	1.8	1.1
VNG_7113	VNG_7113	NA	1.8	0.1
VNG_RS13310	VNG_RS13310	hypothetical protein	1.8	0.7
VNG_RS05460	VNG_RS05460	DUF309 domain -containing protein	1.8	0.8
VNG_RS06985	VNG_RS06985	hypothetical protein	1.8	0.7
VNG_RS07250	VNG_RS07250	extracellular solute-binding protein	1.8	0.8
VNG_RS09745	VNG_RS09745	cupin domain -containing protein	1.8	0.1
VNG_0277a	VNG_0277a	NA	1.8	1
VNG_RS07310	VNG_RS07310	TrmB family transcriptional regulator	1.8	0.6

VNG_RS01155	VNG_RS01155	YkgJ family cysteine cluster protein	1.7	0.1
VNG_RS05310	VNG_RS05310	DUF3054 domain -containing protein	1.7	0.2
VNG_RS05580	VNG_RS05580	deoxyhypusine synthase	1.7	0.5
VNG_RS07065	VNG_RS07065	molybdenum cofactor biosynthesis protein MoaB	1.7	0.2
VNG_RS10315	VNG_RS10315	hypothetical protein	1.7	0.6
VNG_RS13500	VNG_RS13500	hypothetical protein	1.7	0.6
VNG_RS13630	VNG_RS13630	hypothetical protein	1.7	0.4
VNG_RS11805	VNG_RS11805	type BDNA-directed DNA polymerase	1.7	0.1
VNG_RS05740	VNG_RS05740	hypothetical protein	1.7	0.8
VNG_RS08625	VNG_RS08625	DEAD/DEAHboxhelicase	1.7	0.8
VNG_0280H	VNG_0280H	NA	1.7	0.3
VNG_RS05070	VNG_RS05070	tRNA4-thiouridine(8) synthase Thil	1.7	1
VNG_RS09735	VNG_RS09735	PHP domain -containing protein	1.7	0.4
VNG_RS11595	VNG_RS11595	6-pyruvoyl tetrahydropterin synthase family protein	1.7	0
VNG_RS12090	VNG_RS12090	helix-turn-helix domain -containing protein	1.7	1
VNG_0535C	VNG_0535C	NA	1.7	0.6
VNG_RS06835	VNG_RS06835	HAMP domain -containing protein	1.7	0.1
VNG_RS10120	<i>engB</i>	GTP-binding protein EngB	1.7	0.1
VNG_RS02480	VNG_RS02480	CoA esterlyase	1.6	0.2

VNG_RS04075	VNG_RS04075	hypothetical protein	1.6	0.5
VNG_RS07405	VNG_RS07405	hypothetical protein	1.6	0.2
VNG_2335H	VNG_2335H	NA	1.6	0.5
VNG_RS10100	VNG_RS10100	rhodanese-like domain -containing protein	1.6	0.5
VNG_RS04650	VNG_RS04650	hypothetical protein	1.6	0.1
VNG_RS05080	VNG_RS05080	PLP-dependent cysteine synthase family protein	1.6	0.8
VNG_RS07080	VNG_RS07080	hypothetical protein	1.6	0.4
VNG_RS13190	VNG_RS13190	hypothetical protein	1.6	0.8
VNG_RS04340	VNG_RS04340	VOC family protein	1.6	0
VNG_RS04470	VNG_RS04470	VOC family protein	1.6	0.6
VNG_RS04860	<i>surE</i>	5'/3'-nucleotidase SurE	1.6	0.6
VNG_RS05040	VNG_RS05040	hypothetical protein	1.6	0.6
VNG_RS05765	VNG_RS05765	hypothetical protein	1.6	0.1
VNG_RS08725	<i>rnz</i>	ribonuclease Z	1.6	0.4
VNG_RS09470	VNG_RS09470	MFS transporter	1.6	0.6
VNG_RS09645	VNG_RS09645	hypothetical protein	1.6	0.7
VNG_RS10810	VNG_RS10810	site-specific integrase	1.6	0.6
VNG_RS02275	VNG_RS02275	hypothetical protein	1.5	0.1
VNG_RS05035	VNG_RS05035	RNA-binding protein	1.5	0.4

