



Uma abordagem prática:

LEGIONELLA PNEUMOPHILA - Análise genómica

UC:

Bioinformática

Docentes:

Miguel Rocha

Pedro Soares

Grupo 1

Ana Paula Martins A62590

Ivo Ramalhosa A68362

Luís Costa A66516

Nuno Freitas A68366

Data da realização do trabalho: Novembro de 2014

Introdução

Este trabalho teve como principal objetivo a análise genómica da bactéria *Legionella Pneumophila* no âmbito da unidade curricular de Bioinformática. Foi assim proposto o estudo de uma parte do genoma desta bactéria recorrendo a ferramentas de programação. Toda a análise realizada e dados complementares (literatura, scripts, etc) pode ser consultada em <http://ebbioinformatica.weebly.com/scripts.html>

A bactéria *Legionella pneumophila* é o principal agente etiológico da doença do legionário (ou legionelose), uma forma de pneumonia lobar. Presente em diversos ambientes aquáticos, esta bactéria é um parasita intracelular de protozoários. A patogénese da doença do legionário é largamente devida à capacidade da *L. pneumophila* para invadir e proliferar dentro dos macrófagos alveolares, acreditando-se que esta habilidade resulta de uma adaptação anterior de nichos intracelulares na natureza. Na verdade, a *Legionella* intracelular exhibe uma notável capacidade de evitar as atividades endossomais e lisossomais bactericidas e de originar um único fagossoma replicativo. Nos últimos anos, muitos progressos tem sido feitos no sentido de identificar os fatores bacterianos que promovem a infeção intracelular e virulência.

Estes organismos são bastonetes gram-negativos pleomórficos aeróbios, bastante móveis, e muito exigentes nutricionalmente. O seu crescimento depende da presença de L-cisteína e de ferro em meios especiais. O organismo foi isolado em habitats aquáticos naturais (água doce e lagos, reservatórios de água) e fontes artificiais (torres de refrigeração, sistemas de distribuição de água potável). A temperatura de crescimento ótima situa-se entre os 28-40 °C. Estes organismos são latentes abaixo de 20 °C e são mortos a temperaturas acima dos 60 °C. A resposta do hospedeiro a este tipo de organismo é semelhante a outras respostas perante outros micro-organismos intracelulares oportunistas e apresenta ambos os mecanismos de resposta imunitária: inata e adquirida. A imunidade inata inclui as respostas de uma variedade de células hospedeiras e citoquinas, incluindo aquelas produzidas por macrófagos estimulados por antigénios microbianos. A imunidade adquirida consiste em linfócitos T ativados e nas citoquinas por eles produzidos, como o interferão (IFN) e outras citoquinas que ativam macrófagos com o objetivo de restringir o crescimento e propagação de bactérias intracelulares. O papel das citoquinas no que respeita especificamente à resistência e imunidade perante a *Legionella* é exemplificado por estudos sobre a natureza e o mecanismo pelo qual interferão produzido pelos linfócitos T ativados influencia macrófagos para resistir à infeção por parte desta bactéria, não só através da restrição do seu crescimento, mas também através da morte desta. Esta citoquina apresenta um papel chave na ativação de macrófagos na imunidade adquirida à *Legionella* e a outras bactérias intracelulares. Em particular, o interferão é conhecido por ter um papel crucial na ativação de macrófagos para resistir à infeção por *L. pneumophila*.

A doença do legionário (LD) foi reconhecida em 1976, após um surto de pneumonia numa convenção da Legião Americana em Filadélfia. Posteriormente, o agente etiológico foi identificado como uma fastidiosa bactéria gram-negativa chamada *Legionella pneumophila*. Apesar de várias espécies do género *Legionella* terem sido posteriormente identificadas, a *L. pneumophila* é a causa mais frequente da legionelose humana e uma causa relativamente comum de pneumonia adquirida em adultos, sendo esta mais rara em crianças.

Apesar de mais de 70 soro-grupos de Legionella terem sido identificados entre 50 espécies, a *L pneumophila* soro-grupo 1, sozinha, é responsável por 70-90% dos casos verificados em adultos. A legionelose apresenta duas síndromes clínicas diferentes: a **doença dos legionários**, que na maioria das vezes se manifesta como uma pneumonia severa acompanhada de uma doença multi-sistêmica, e a **febre de Pontiac**.

