UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

ELEMENTOS REPETITIVOS NA REGULAÇÃO DA TRANSCRIÇÃO DE *Mycoplasma hyopneumoniae*

Dissertação de Mestrado

Amanda Malvessi Cattani

Porto Alegre, março de 2016

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

ELEMENTOS REPETITIVOS NA REGULAÇÃO DA TRANSCRIÇÃO DE *Mycoplasma hyopneumoniae*

Dissertação submetida ao Programa de Pós-Graduação em Biologia Celular e Molecular do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Mestre.

Amanda Malvessi Cattani

Orientadora: Prof.ª Dra. Irene Silveira Schrank
Co-orientadora: Dra. Franciele Maboni Siqueira

Porto Alegre, março de 2016

Este trabalho foi realizado no Laboratório de Micro-organismos Diazotróficos do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul. Recebeu auxilio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

AGRADECIMENTOS

À minha orientadora Prof.ª Dra. Irene Silveira Schrank, pelo acolhimento em seu laboratório desde a iniciação científica, pela sempre disponibilidade em discussões enriquecedoras, fundamentais para o bom andamento do trabalho, além de todos os ensinamentos transmitidos ao longo desses anos.

À minha co-orientadora, Dra. Franciele Maboni Siqueira, pela amizade, carinho e palavras de incentivo sempre que necessário. Agradeço também todos os ensinamentos lapidados pouco a pouco desde a graduação.

Aos colegas do laboratório 217 e 220 pela gentil colaboração diária.

A todos os professores do Centro de Biotecnologia da UFRGS, especialmente aos professores: Augusto Schrank, Arnaldo Zaha, Charley Staats, Henrique Ferreira e Marilene Veinstein por todos os ensinamentos e conselhos.

Á minha amiga e colega Laura, pela amizade e parceria, agradeço também pela sua disponibilidade de ajuda aos colegas em quaisquer situações (inclusive na confecção de tabelas), além das caronas que sempre facilitaram muito minha vida. Agradeço ao Tiago, amigo, colega e companheiro nos intervalos para o Café, sempre disposto a uma boa conversa sobre temas diversos, que tornavam o dia mais leve. As amigas e colegas Gabriela e Camila, agradeço por toda a ajuda na confecção das tabelas desse trabalho, que não foram nada simples, pela amizade e parceria além da constante ajuda ao longo de todo o trabalho. A todos os colegas do laboratório 222 agradeço muito pelo convívio diário, amizade, carinho e atenção.

Aos meus pais, pelo apoio e amor incondicional em todos os momentos. Por estarem sempre disponíveis para uma palavra de incentivo, mesmo nas situações mais difíceis. Agradeço também por me propiciarem um estudo de qualidade e o desejo de um futuro melhor. Sem vocês, essa e outras conquistas não seriam possíveis. Agradeço ao meu irmão, pelo também apoio e amor em todos os momentos e principalmente a constante valorização dos estudos executados.

Ao meu namorado, pela paciência e compreensão nos intermináveis dias de estudo, além do apoio e carinho que sempre demonstrou ao longo de todos esses anos.

As minhas amigas Cristine, Luiza, Mayara e Thaiane pela amizade, companheirismo, compreensão e confidencias, além da boa companhia em noites de risadas e brincadeiras que tornaram a pós-graduação bem mais leve.

Ao LNCC, especialmente ao Rafael Guedes, por todos os scripts desenvolvidos, que facilitaram muito a análise dos dados brutos.

À Universidade Federal do Rio Grande do Sul pelo ensino de qualidade. Ao CNPq e a CAPES pelo financiamento do trabalho.

ÍNDICE

LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES

ºC - graus Celsius
A – adenina
ATP – adenosina Trifosfato
BLAST – do inlgês, *basic local alignment search tool*
C – citosina
cDNA – DNA complementar
CDS – sequência codificadora (do inglês, *coding sequence*)
CT – do inglês, *threshold cycle*
DNA - ácido desoxirribonucleico (do inglês, *desoxyribunucleic acid*)
G – guanina
GTP – guanosina trifosfato
Kb - quilobase (1000 pares de bases)
kDa – quilodalton (1000 daltons)
LPS – lipopolissacarídeo
MFL – *Mycoplasma flocculare*
MHP_232 - *Mycoplasma hyopneumoniae* linhagem 232
MHP_7448 – *Mycoplasma hyopneumoniae* linhagem 7448
MHP_J - *Mycoplasma hyopneumoniae* linhagem J
min – minutos
mRNA - RNA mensageiro
PAL – repetição palindrome
PALG – repetição palindrome possuindo espaçadores
pb - par de bases
PCR - reação em cadeia da polimerase (do inglês *polymerase chain reaction*)
qPCR – PCR quantitativo
RNA – ácido ribonucleico (do inglês, *ribunucleic acid*)
rRNA – RNA ribossômico
RT – trancrição reversa (do inglês, *reverse transcription*)
s – segundos
SSR – repetição de sequência simples (do inglês, *simple sequence repeat*)