VNG_RS11775	VNG_RS11775	lamintail domain -containing protein	1.5	0.4
VNG_RS07025	VNG_RS07025	erythromycin esterase family protein	1.5	0.2
VNG_RS08620	VNG_RS08620	DUF91 domain -containing protein	1.5	0.2
VNG_RS13030	VNG_RS13030	helix-turn-helix domain -containing protein	1.5	0.5
VNG_RS00385	VNG_RS00385	RimK family alpha-L-glutamate ligase	1.5	0.7
VNG_RS02330	<i>nth</i>	endonuclease III	1.5	0
VNG_RS03450	VNG_RS03450	HAD-IIA family hydrolase	1.5	0.3
VNG_RS04610	<i>msrA</i>	peptide-methionine(S)-S-oxidereductaseMsrA	1.5	0.3
VNG_RS08375	VNG_RS08375	radical SAM protein	1.5	0.4
VNG_RS10875	VNG_RS10875	SOS response-associated peptidase	1.5	0.1
VNG_RS12950	VNG_RS12950	nucleotidyl transferase AbiEii/AbiGii toxin family protein	1.5	0.4
VNG_RS01680	VNG_RS01680	NUDIX domain -containing protein	1.4	0.5
VNG_RS03130	VNG_RS03130	ABC transporter ATP-binding protein	1.4	0.2
VNG_RS04320	VNG_RS04320	50Sribosomal protein L11	1.4	0.5
VNG_RS06540	VNG_RS06540	MFS transporter	1.4	0.4
VNG_RS09920	VNG_RS09920	dihydroorotase	1.4	0.2
VNG_RS09940	VNG_RS09940	cysteine hydrolase	1.4	0.5
VNG_RS09960	VNG_RS09960	aminopeptidase	1.4	0.2
VNG_RS10170	<i>trxA</i>	thioredoxin	1.4	0.2

VNG_RS10235	<i>tadA</i>	Flp pilus assembly complex ATPase component TadA	1.4	0.1
VNG_RS10310	<i>uvrA</i>	excinuclease ABC subunit UvrA	1.4	0.1
VNG_0057H	VNG_as00215_76	NA	1.4	0.4
VNG_RS01080	VNG_RS01080	SDR family oxidoreductase	1.4	0.4
VNG_RS03600	VNG_RS03600	iron ABC transporter substrate-binding protein	1.4	0.6
VNG_RS09530	<i>argH</i>	argininosuccinate lyase	1.4	0.1
VNG_RS09590	VNG_RS09590	hypothetical protein	1.4	0.3
VNG_RS05405	VNG_RS05405	acyl transferase	1.3	0.4
VNG_RS08250	VNG_RS08250	4-demethyl wyosine synthase TYW1	1.3	0.1
VNG_RS00730	VNG_RS00730	50S ribosomal protein L15e	1.3	0.1
VNG_RS01235	VNG_RS01235	3-dehydroquinate synthase I	1.3	0.4
VNG_RS04815	<i>aroC</i>	chorismate synthase	1.3	0
VNG_RS06875	VNG_RS06875	4a-hydroxytetrahydrobiopterin dehydratase	1.3	0.4
VNG_RS08605	VNG_RS08605	hypothetical protein	1.3	0.4
VNG_RS08955	VNG_RS08955	(2Fe-2S)-binding protein	1.3	0
VNG_RS02280	VNG_RS02280	hypothetical protein	1.2	0.1
VNG_0892a	VNG_0892a	NA	1.2	0.1
VNG_RS05475	VNG_RS05475	helix-turn-helix transcriptional regulator	1.2	0.1