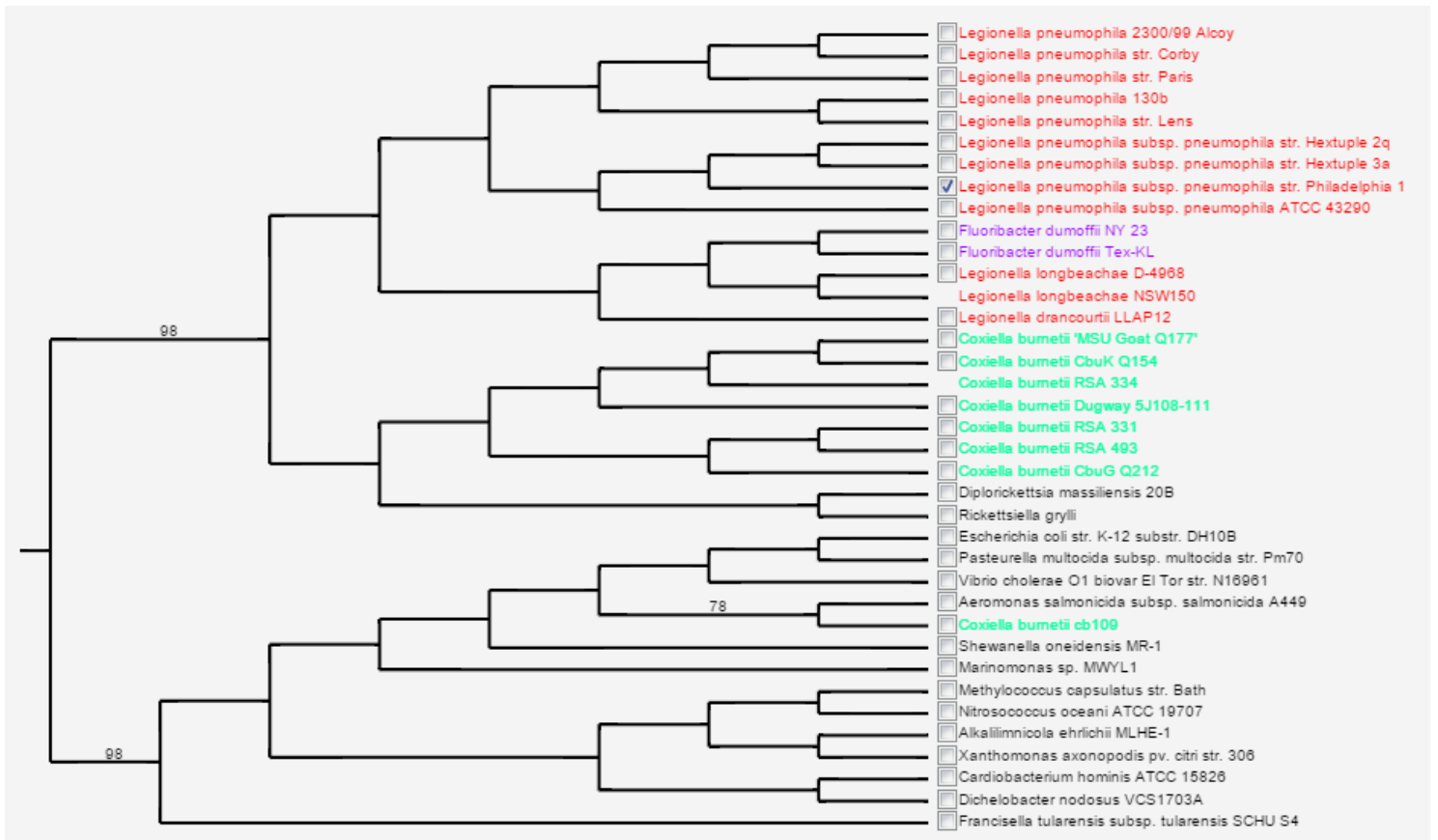


Figura 1 – Árvore filogenética (PATRIC)

A bactéria em questão apresenta a sua classificação Científica dirigida da seguinte forma:

Reino- Bactérias

Filo- Proteobactérias

Classe- GammaProteobacterias

Ordem- Legionellales

Famílias- Legionellaceae

Gênero- Legionella

Espécie- Legionella Pneumophila

Subespécie- *Legionella pneumophila subsp. Pneumophila* (figura 1)

Resultados e Discussão

Numa primeira fase procedeu-se à leitura da informação recolhida da base de dados genbank, diretamente da plataforma NCBI, utilizando apenas a sequência que corresponde ao intervalo de genes [lpg0001 a lpg0215], guardou-se essa sequência no ficheiro `sequence.gb` e leram-se as features do tipo CDS (*coding sequence for protein*) e o gene da sequência em questão.

A região codificante de um gene, também conhecida como a **sequência de codificação ou CDS** (sequência de codificação de DNA), é a porção de DNA (ou RNA) de um gene, composta por exões, que codifica a proteína. A região é limitada mais perto da extremidade 5' por um codão de iniciação e mais perto da extremidade 3' por um codão stop. Enquanto a identificação de **ORF's** (open reading frame) de uma sequência de DNA é simples, a identificação de sequências de codificação não o é, porque a célula traduz apenas um subconjunto de todas as ORF's para proteínas. A predição de CDS é a predição de um subconjunto de genes, incluindo a predição de sequências de DNA que codificam não só a proteína, mas também outros elementos funcionais, tais como os genes de RNA e sequências reguladoras.

O ficheiro `geral.py`, permite a obtenção de várias informações acerca do ficheiro, como a sequência e qual o seu alfabeto, a sua identificação, os grupos taxonómicos associados, bem como alguns qualificiers importantes para a sua identificação científica. Os qualificiers das features do tipo CDS não são os mesmos que os das features do tipo gene. O ficheiro `functionsizprotein.py` dá o tamanho de uma sequência e é utilizado para ver o tamanho das proteínas do tipo CDS.

A validação das features CDS e Gene com a tabela fornecida pelo NCBI, foi efetuada vendo se os Gene ID's das sequências de cada feature apresentam correspondência com os Gene ID's dos genes na tabela fornecida pelo NCBI. Como anteriormente tinham sido criadas listas com cada qualificier de cada feature, experimentalmente, utilizou-se a lista do Gene ID para efetuar essa comparação. Procurou-se obter o Gene ID da tabela, e colocar numa lista. Concluiu-se que o gene com Gene ID: 19831636 não se encontra na tabela de validação, pelo que não é validado.

Podemos ainda constatar que foram obtidas **432** features das quais **215** são genes (113 genes com funções desconhecidas e 101 genes com funções conhecidas) e **214** são CDS's. Das funções atribuídas aos genes podemos verificar que:

- 100 proteínas são **hipotéticas** - na bioquímica, uma proteína hipotética é aquela cuja existência foi prevista, mas para a qual não existe qualquer evidência experimental de que esta é expressa in vivo não tendo por isso função associada;
- A maioria dos genes presentes na porção do genoma estudado são de transporte e ligação (20), funções metabólicas (37) e ainda funções de degradação, transcrição e ligação de DNA e/ou RNA (23).

De forma a recolher alguma informação sobre a *Legionella pneumophila subsp. pneumophila str. Philadelphia 1*, criou-se um script que procura todos os artigos relacionados com a *legionella Pneumophila* na base de dados PubMed. Depois de recolhida essa informação, guardou-se num ficheiro `.txt`, o título do artigo, o autor e a companhia que publicou o artigo (Source).

Para facilitar a visualização das funções dos genes a informação foi distribuída sobre a forma de um mapa (baseado no do grupo 7), que se encontra no ficheiro `proteinfuction.py`. Pelo simples facto de alguns genes não terem funções associadas, procedeu-se a uma análise mais

aprofundada dos mesmos recorrendo a bases de dados como a Uniprot, NCBI, Swissprot, Prosite etc. Os resultados obtidos encontram-se no ficheiro func_genes.xls. Assim, foram encontrados registos de 18 proteínas traduzidas pelo nosso genoma. Os resultados obtidos estão resumidos na tabela 1.

Tabela 1 – Resultados obtidos por correlação com a base de dados da Uniprot e outras bases de dados

UniprotID	Nome da proteína	Função	Localização celular	Note
YP_094056.1	chromosomal replication initiator DnaA	Replication and Repair	Cytoplasm	ATP binding; cytoplasm; DNA replication initiation; DNA replication origin binding; regulation of DNA replication
YP_094060.1	peptidylarginine deiminase	Agmatine deiminase activity		Protein-arginine deiminase activity; putrescine biosynthetic process
YP_094064.1	host factor-I protein for bacteriophage Q beta replication	Signal transduction / other regulatory functions, Viral functions / Phage / Transposases		regulation of transcription, DNA-templated; RNA binding
YP_094077.1	23S rRNA m(2)G2445 methyltransferase	Nucleotide Metabolism	Cytoplasm	23S rRNA (guanine(2445)-N(2))-methyltransferase activity; cytoplasm; RNA binding; rRNA (guanine-N7-)-methyltransferase activity
YP_094080.1	palmitoyl transferase	Toxin production / other pathogen functions, Detoxification / adaptation	Cell outer membrane	cell outer membrane; integral component of membrane; lipid A biosynthetic process; O-palmitoyltransferase activity
YP_094083.1	ubiquinone biosynthesis protein COQ7	Energy Metabolism	Cell membrane; Peripheral membrane protein	metal ion binding; monooxygenase activity; plasma membrane; protein metabolic process; ubiquinone biosynthetic process
YP_094126.1	excinuclease ABC subunit B	DNA/RNA degradation / restriction	Cytoplasm	ATP binding; cytoplasm; DNA binding; excinuclease ABC activity; helicase activity; nucleotide-excision repair; SOS response

UniprotID	Nome da proteína	Função	Localização celular	Nota
YP_094148.1	ribose-5-phosphate isomerase A	Carbohydrate Metabolism, Energy Metabolism		pentose-phosphate shunt, non-oxidative branch; ribose-5-phosphate isomerase activity
YP_094154.1	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	metabolism of Complex Carbohydrates		lipid A biosynthetic process; transferase activity, transferring acyl groups other than amino-acyl groups
YP_094168.1	glycine dehydrogenase subunit 2	Amino Acid Metabolism		glycine decarboxylation via glycine cleavage system; glycine dehydrogenase (decarboxylating) activity; pyridoxal phosphate binding
YP_094170.1	glycine dehydrogenase subunit 1	Amino Acid Metabolism		glycine catabolic process; glycine dehydrogenase (decarboxylating) activity; nucleoside metabolic process; pyridoxal phosphate binding
YP_094171.1	glycine cleavage system protein H	Amino Acid Metabolism		glycine cleavage complex; glycine decarboxylation via glycine cleavage system
YP_094172.2	glycine cleavage system(GCS) aminomethyltransferase T	Amino Acid Metabolism, Metabolism of Cofactors and Vitamins, Energy Metabolism		aminomethyltransferase activity; cytoplasm; glycine catabolic process; transaminase activity
YP_094179.1	ribosome biogenesis GTP-binding protein YsxC	Signal transduction / other regulatory functions		barrier septum assembly; GTPase activity; GTP binding; magnesium ion binding

UniprotID	Nome da proteína	Função	Localização celular	Note
YP_094185.1	dihydrodipicolinate reductase	Amino Acid Metabolism	Cytoplasm	4-hydroxy-tetrahydrodipicolinate reductase; cytoplasm; diaminopimelate biosynthetic process; lysine biosynthetic process via diaminopimelate; NAD binding; NADPH binding; oxidoreductase activity, acting on CH or CH2 groups, NAD or NADP as acceptor
YP_094191.1	phosphoglycerate kinase	Carbohydrate Metabolism, Energy Metabolism	Cytoplasm	ATP binding; cytoplasm; glycolytic process; phosphoglycerate kinase activity
YP_094234.1	M48 family peptidase	Detoxification / adaptation, Protein fate / hydrolases / secretion		
YP_094248.1	catalase/(hydro)peroxidase KatG	Amino Acid Metabolism, Energy Metabolism		catalase activity; heme binding; hydrogen peroxide catabolic process; metal ion binding

De modo a obter uma análise mais aprofundada, numa primeira fase, procedeu-se ao estudo das **6** proteínas com atividade enzimática em cima encontradas (a azul).

A vida depende da realização de inúmeras reações químicas que ocorrem no interior das células e também fora delas (em cavidades de órgãos, por exemplo). Deste modo, é necessário que as reações químicas que a sustentam se deem a uma determinada velocidade. Muitas das reações bioquímicas não se realizariam a uma velocidade relativamente elevada se não fossem catalisadas. Em sistemas vivos, a catálise de reações químicas é feita por enzimas.

As enzimas são proteínas que, atuando como catalisadores na maioria das reações bioquímicas, baixam a energia de ativação necessária para que se dê uma reação química. Por serem catalisadores eficientes, são aproveitadas para aplicações industriais, como na indústria farmacêutica ou na alimentar. Como intervêm nas reações químicas que sustentam a vida, a compreensão do seu funcionamento é importante em áreas como a investigação de patologias com origem em deficiências enzimáticas.

A esmagadora maioria das reações bioquímicas dá-se em vias metabólicas, que são seqüências de reações em que o produto de uma reação é utilizado como reagente na reação seguinte. Diferentes enzimas catalisam diferentes passos de vias metabólicas, agindo de forma concentrada de modo a não interromper o fluxo nessas vias. Por outro lado, cada enzima pode sofrer regulação da sua atividade, aumentando-a, diminuindo-a ou mesmo interrompendo-a. Assim, o fluxo de uma via metabólica depende da velocidade de catálise das enzimas que nela participam.

A International Union of Biochemistry and Molecular Biology (IUBMB) classificou as enzimas em seis grandes grupos (Classes), de acordo com o tipo de reação que catalisam. Cada enzima descrita recebe um número de classificação, conhecido por "E.C." (Enzyme Commission of the IUBMB), que é composto por 4 dígitos:

1. Classe
2. Subclasse dentro da classe
3. Grupos químicos específicos que participam na reação.
4. A enzima, propriamente dita

Através da base de dados KEGG, Metacyc entre outras recolheu-se a seguinte informação:

- [Ribose 5-phosphate isomerase A](#)

E.C: **5.3.1.6**

5. Isomerases

5.3 Oxidorredutases intramoleculares

5.3.1 Interconversão de aldoses e cetoses, e compostos relacionados

5.3.1.6 Ribose-5-fosfato isomerase

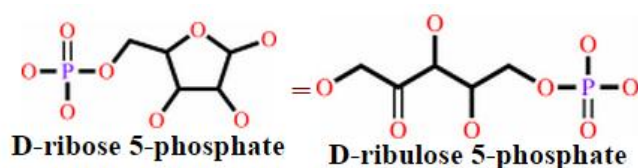


Figura 2 – Ação da Ribose-5-fosfato isomerase

Classe: As isomerases catalisam alterações geométricas e/ou estruturais no interior de uma molécula. De acordo com o tipo de isomerismo, podem ser chamadas epimerases, cis-trans-isomerases, isomerases, tautomerases, mutases entre outras.

Subclasse: As oxidorredutases intramoleculares catalisam as reações de oxidação-redução, oxidando uma parte da molécula ocorrendo a correspondente redução na outra parte. Incluem as enzimas que convertem, na série dos açúcares, aldoses em cetoses e vice-versa.

A Ribose-5-fosfato isomerase desempenha um papel vital no metabolismo bioquímico, tanto na via das pentose-fosfato como no ciclo de Calvin. Esta enzima catalisa a transformação **reversível** de Ribose-5-fosfato em Ribulose-5-fosfato (figura 2), um intermediário de extrema importância na **Via das pentoses-fosfato** (figura 3). A via das pentoses-fosfato é um dos 3 destinos possíveis para a glicose-6-fosfato (para além da glicogénese e via glicolítica). É uma via alternativa da oxidação da glicose-6-fosfato que origina a produção de 3 compostos: a ribose-5-fosfato, CO₂ e NADPH. A Ribose-5-fosfato é um substrato de vias metabólicas que levam à formação de nucleótidos e de ácidos nucleicos, sendo ainda a pentose constituinte de muitas coenzimas como o ATP, NADH, FADH₂ e coenzima A.