VNG_RS07910	<i>rdgB</i>	RdgB/HAM1 family non-canonical purine NTP pyrophosphatase	1.2	0.1
VNG_RS08780	VNG_RS08780	nucleoid-structuring protein H-NS	1.2	0.2
VNG_RS09885	VNG_RS09885	hypothetical protein	1.2	0.1
VNG_7133	MITEHsaI5-ISNCY	NA	1.2	0.2
VNG_RS02135	VNG_RS02135	proteasome assembly chaperone family protein	1.2	0.3
VNG_RS04180	VNG_RS04180	transcription factor	1.2	0.0
VNG_RS08915	<i>hisI</i>	phosphoribosyl -AMP cyclohydrolase	1.2	0.3
VNG_6367a	VNG_6367a	NA	1.2	0
VNG_RS12730	VNG_RS12730	DEAD/DEAHbox helicase family protein	1.2	0
VNG_1675a	VNG_1675a	NA	1.2	0.1
VNG_RS09445	VNG_RS09445	helix-turn-helix transcriptional regulator	1.2	0.1
VNG_RS09595	VNG_RS09595	hypothetical protein	1.2	0.1
VNG_RS09620	<i>ftsY</i>	signal recognition particle-docking protein FtsY	1.2	0.1
VNG_RS09775	VNG_RS09775	phenyl alanine--tRNA ligase subunit beta	1.2	0.1
VNG_RS03345	<i>dph2</i>	diphthamide biosynthesis enzyme Dph2	1.1	0.1
VNG_RS09855	VNG_RS09855	BtpA/SgcQ family protein	1.1	0.1
VNG_RS12990	VNG_RS12990	site-specific integrase	1.1	0.1
VNG_RS00380	VNG_RS00380	Hsp20/alpha crystallin family protein	1.1	0.0
VNG_RS08240	<i>ribB</i>	3,4-dihydroxy-2-butanone-4-phosphate synthase	1.1	0.1

VNG_RS09455	VNG_RS09455	S26 family signal peptidase	1.1	0.0
NC_002608.1: 287164..28889 5:plus NA ISH4 0-IS200_IS605	ISH40-IS200_IS60 5	NA	1.1	0.0
VNG_RS04905	VNG_RS04905	DEAD/DEAHboxhelicase	1	0.0

Tabela S3: Tabela de replicatas do RNA-seq vs microarranjos de DNA em hipersalinidade.

Locus	Nome do gene	Produto	rep1	rep2	MA
VNG_RS09260	VNG_RS09260	malate dehydrogenase	0,1	0,4	0,4
VNG_RS09310	VNG_RS09310	ABC transporter permease	0,5	0	0,5
VNG_RS08065	VNG_RS08065	phosphonateABC transporter ATP-binding protein	0,4	2,3	-0,5
VNG_RS08490	VNG_RS08490	histidine kinase	0,4	0,9	-0,3
VNG_RS07035	carA	glutamine-hydrolyzing carbamoyl -phosphatesynthasessmallsubunit	-0,5	-0,8	1,2

Tabela S4: Tabela de replicatas do RNA-seq vs proteômica em hipersalinidade.

Locus	Nome do gene	Produto	rep1	rep2	Proteômica
VNG_RS00860	VNG_RS00860	hypothetical protein	1,1	1	0,7
VNG_RS01050	VNG_RS01050	helix-turn-helix transcriptional regulator	0	0,5	0,2
VNG_RS01290	metG	methionine--tRNAligase	0,4	0,5	0,4
VNG_RS01850	VNG_RS01850	2-oxoacid:acceptor oxidoreductase subunit alpha	0,9	1,6	0,3
VNG_RS02485	VNG_RS02485	Glu/Leu/Phe/Val dehydrogenase	0	0,5	0,8

VNG_RS03190	VNG_RS03190	NADP-dependent oxidoreductase	0,5	1,4	0,7
VNG_RS03480	VNG_RS03480	Rrf2 family transcriptional regulator	0,3	0,5	0,5
VNG_RS04420	VNG_RS04420	30S ribosomal protein S11	0,3	2,4	0,2
VNG_RS04460	<i>eno</i>	phosphopyruvate hydratase	0,5	1,8	0,5
VNG_RS04685	VNG_RS04685	NAD(P)/FAD-dependent oxidoreductase	0,1	0,2	0,3
VNG_RS04780	VNG_RS04780	hypothetical protein	0	0,2	0,4
VNG_RS05290	VNG_RS05290	class II fumarate hydratase	0	0,9	0,3
VNG_RS05955	VNG_RS05955	acetyl /propionyl /methyl crotonyl-CoA carboxylase subunit alpha	0	0,4	0,3
VNG_RS06370	VNG_RS06370	adenosyl cobalamin-dependent ribonucleoside-diphosphate reductase	0,4	0,5	0,2
VNG_RS06850	VNG_RS06850	translation initiation factor IF-5A	0	1	0,1
VNG_RS07430	VNG_RS07430	peptidyl prolyl isomerase	0	0,7	1
VNG_RS08070	VNG_RS08070	phosphate/phosphite/phosphonate ABC transporter substrate-binding protein	0,8	0,3	0,1
VNG_RS08225	VNG_RS08225	branched-chain amino acid transaminase	0,1	0	0,5
VNG_RS08295	VNG_RS08295	ATP synthase subunit A	0,8	0,9	0,3
VNG_RS09205	VNG_RS09205	twin-arginine translocation signal domain -containing protein	0,8	0	1,3
VNG_RS09665	<i>radA</i>	DNA repair and recombination protein RadA	0,5	0,1	0,6
VNG_RS10425	VNG_RS10425	30S ribosomal protein S12	0,3	1,2	0,5

VNG_RS10505	<i>csg</i>	major cell surface glycoprotein	0,1	0,2	1,1
VNG_RS02195	VNG_RS02195	purine phosphoribosyl transferase family protein	-0,1	-0,1	-0,5
VNG_RS04390	VNG_RS04390	2-oxoacid:acceptor oxidoreductase subunit alpha	-0,3	-1	-0,4
VNG_RS06100	VNG_RS06100	precorrin-8Xmethyl mutase	-0,7	-0,7	-0,5
VNG_RS07035	<i>carA</i>	glutamine-hydrolyzing carbamoyl-phosphatesynthasessmallsubunit	-0,5	-0,8	-0,7
VNG_RS09495	VNG_RS09495	phosphoglycerate dehydrogenase	-1,2	-0,5	-0,3
VNG_RS09530	<i>argH</i>	argininosuccinate lyase	-2,8	-2,9	-0,6
VNG_RS09535	VNG_RS09535	argininosuccinate synthase	-1,7	-2,5	-0,8
VNG_RS11625	<i>argF</i>	ornithinecarbamoyl transferase	-0,2	-0,3	-0,3
VNG_RS00730	VNG_RS00730	50Sribosomal protein L15e	0,4	1,4	-0,5
VNG_0287a	VNG_0287a	NA	0,7	0,8	-0,4
VNG_RS01540	VNG_RS01540	coiled-coil protein	1,6	0	-0,3
VNG_RS03095	VNG_RS03095	30Sribosomal protein S15	1	3	-0,2
VNG_RS03340	VNG_RS03340	yl bF family regulator	0,4	0,4	-0,5
VNG_RS03940	VNG_RS03940	flagellin	0	0	-0,5
VNG_RS04315	VNG_RS04315	50Sribosomal protein L1	0,2	1,4	-0,4
VNG_RS04320	VNG_RS04320	50Sribosomal protein L11	0,1	1,6	-0,5
VNG_RS04410	VNG_RS04410	30Sribosomal protein S13	0,4	2,6	-0,3
VNG_RS04440	VNG_RS04440	50Sribosomal protein L13	0,8	1,4	-0,4
VNG_RS04530	<i>ndk</i>	nucleoside-diphosphate kinase	0,3	2,3	-0,3