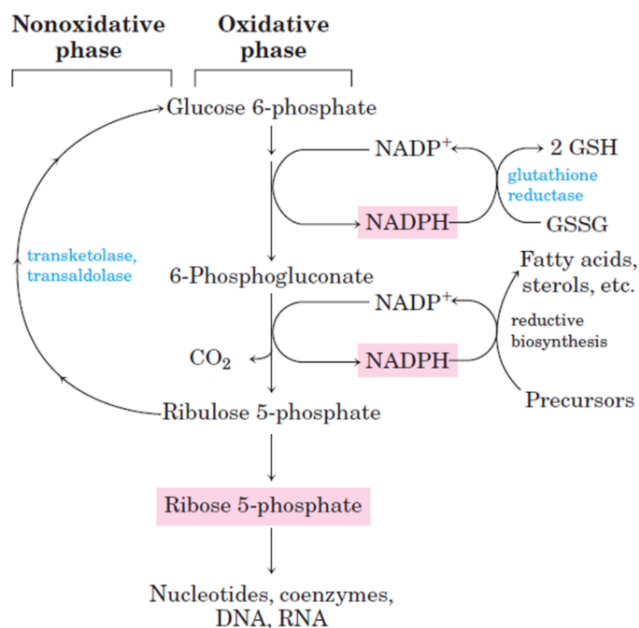


Figura 3 – Via pentoses-fosfato

A via das pentoses é uma via citoplasmática, anaeróbica ocorrendo no fígado, glândulas mamárias, tecido adiposo e nas hemácias. Ocorre em duas etapas: **fase oxidativa**, onde ocorre a produção de pentoses, sendo esta uma importante fonte de NADPH que é necessário para reações biossintéticas e proteção contra espécies reativas de oxigénio e a **fase não oxidativa** onde ocorre a interconversão de pentoses intermediárias da via glicolítica.

O metabolismo da glucose na via das pentoses fosfato influencia o desenvolvimento de diversas patologias. A deficiência em G6PD (glucose-6-fosfatosdesidrogenase) é o defeito genético mais comum na população humana, levando a níveis reduzidos de NADPH nos eritrócitos e predispondo ao aparecimento de anemia hemolítica induzida por stress oxidativo. Recentemente, foram descritas duas novas deficiências enzimáticas na via das pentoses fosfato: a deficiência em ribose-5-fosfato isomerase (RPI) e a deficiência em transaldolase (TALDO). A primeira deficiência enzimática está associada a leucoencefalopatia progressiva e ligeira neuropatia periférica. O perfil metabólico da doença revela níveis elevados dos polióis: ribitol e arabitol no cérebro, tal como demonstrado por ressonância magnética, na urina, plasma e líquido cefalorraquidiano (LCR). Por outro lado, a deficiência em transaldolase está associada a alterações hepáticas, havendo um grau variado de comprometimento por parte de outros órgãos.

No ciclo de Calvin, a energia a partir dos transportadores de eletrões é utilizada na fixação de carbono, na conversão de dióxido de carbono em água e hidratos de carbono. Esta enzima é essencial no ciclo uma vez que a Ribulose 5-fosfato, gerada a partir da Ribose 5-fosfato, é subsequentemente convertida em Ribulose bifosfato, o aceitador de dióxido de carbono na primeira reação da fase química da fotossíntese. É formado um composto intermédio instável de seis carbonos, que rapidamente forma duas moléculas com três carbonos – ácido fosfoglicérico ou PGA. Estas reações de fixação de CO₂ são catalisadas pela enzima ribulose difosfato carboxilase-oxidase (RuBisCo). As moléculas de PGA são fosforiladas pelo ATP e posteriormente reduzidas pelo NADPH proveniente da fase foto-dependente, formando o

aldeído fosfoglicérico (PGAL). As reações seguintes do ciclo têm como objetivo produzir mais RuDP e moléculas orgânicas mais complexas, como a glicose. Por cada 12 moléculas de PGAL formadas, 10 serão utilizadas para regenerar RuDP e as duas restantes para sintetizar compostos orgânicos mais complexos (glicose e outros glícidos). O PGAL pode também ser convertido noutros compostos orgânicos como lípidos (glicerol e ácidos gordos) ou prótidos (aminoácidos).

- [Glicina desidrogenase \(subunidade 1 e 2\)](#)

E.C: **1.4.4.2**

1. Oxidorredutases

1.4 Atuam nos grupos CH-NH₂ (dadores de hidrogénio)

1.4.4 Com o dissulfeto como recetor

1.4.4.2 Glicina desidrogenase

Classe: As oxidorredutases catalisam reações de oxidação-redução. O substrato oxidado é considerado um dador de hidrogénio ou de eletrões.

A classificação EC para estas enzimas é feita em primeiro lugar para o grupo dador que sofre oxidação, seguidamente pelo recetor, e, finalmente, pela enzima. O nome recomendado para estas enzimas é desidrogenase ou redutase. O nome oxidase só é utilizado quando o oxigénio é o aceitador.

Subclasse: São as desidrogenases de aminoácidos e as oxidases de amina. Na maioria dos casos, a imina formada é hidrolisada para originar um grupo oxo e –NH₃ (desaminação). As subclasses são formadas de acordo com o aceitador, neste caso com o dissulfeto.

A glicina desidrogenase é uma enzima dependente de fosfato de piridoxal que catalisa a descarboxilação da glicina (figura 3) com a transferência de um grupo amino-metil para o resíduo de ácido lipoico da proteína-h do complexo glicina descarboxilase. É uma das quatro subunidades do **complexo glicina descarboxilase** (Complexo enzimático que catalisa a descarboxilação oxidativa e a desaminação da glicina em dióxido de carbono, amónia, NADH e N5N10-metilenotetrahidrofolato. É composto por quatro componentes diferentes citados como componentes proteicos H, P, L e T).

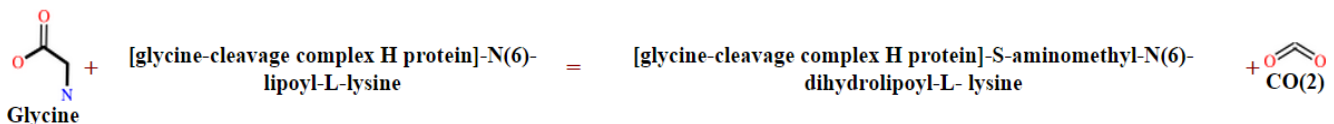


Figura 4 – Descarboxilação da glicina

- [GCS – Sistema de clivagem da glicina \(ou complexo glicina descarboxilase\)](#)

E.C: **2.1.2.10**

2. Transferases

2.1 Transferência de grupos carbono

2.1.2 Hidroximetil-, formil- e transferases relacionadas

2.1.2.10 aminometiltransferase

Classe: As transferases são enzimas que transferem um grupo (por exemplo, um grupo metilo ou glicosil) a partir de um composto (que é geralmente considerado como dador) para um outro composto (recetor). A classificação EC para estas enzimas é realizada em primeiro lugar para o grupo dador, numa segunda fase para o recetor, e, finalmente, para a enzima.

Subclasse: Esta subclasse contém hidroximetil-, formil- e transferases relacionadas

O sistema de clivagem da glicina é um complexo multienzimático que catalisa a oxidação reversível de glicina, originando dióxido de carbono, amoníaco, 5,10- metilenotetrahidrofolato e um nucleótido de piridina reduzido (figura 5). O tetrahydrofolato serve como um recetor para as unidades de um carbono gerado durante a clivagem de glicina para formar o grupo metileno. O sistema consiste em quatro componentes de proteínas: a proteína P, proteína H, proteína T, G e L. Proteína P catalisa a libertação de piridoxal-fosfato dependente de CO₂ a partir de glicina, deixando uma porção de metilamina. A porção de metilamina é transferida para o grupo do ácido lipóico da proteína H, que está ligado à proteína P antes da descarboxilação de glicina. A proteína T catalisa a libertação de NH₃ a partir do grupo de metilamina e transfere a unidade C1 restante de THF, formando-MTHF 5,10. A proteína G, em seguida, oxida o ácido lipóico da proteína H e os transfere eletrões para o NAD⁺, formando NADH.

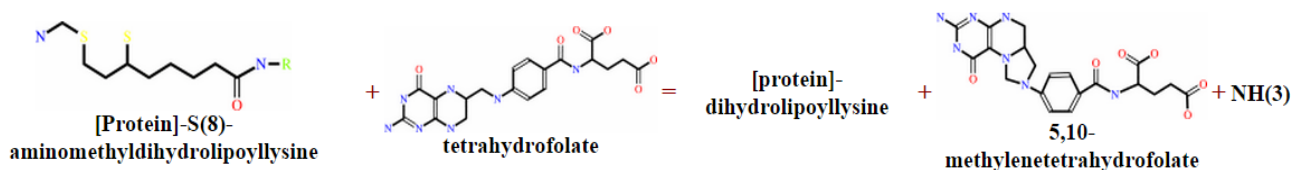


Figura 5 – Ação do complexo glicina descarboxilase

- [Fosfoglicerato cinase](#)

E.C: **2.7.2.3**

2. Transferases

2.7 Transferência de grupos contendo fosfato

2.7.2 Fosfotransferases com grupo carboxilo como recetor

2.7.2.3 fosfoglicerato cinase

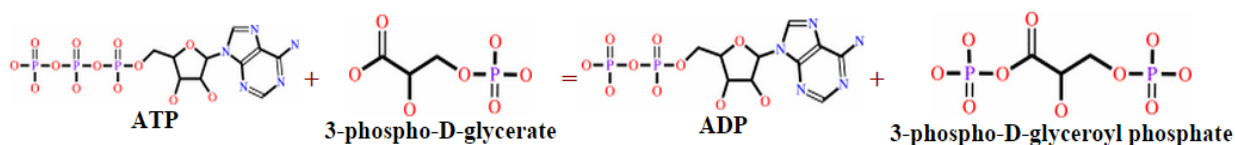


Figura 6 – Fosforilação do 1,3-Bisfosfoglicerato

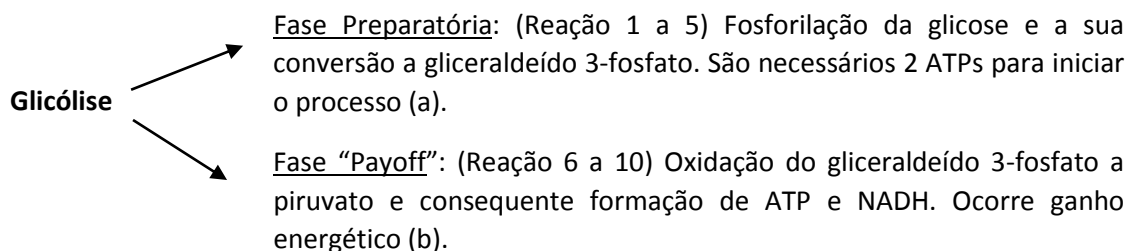
Classe: transferases

Subclasse: Este é um grande grupo de enzimas que compreendem não só aquelas que transferem fosfato mas também difosfato, resíduos de nucleótidos entre outros.

A fosfoglicerato cinase promove a transferência de um grupo fosforilo do 1,3- Bisfosfoglicerato (figura 5), interveniente na **oxidação metabólica das hexoses**, para o ADP. A enzima fosfoglicerato cinase transfere o grupo fosforilo de alta energia do grupo carboxilo do 1,3- Bisfosfoglicerato para o ADP formando ATP e 3-fosfoglicerato. Nesta reação ocorre uma fosforilação ao nível do substrato porque a energia libertada na hidrólise do grupo fosforilo é imediatamente utilizada na síntese de um ATP, a partir do ADP. O gene que codifica esta enzima (PGK) está localizado no cromossoma X. Uma deficiência nesta enzima é uma causa comum de anemia hemolítica crónica e mioglobinúria.

As hexoses são monossacarídeos de 6 carbonos, que obedecem à fórmula geral – $C_nH_{2n}O_n$ ($n=6$). As hexoses mais importantes são a glicose, a frutose e a galactose, principais fontes de energia para os seres vivos. Ricas em energia, as hexoses constituem os principais combustíveis das células. São naturalmente sintetizadas por fotossíntese, processo de absorção de energia da luz.

A principal via de oxidação metabólica das hexoses é a **glicólise** (figura 7).