VNG_RS06250	VNG_RS06250	DUF2150 family protein	0,4	0,9	-0,4
VNG_RS06355	VNG_RS06355	hypothetical protein	0	0	-0,2
VNG_RS06555	VNG_RS06555	50Sribosomal protein L3	0,5	2,4	-0,3
VNG_RS06560	<i>rpl4p</i>	50Sribosomal protein L4	0,3	2,2	-0,4
VNG_RS06575	VNG_RS06575	30Sribosomal protein S19	0,7	2,7	-0,4
VNG_RS06585	VNG_RS06585	30Sribosomal protein S3	0,5	2,6	-0,4
VNG_RS06625	VNG_RS06625	30Sribosomal protein S14	0,5	3,3	-0,3
VNG_RS06640	VNG_RS06640	50Sribosomal protein L32e	0	2,6	-0,4
VNG_RS06645	VNG_RS06645	50Sribosomal protein L19e	0	2,9	-0,4
VNG_RS06650	VNG_RS06650	50Sribosomal protein L18	0,1	2,8	-0,5
VNG_RS07740	<i>infB</i>	translationinitiationfactorIF-2	0,1	0,8	-0,9
VNG_RS07930	<i>rps24e</i>	30Sribosomal protein S24e	0,8	1,5	-0,3
VNG_RS08405	VNG_RS08405	RPA family protein	0,5	0,2	-0,5
VNG_RS09210	VNG_RS09210	CBS domain -containing protein	0,3	0,7	-0,1
VNG_RS09775	VNG_RS09775	phenyl alanine--tRNAligasesubunitbeta	0,1	0,2	-0,2
VNG_RS10380	VNG_RS10380	30Sribosomal protein S10	0,2	1,4	-0,3
VNG_RS01305	<i>ppsA</i>	pyruvate,water dikinase	-0,6	-0,4	0,4
VNG_RS02920	VNG_RS02920	universal stress protein	-0,2	-1,2	1,1
VNG_RS03880	<i>acs</i>	acetate--CoA ligase	-0,2	-2	0,9
VNG_RS06045	VNG_RS06045	cobalt-precorrin-4/precorrin-4C(11)-m ethyl transferase	-0,2	-0,2	0,5
VNG_RS06070	VNG_RS06070	hypothetical protein	-0,5	-0,1	0,4

VNG_RS06085	VNG_RS06085	transcriptional regulator	-0,9	-0,4	1,1
VNG_RS07865	VNG_RS07865	NAD+synthase	-0,9	-0,5	0,3
VNG_RS09295	VNG_RS09295	tryptophanase	-0,5	-0,3	0,3

Tabela S5: Tabela de replicatas do RNA-seq vs microarranjos de DNA em hiposalinidade.

<i>Locus</i>	Nome do gene	Product	rep1	rep2	MA
VNG_RS00395	VNG_RS00395	cold-shock protein	0,5	1,2	0,5
VNG_RS03955	VNG_RS03955	methyl -accepting chemotaxis protein	2	0,4	0,6
VNG_RS04655	VNG_RS04655	Superoxide dismutase	3,5	3,2	0,9
VNG_RS12735	VNG_RS12735	TATA-box-binding protein	3	1,3	0,7
VNG_RS05495	VNG_RS05495	5'-nucleotidase C-terminal domain -containing protein	-2,2	-0,8	-0,5
VNG_RS06330	<i>cbiQ</i>	Cobalt ECF transporter T component CbiQ	-1,9	-1,6	-0,3
VNG_RS06335	VNG_RS06335	Energy-coupling factor ABC transporter substrate-binding protein	-1,8	-1,2	-0,6
VNG_RS08195	VNG_RS08195	NCS2 family permease	-0,5	-0,2	-0,6
VNG_RS08295	VNG_RS08295	ATP synthase subunit A	-1,6	-0,6	-0,6
VNG_RS08300	VNG_RS08300	V-type ATP synthase subunit F	-1,3	-0,3	-0,6
VNG_RS08305	VNG_RS08305	V-type ATP synthase subunit C	-1,5	-0,4	-0,7
VNG_RS08310	VNG_RS08310	V-type ATP synthase subunit E	-1,3	-0,8	-0,8
VNG_RS08315	VNG_RS08315	F0F1 ATP synthase subunit C	-2,1	-0,6	-0,7
VNG_RS08320	VNG_RS08320	V-type ATP synthase subunit I	-1,4	-0,1	-0,8