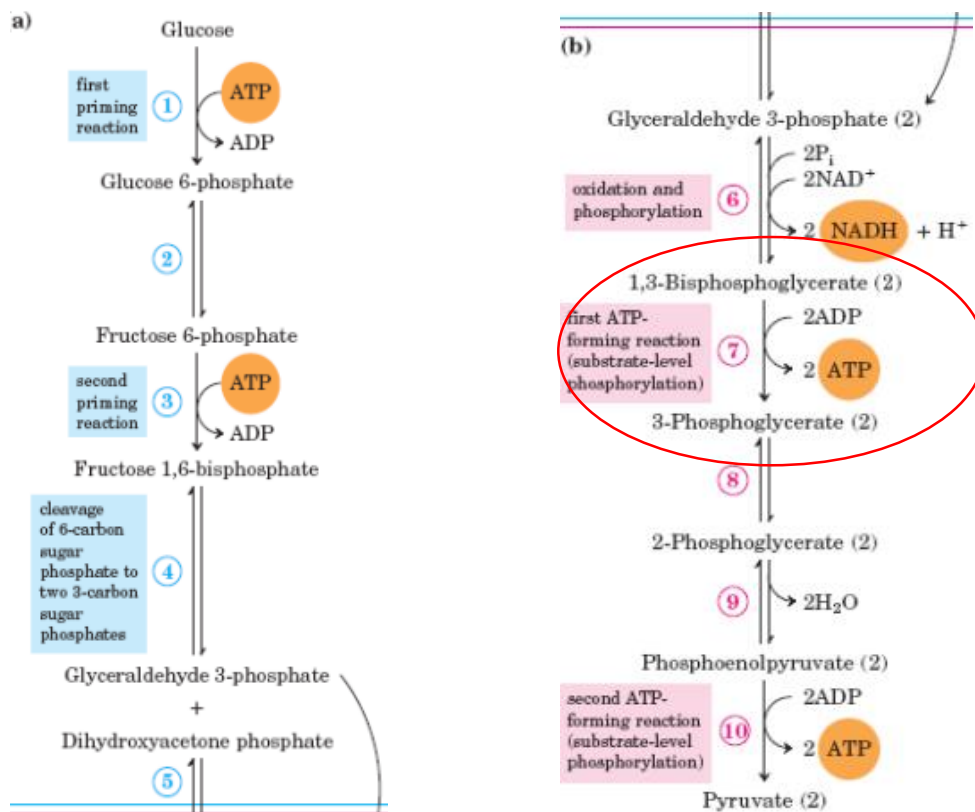


Figura 7 – Glicólise

Na maior parte dos organismos, outras hexoses como a frutose, a manose e a galactose, podem participar na glicólise, após conversão para o derivado fosforilado.

- [Catalase/\(hidro\)peroxidase \(KatG\)](#)

E.C: **1.11.1.6**

1. Oxidorredutases
- 1.11 Peróxido como recetor de eletrões
- 1.11.1 Peroxidases
- 1.11.1.6 Catalase

Classe: Oxidorredutases

Subclasse: Esta simples subclasse contém as peroxidases

A catálase (normalmente denominada hidropoxidase) é uma enzima intracelular, encontrada na maioria dos organismos, que converte o **peróxido de hidrogénio** (H₂O₂) em água segundo a reação química em baixo descrita.

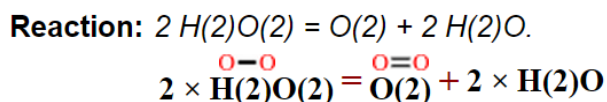


Figura 8 – Decomposição do peróxido de hidrogénio

A catalase é uma enzima que está relacionada com a dissipação de radicais livres dentro das células. Está normalmente localizada num organelo celular (peroxissoma), sendo também encontrada nas mitocôndrias (mamíferos) sendo esta enzima homotetramérica de 230.813 kDa que contém uma molécula de NADPH e um grupo heme por subunidade (hemoproteína). Esta enzima por decompor o peróxido de hidrogénio resultante do metabolismo celular normal cuja ação oxidativa é tóxica para algumas células como os eritrócitos. Assim torna-se necessário a sua degradação para evitar danos que comprometam a viabilidade celular.

A acatalasemia é uma deficiência genética de carácter autossómica recessiva, que interfere no metabolismo da catálase no sangue (acatalasemia) e nos tecidos (acatalasia).

Depois de consultada a base de dados **KEGG** (Kyoto Encyclopedia of Genes and Genomes) e de modo a construir um mapa metabólico teve-se em especial atenção 4 enzimas:

KEGG ORTHOLOGY	NAME	KEGG PATHWAY	KEGG MODULE	REACTION	DESEASE
K01807	Ribose-5-phosphate isomerase A	Ko00030 Pentose phosphate pathway	M00004 Pentose phosphate pathway (Pentose phosphate cycle)	R01056	H01135 Ribose 5-phosphate isomerase (RPI) deficiency
		Ko00710 Carbon fixation in photosynthetic organisms	M00007 Pentose phosphate pathway, non-oxidative phase, fructose 6P => ribose 5P		
		Ko01200 Carbon metabolism	M00165 Reductive pentose phosphate cycle (Calvin cycle)		
		Ko01230 Biosynthesis of amino acids	M00167 Reductive pentose phosphate cycle, glyceraldehyde-3P => ribulose-5P		
			M00580 Pentose phosphate pathway, archaea, fructose 6P => ribose 5P		
K00927	phosphoglycerate kinase	ko00010 Glycolysis / Gluconeogenesis	M00001 Glycolysis (Embden-Meyerhof pathway), glucose => pyruvate	R01512	H00654 Anemia due to disorders of glycolytic enzymes
		koo0710 Carbon fixation in photosynthetic organisms	M00002 Glycolysis, core module involving three-carbon compounds		

		ko01200 Carbon metabolism	M00003 Gluconeogenesis, oxaloacetate => fructose-6P		
		ko01230 Biosynthesis of amino acids	M00165 Reductive pentose phosphate cycle (Calvin cycle) M00166 Reductive pentose phosphate cycle, ribulose-5P => glyceraldehyde-3P M00308 Semi-phosphorylative Entner-Doudoroff pathway, gluconate => glycerate-3P M00552 D-galactonate degradation, De Ley-Doudoroff pathway, D-galactonate => glycerate-3P		
K03781	catalase/(hydro) peroxidase KatG	ko00380 Tryptophan metabolism ko00630 Glyoxylate and dicarboxylate metabolism ko04068FoxO signaling pathway ko04146 Peroxisome koo5014 Amyotrophic lateral sclerosis (ALS)	M00532 Photorespiration	R00009	H00203 Acatlasia
K00215	4-hydroxy-tetrahydrodipicolinate reductase	ko00300 Lysine biosynthesis ko01230 Biosynthesis of amino acids	M00016 Lysine biosynthesis, succinyl-DAP pathway, aspartate => lysine M00525 Lysine biosynthesis, acetyl-DAP pathway, aspartate => lysine M00526 Lysine biosynthesis, DAP dehydrogenase pathway, aspartate => lysine M00527 Lysine biosynthesis, DAP aminotransferase pathway, aspartate => lysine	R04198 R04199	

Legenda:

KEGG orthology → A base de dados KEGG ortologia (KO) é um conjunto de grupos de ortólogos que correspondem aos nós (caixas) dos mapas KEGG (pathway KEGG).

KEGG module → Os módulos KEGG contêm esta entrada dos grupos ortólogos. Ao clicar no link, o objeto retangular correspondente é marcado a **vermelho** no mapa do módulo (figura 8).

KEGG reaction → Todas as reações em que o número E.C está envolvido

O mapa metabólico construído é apresentado na figura 9.

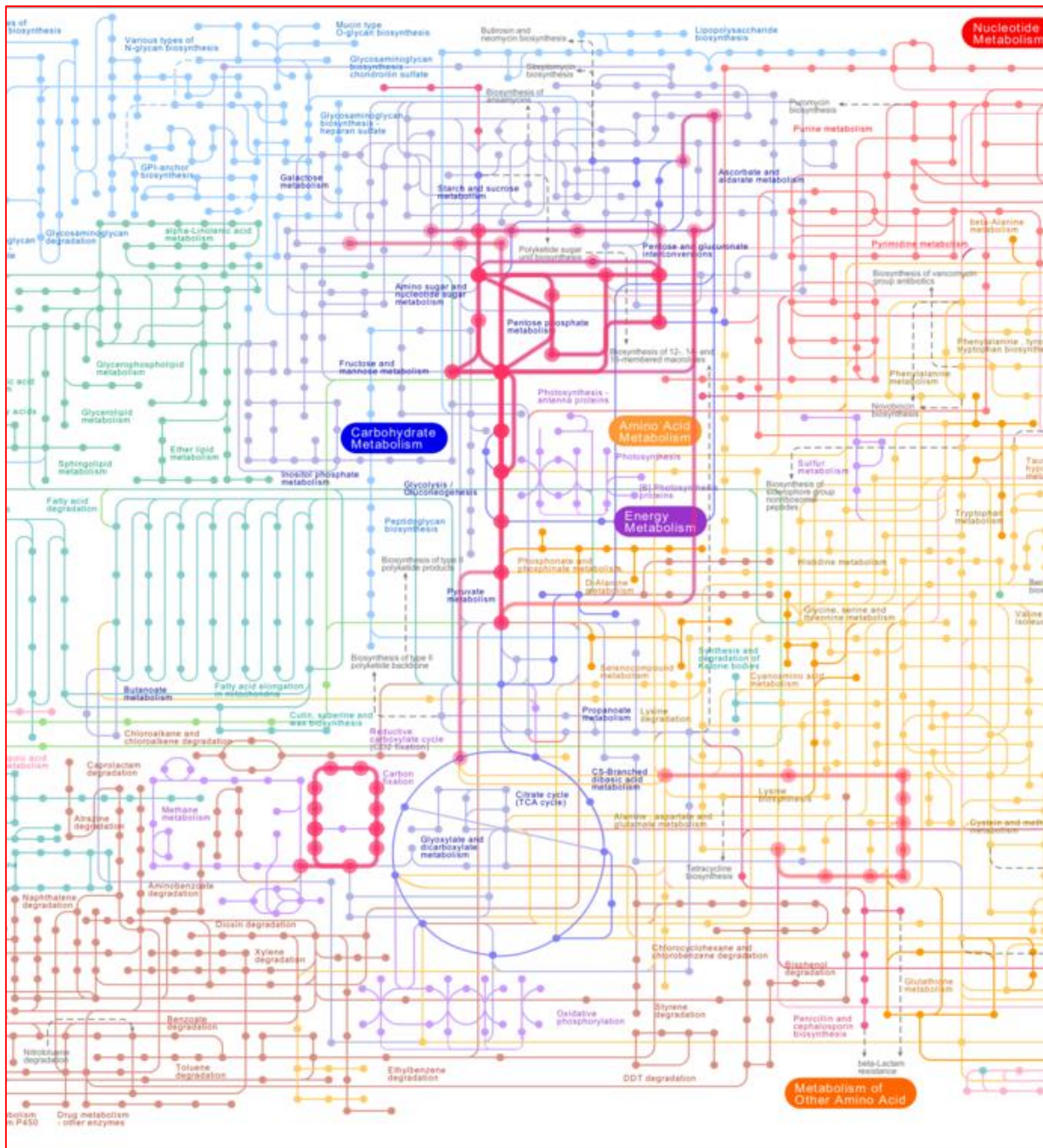


Figura 9 – Mapa metabólico

Numa segunda fase procedeu-se a análise das seguintes proteínas existentes na tabela 1.

- **Iniciador da replicação cromossomal DnaA**

Liga-se ao “dnaA-box” como um complexo ligando de ATP na origem da replicação durante a iniciação da replicação cromossómica; Também pode afetar a transcrição de vários genes, incluindo ela própria.

- **Peptidil arginina desaminase**

Através da base de dados PATRIC foi feita a correlação entre esta proteína e a agmatina desaminase (EC 3.5.3.12). A enzima também catalisa as reações de EC 2.1.3.3 (carbamoiltransferase ornithine), EC 2.1.3.6 (carbamoiltransferase putrescina) e EC 2.7.2.2 (carbamato quinase), funcionando, assim, como uma putrescina sintase, convertendo agmatina e ornitina em putrescina e citrulina, respetivamente. Intervém no metabolismo de aminoácidos como a arginina e a prolina.

- **Proteína Hfq**

A proteína Hfq (também conhecida como proteína HF-I), codificada pelo gene HFQ foi descoberta em 1968 sendo um este um fator hospedeiro da Escherichia Coli que foi essencial para a replicação do bacteriófago Q ϕ . É agora sabido que a Hfq é uma abundante proteína de ligação ao RNA bacteriano que tem muitas funções fisiológicas importantes (transdução de sinais, funções regulatórias, regulação da transcrição) que são geralmente mediadas por interação desta proteína com o sRNA.

- **23S rRNA m(2)G2445 metiltransferase**

Esta proteína está envolvida no metabolismo de aminoácidos, nomeadamente, catalisa a modificação da N2-metil guanosina do resíduo G2445 da 23S rRNA.

- **Palmitoil transferase**

Esta proteína catalisa a transferência de palmitato para o lípido A. Intervém na produção de toxinas e outras funções patogénicas. Representa um componente integrante da membrana celular externa e está também envolvida na biossíntese de lípidos.

- **COQ7 (proteína envolvida na síntese de ubiquinona)**

Proteína muito importante no metabolismo energético. É uma proteína de membrana, sendo uma proteína periférica. Participa na ligação a iões metálicos e na síntese de ubiquinona.

- **Excinuclease ABC (subunidade B- uvrB)**

A reparação por excisão de nucleótidos (NER, do inglês *nucleotid excision repair*) sinaliza lesões amplas e volumosas e retira porções de DNA de ambos os lados. Tal atividade enzimática é denominada excinuclease. O complexo uvrA reconhece a lesão enquanto em associação com o complexo uvrB que também está envolvido no desenrolamento da molécula de DNA. Posteriormente o complexo uvrC é responsável pelas incisões na molécula de DNA e o complexo uvrD (helicase II) é responsável pela eliminação da lesão.

- **UDP-3-O-[3-hydroximiristoil] glucosamina-N-aciltransferase**

Esta proteína apresenta atividade enzimática (EC 2.3.1.-) estando envolvida na transferência de grupos acilo para outros grupos amino-acil. Está envolvida no metabolismo de carboidratos e aminoácidos.

- **Proteína ligando de GTP envolvida na biogénese de ribossomas**

Esta proteína está envolvida na transdução de sinal e outras funções reguladoras. Atividade de GTPase e de ligação ao íon magnésio.

- **Dihidrodipicolinato redutase**

Também conhecida como 4-hidroxi-tetrahidrodipicolinato (EC 1.3.1.26) esta enzima pertence à família das oxidoredutases, especificamente aquelas que atuam sobre o grupo CH-CH de dadores, sendo NAD⁺/ NADP⁺ o recetor. Está envolvida no metabolismo de aminoácidos, nomeadamente no processo de biossíntese de lisina via diaminopimelato.