VNG_RS09190	VNG_RS09190	ABC transporter ATP-binding protein	-2,7	-1,4	-0,6
VNG_RS09195	VNG_RS09195	ABC transporter permease	-3,1	-1,4	-0,5
VNG_RS09200	VNG_RS09200	ABC transporter permease	-3,3	-1,2	-0,6
VNG_RS11620	<i>arcD</i>	arginine/ornithine antiporter ArcD	-2,1	-2	-0,3
VNG_RS07030	<i>carB</i>	carbamoyl -phosphatesynthaselargesubunit	1,7	2,7	-0,6
VNG_RS07035	<i>carA</i>	glutamine-hydrolyzing carbamoyl -phosphate synthase small subunit	1,3	3,4	-0,5
VNG_RS04375	VNG_RS04375	amino transferase class I/II-fold pyridoxal phosphate-dependent enzyme	-0,9	-1,5	0,7
VNG_RS05200	VNG_RS05200	Superoxide dismutase	-1,1	-1,3	1
VNG_RS06995	VNG_RS06995	Hsp20/alpha crystallin family protein	-0,8	-4	1
VNG_RS09230	VNG_RS09230	twin-arginine translocation signal domain -containing protein	-2,8	-3,2	0,8
VNG_RS09235	VNG_RS09235	ABC transporter permease	-2,9	-3,3	0,5

Tabela S6: Tabela de replicatas do RNA-seq vs proteômica em hiposalinidade.

Locus	Nome do gene	Produto	rep1	rep2	PTN
VNG_RS00375	VNG_RS00375	typell glyceraldehyde-3-phosphate dehydrogenase	1,3	2,7	1,1
VNG_RS02095	VNG_RS02095	metal-dependent transcriptional regulator	1,8	2	2,2
VNG_RS02490	VNG_RS02490	pyridoxal phosphate-dependent amino transferase	0,1	0	1

VNG_RS02695	VNG_RS02695	3-hydroxybutyryl-CoA dehydrogenase	0,8	1,3	1,5
VNG_RS03640	VNG_RS03640	nitroreductase family protein	3,7	3	1,4
VNG_RS04340	VNG_RS04340	VOC family protein	1,6	1,6	1
VNG_1173a	VNG_1173a	NA	2,6	1,9	0,7
VNG_RS04930	VNG_RS04930	hypothetical protein	1,2	0,7	1,3
VNG_RS05125	VNG_RS05125	enoyl-CoA hydratase/isomerase family protein	0	0	0,9
VNG_RS05735	VNG_RS05735	CDC48 family AAAATPase	1,6	0	0,4
VNG_RS05990	<i>sucC</i>	ADP-forming succinate--CoAligase subunit beta	0,3	1,1	0,3
VNG_RS06235	VNG_RS06235	NYN domain -containing protein	0,8	2,7	0,8
VNG_RS06370	VNG_RS06370	adenosyl cobalamin-dependent ribonucleoside-diphosphate reductase	3	2,7	1,6
VNG_RS07060	VNG_RS07060	zinc-binding dehydrogenase	0,2	0,4	0,7
VNG_RS07635	<i>glpA</i>	anaerobic glycerol-3-phosphate dehydrogenase subunit A	2,8	2,9	0,9
VNG_RS07700	VNG_RS07700	pre protein translocase subunit SecD	0,5	0,9	0,3
VNG_RS08110	VNG_RS08110	thermosome subunit	0,3	0,8	0,4
VNG_RS08585	VNG_RS08585	ribose-phosphate pyrophosphokinase	0,6	1,7	2
VNG_RS08610	VNG_RS08610	tryptophan--tRNAligase	0,8	1,2	0,8
VNG_RS09085	VNG_RS09085	GNAT family N-acetyl transferase	4,1	3	1,2