- Virulência e resistência a antibióticos

- Depois de analisada a base de dados PATRIC (Pathosystems Resource Integration Center) recolheu-se a seguinte informação genérica acerca da *Legionella pneumophila* subsp. *pneumophila* str. Philadelphia 1

Specialty Gene Summary		
	Source	Genes
Antibiotic Resistance	CARD ▶	22
Drug Target	DrugBank ▶	4
Drug Target	TTD ▶	2
Human Homolog	Human ▶	48
Virulence Factor	VFDB ▶	116
Virulence Factor	Victors ▶	28

Deu-se especial atenção à patogenicidade e à resistência a antibióticos uma vez que apenas aqui se encontram **3** dos genes correspondentes à porção de genoma analisada [lpg0001 a lpg0215]. Os dados obtidos encontram-se no ficheiro SpecialtyGene.xlsx

Property	Source	RefSeq Locus Tag	Gene	Classification	Function
Virulence Factor	VFDB	lpg0103	N-terminal acetyltransferase, GNAT family (vipF)	Secretion, Type IV secretion system	Allows survival and growth in macrophages, prevent phagosome acidification and lysosome fusion, essential for induction of apoptosis in human macrophages
	VFDB	lpg0194	Catalase (EC 1.11.1.6) / Peroxidase (EC 1.11.1.7)	Stress protein	A periplasmic catalase, expressed maximally during the postexponential phase; important for intracellular survival and transmission
Antibiotic Resistance	CARD	lpg0047	Chloramphenicol acetyltransferase (EC 2.3.1.28)	ARO:1000001 ARO:3000122	Chloramphenicol O-acetyltransferase (CAT) catalyzes the acetyl-CoA dependent acetylation of chloramphenicol (Cm), an antibiotic which inhibits prokaryotic peptidyltransferase activity

Através da base de dados VFDB (virulence factors data base) conseguiu-se a seguinte informação acerca do gene **N-terminal acetiltransferase** (GNAT family):

Esta proteína participa como substrato no **sistema de secreção Dot/Icm tipo IV (figura 10)** da *Legionella*.

O maior sistema de virulência da *Legionella pneumophila* é um sistema de secreção especializado codificado por 26 dot/icm (defect in organella trafficking/intracellular multiplication) genes. Este sistema é classificado como tipo IV devido à sua homologia com o aparelho de transferência de plasmídeos conjugativos de classe InCl, uma vez que 14 genes compartilham uma homologia detetável com os genes tra / trb do plasmídeo col1b-P9, um membro da classe InCl de plasmídeos conjugais.

Na *Legionella*, o sistema de secreção Dot / Icm é usado para exportar um grande número de substratos de proteína no citoplasma das células fagocíticas, permitindo assim que a *Legionella* possa sobreviver e replicar dentro destas células.

Funções:

Permite a sobrevivência e crescimento em macrófagos, prevenindo a acidificação do fagossoma e a sua fusão com lisossomas, sendo essencial para a indução de apoptose em macrófagos humanos.

Mecanismo:

Montar e ativar um sistema de secreção do tipo IV na membrana bacteriana que exporta o plasmídeo e de fatores de virulência para inibir a fusão fago-lisossoma e reprogramar o vacúolo “*Legionella-bearing*”.

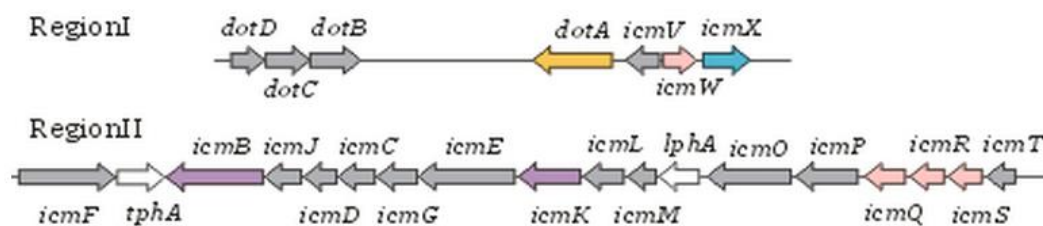


Figura 10 – Loci dot/icm localizado em duas regiões distintas. Reproduzido de: Vogel JP, Isberg RR, 1999. Cell biology of *Legionella pneumophila*. Curr. Opin. Microbiol. 2(1):30-34.)

Através da base de dados CARD (comprehensive antibiotic resistance database), PROSITE, entre outras analisou-se mais detalhadamente a atividade da cloranfenicol acetiltransferase.

2. Transferases

2.3 Aciltransferases

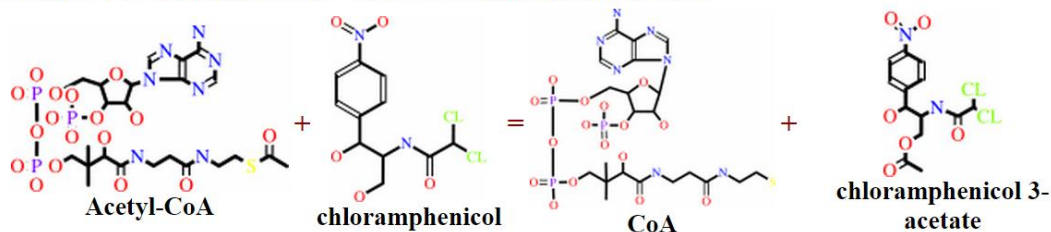
2.3.1 Transferência de grupos a partir de outros grupos aminoacil

2.3.1.28 cloranfenicol O-acetiltransferase

Classe: transferases

Subclasse: estas enzimas transferem grupos acilo formando ésteres ou amidas. O dador é na maioria dos casos um derivado de acetil-CoA.

Reaction: *Acetyl-CoA + chloramphenicol = CoA + chloramphenicol 3-acetate.*



O cloranfenicol é um antibiótico inibidor da síntese proteica bacteriana com efeito bacteriostático e bactericida. Exerce a sua atividade ligando-se reversivelmente à peptidil-transferase da subunidade 50S do ribossoma impedindo a ligação do tRNA, evitando a codificação do novo aminoácido. O efeito bactericida acontece na presença de concentrações elevadas de antibióticos. A razão para a sua utilização ser limitada é que, para além de interferir na síntese de proteínas bacterianas, este pode interromper a síntese de proteínas nas células da medula óssea humana e produzir discrasias sanguíneas, como anemia aplásica.

A produção de enzimas é o mecanismo de resistência ao cloranfenicol mais usado pelas bactérias. O gene CAT adquirido pela bactéria via plasmídeo é o responsável pela produção da enzima cloranfenicol-acetiltransferase responsável pela hidrolisação do grupo amida e modificação dos grupos hidroxilos do fármaco, inativando-o.

A **cloranfenicol acetiltransferase** (CAT), é uma enzima bacteriana que desintoxica o antibiótico cloranfenicol e é responsável pela resistência ao cloranfenicol em bactérias. Esta enzima atribui covalentemente um grupo acetilo a partir da acetil-CoA ao cloranfenicol o que impede que o cloranfenicol se ligue aos ribossomas. A acetilação do Cm pela CAT inativa o antibiótico. Um resíduo de histidina, localizada na parte C-terminal da enzima, desempenha um papel central no mecanismo catalítico.