VNG_RS09290	VNG_RS09290	metallophosphoesterase family protein	0,1	0,1	0,5
VNG_RS09655	VNG_RS09655	cysteine desulfurase	3	1,6	1,6
VNG_RS09665	<i>radA</i>	DNA repair and recombination protein RadA	2,2	2,4	1,1
VNG_RS10170	<i>trxA</i>	thioredoxin	1,5	1,2	0,2
VNG_RS10325	VNG_RS10325	UDP-N-acetyl glucosamine2-epimerase	0,4	1,9	0,4
VNG_RS11430	VNG_RS11430	antibiotic biosynthesis monooxygenase	0,9	0,1	0,5
VNG_RS11625	<i>argF</i>	ornithine carbamoyl transferase	1,3	0,9	1,2
VNG_RS11630	<i>arcC</i>	carbamate kinase	2,6	2	0,8
VNG_RS11635	<i>arcA</i>	arginine deiminase	1	0,3	0,7
VNG_RS00630	VNG_RS00630	rod shape-determining protein	-0,8	-2	-1
VNG_RS00665	VNG_RS00665	Glu/Leu/Phe/Val dehydrogenase	-0,9	-0,5	-1,2
VNG_RS00860	VNG_RS00860	hypothetical protein	-1,9	-1	-1,4
VNG_RS00955	VNG_RS00955	prefoldin subunit beta	-1,3	-1	-0,9
VNG_0287a	VNG_0287a	NA	-1,2	-0,7	-0,6
VNG_RS01305	<i>ppsA</i>	pyruvate,water dikinase	-0,7	-0,1	-0,8
VNG_RS02110	VNG_RS02110	TAXI family TRAP transporter solute-binding subunit	-2,4	-1,2	-0,7
VNG_RS02195	VNG_RS02195	purine phosphoribosyl transferase family protein	-1,3	-1,8	-1,2
VNG_RS03735	VNG_RS03735	flagellin	-1,6	-1,7	-1,2

VNG_RS03740	VNG_RS03740	flagellin	-1,8	-1,9	-3,5
VNG_RS03745	VNG_RS03745	flagellin	-1,7	-1,9	-1,7
VNG_RS03940	VNG_RS03940	flagellin	-3	-2,9	-2,3
VNG_1158G	VNG_1158G	NA	-0,5	-0,7	-0,6
VNG_RS06060	<i>cobJ</i>	precorrin-3BC(17)-methyl transferase	-1,1	-0,5	-1,3
VNG_RS06100	VNG_RS06100	precorrin-8Xmethyl mutase	-3	-2,3	-2,7
VNG_RS06435	VNG_RS06435	HAMP domain -containing protein	-1,7	-1,9	-1,3
VNG_RS08070	VNG_RS08070	phosphate/phosphite/phosphonateA BC transporter substrate-binding protein	-2,1	-0,6	-0,8
VNG_RS08145	VNG_RS08145	citrate(Si)-synthase	-2,1	-2	-0,7
VNG_RS08225	VNG_RS08225	branched-chain amino acid transaminase	-1,6	-0,5	-0,5
VNG_RS08290	VNG_RS08290	ATPsynthasesubunitB	-1,4	-0,7	-0,5
VNG_RS08295	VNG_RS08295	ATPsynthasesubunitA	-1,6	-0,6	-1,2
VNG_RS08320	VNG_RS08320	V-type ATP synthase subunit I	-1,4	-0,1	-0,4
VNG_RS09205	VNG_RS09205	twin-arginine translocation signal domain -containing protein	-4,6	-3,1	-1,5
VNG_RS10505	<i>csg</i>	majorcellsurfaceglyco protein	-2,3	-1,6	-0,6
VNG_RS10600	<i>gvpH</i>	gas vesicle protein GvpH	-0,8	-1,5	-0,7
VNG_RS10730	VNG_RS10730	cytochromeubiquinoloxidasesubunitI	-1,4	-1	-0,5
VNG_RS12335	<i>gvpO</i>	gas vesicle protein GvpO	-1,5	-2,7	-0,3

VNG_RS12520	VNG_RS12520	type 2 isopentenyl-diphosphate Delta-isomerase	-0,1	-0,1	-0,8
VNG_RS12750	VNG_RS12750	hypothetical protein	-2,2	-3,4	-0,7
VNG_RS00730	VNG_RS00730	50Sribosomal protein L15e	1,2	1,4	-1,8
VNG_RS01845	VNG_RS01845	2-oxoacid:ferredoxin oxidoreductase subunit beta	0,7	0,7	-0,8
VNG_RS01850	VNG_RS01850	2-oxoacid:acceptor oxidoreductase subunit alpha	1,1	0,9	-1,1
VNG_RS02295	VNG_RS02295	cytochrome bc complex cytochrome b subunit	0,8	0,1	-1,6
VNG_RS03080	VNG_RS03080	30Sribosomal protein S3ae	0,1	1,5	-0,4
VNG_RS03095	VNG_RS03095	30Sribosomal protein S15	0,2	1,4	-0,4
VNG_RS04320	VNG_RS04320	50Sribosomal protein L11	1	1,7	-0,5
VNG_RS04410	VNG_RS04410	30Sribosomal protein S13	0,1	1,4	-0,6
VNG_RS04465	VNG_RS04465	30Sribosomal protein S2	0,3	0,9	-0,2
VNG_RS04935	VNG_RS04935	FAD-dependent oxidoreductase	2,3	0,2	-1,5
VNG_RS07085	<i>guaA</i>	glutamine-hydrolyzingGMPsynthase subunitGuaA	0,7	0	-0,2
VNG_RS07755	VNG_RS07755	proteasome-activating nucleotidase	0,3	0,4	-0,4
VNG_RS07790	VNG_RS07790	30Sribosomal protein S19e	0,3	1,7	-0,5
VNG_RS07930	<i>rps24e</i>	30Sribosomal protein S24e	0,1	0,5	-0,2
VNG_RS09535	VNG_RS09535	argininosuccinate synthase	1,3	0,8	-0,4
VNG_RS09625	<i>pfdA</i>	prefoldin subunit alpha	0,9	2,1	-0,7
VNG_RS09920	VNG_RS09920	dihydroorotase	1,2	1,5	-0,9

VNG_RS10310	<i>uvrA</i>	excinucleaseABC subunitUvrA	1,3	1,4	-0,6
VNG_RS10380	VNG_RS10380	30Sribosomal protein S10	0,7	1	-0,6
VNG_RS10385	<i>tuf</i>	translation elongation factor EF-1 subunit alpha	0,4	1,2	-0,4
VNG_RS11485	VNG_RS11485	glycerol dehydrogenase	1,5	0,6	-0,6
VNG_RS00670	VNG_RS00670	AMP-binding protein	-0,5	-0,5	0,9
VNG_RS01290	<i>metG</i>	methionine--tRNAligase	-0,9	-0,6	0,5
VNG_RS01665	VNG_RS01665	enoyl-CoA hydratase/isomerase family protein	-0,3	-0,8	0,4
VNG_RS02455	VNG_RS02455	S49 family peptidase	-1,1	-0,8	0,7
VNG_RS02595	<i>ctaD</i>	cytochromecoxidasesubunitI	-0,2	-0,2	1,7
VNG_RS05135	VNG_RS05135	hypothetical protein	-2,1	-1,5	3,1
VNG_RS05715	VNG_RS05715	bacteriorhodopsin	-3,9	-5,2	1,1
VNG_RS07000	VNG_RS07000	hypothetical protein	-0,6	-4,1	2,5
VNG_RS08675	VNG_RS08675	thermosome subunit	-0,6	-0,7	0,3
VNG_RS09755	VNG_RS09755	acyl-CoA dehydrogenase family protein	-1,4	-0,9	1,4

Tabela S7: Tabela contendo links para resultados detalhados do projeto:

Identificação	Conteúdos	Links
A	Pasta compartilhada Master's Dissertation - Todos os dados	http://bit.ly/3oCdu4g
B	Tabela genes menos expressos em hiposalinidade	https://bit.ly/3pEVo2O
C	Tabela genes mais expressos em hiposalinidade	http://bit.ly/3tcEJGa
D	Tabela de replicatas do RNA-seq vs microarranjos de DNA e proteômica em hipersalinidade	http://bit.ly/3q7wHMN

E	Tabela de replicatas do RNA-seq vs microarranjos de DNA e proteômica em hiposalinidade	http://bit.ly/3rHrQ5v
F	Dados usados na confecção das figuras 6 e 9.	http://bit.ly/2OvER3K
G	Dados usados na confecção das figuras 15, 16, 17 e 18.	http://bit.ly/3rKyDex
H	Quantificação e dimensionamento de vesículas de gás.	https://bit.ly/3eBS83Z
I	Resultados Blastn	https://bit.ly/2QD5Rjc
J	Resultados Blastx	https://bit.ly/3dWoxmK

Figura S1: Cobertura dos reads no gene *kef1* (VNG_RS07995). O gene *kef1* se encontra na fita reversa por isso os transcritos estão representados no sentido 3'-5'.