SSRM - repetição de sequência simples de mononuleotideo (do inglês, *simple sequence repeat of mononucleotide*)

T - timina

Tm - temperatura de fusão

tRNA - RNA transportador

RESUMO

Mycoplasma hyopneumoniae é uma bactéria de tamanho diminuto, caracterizada por um genoma pequeno, com baixo conteúdo GC. Está associada com doenças respiratórias de suínos, resultando em prejuízos produtivos e econômicos na indústria animal. A presença de sequências de DNA repetitivas, que ocorrem em grandes quantidades em células eucarióticas, vem sendo cada vez mais identificadas em genomas de procariotos, sendo também associadas a um potencial papel regulador. Uma vez que a regulação da transcrição nesses organismos ainda é pouco entendida, o objetivo do presente estudo foi realizar uma busca *in silico* por elementos repetitivos nas regiões intergênicas do genoma de *M. hyopneumoniae* linhagem 7448. Dois tipos de repetições foram selecionados para a busca inicial: tandem e palindromes. Regiões intergênicas de até 500 pb a montante do sítio de início da tradução de todas as CDSs do genoma de *M. hyopneumoniae* linhagem 7448 foram utilizadas para a predição. Para cada tipo de elemento dois programas computacionais independentes foram utilizados. As predições *in silico* resultaram em 144 repetições em tandem e 1.171 palindromes. O DNA repetitivo se encontra distribuído a montante de 86% das unidades transcricionais de *M. hyopneumoniae* linhagem 7448. Análises comparativas entre genomas de micoplasmas demonstraram diferentes níveis de conservação dos elementos repetitivos entre linhagens patogênicas e não-patogênicas. Linhagens patogênicas revelaram uma conservação de 59%, enquanto que a não patogênica, somente de 46%. Através de ensaios de amplificação quantitativa de DNA, foi observado diferentes níveis de expressão em genes codificantes para importantes proteínas, como glicina hidroximetiltransferase, lipoproteína, adesinas e proteína ligadora de GTP. Os genes codificantes para essas proteínas divergiam no número de repetições palindromes e tandens na sua respectiva região intergênica. Além disso, repetições encontradas em 206 genes já descritos como regulados em diferentes condições em *M. hyopneumoniae* linhagem 232 mostraram aproximadamente 80% de conservação em relação à linhagem *M. hyopneumoniae* linhagem 7448. Todos esses resultados sugerem um potencial papel regulador das repetições de DNA em tandem e palindromes em *Mycoplasma*.

Palavras-chave: Tandem, Palindrome, DNA repetido, Regulação da transcrição

ABSTRACT

Mycoplasma hyopneumoniae is a diminutive bacterium, characterized by a small genome with a low GC content. It is commonly associated with swine respiratory diseases, resulting in productivity and economic losses in the animal industry. Repetitive DNA, which occurs in large quantities in eukaryotic cells, has been increasingly identified in prokaryotic genomes, and has been associated with a potential regulatory function. Once transcription regulation in these organisms is still poorly understood, the aim of the current study was to perform an *in silico* search of repeat elements in the genomic intergenic regions of *M. hyopneumoniae* strain 7448. Two types of repeats were selected for initial search: Tandem and Palindromic. Intergenic regions up to 500 bp upstream from start codon of *M. hyopneumoniae* strain 7448 CDSs were used as input for the software's prediction. For each type of repeat sequence, two independent software packages were used. Computational analysis results in 144 tandem repeats and 1,171 palindrome elements. The repeats were distributed in the upstream region of 86% of transcriptional units of *M. hyopneumoniae* strain 7448. Comparative analysis between distinct mycoplasmas, demonstrate different indices of repeat conservation among pathogenic and non-pathogenic strains. Pathogenic strains revealed 59% conservation, while non-pathogenic only 46%. Through assays of quantitative amplification of DNA, different levels of expression in genes coding important proteins have been demonstrated, as glycine hydroxymethyltransferase, lipoprotein, adhesins and GTP-binding protein. These protein coding genes differ in number of palindromes or tandem repeats in respective upstream regions. In addition, repeats found in 206 genes already described to be regulated in different grow conditions in *M. hyopneumoniae* strain 232 showed almost 80% of conservation in relation to *M. hyopneumoniae* strain 7448. All these findings, suggests a potential regulatory role of tandem and palindrome DNA repeats.

Key-words: Tandem, Palindrome, DNA repeats, Transcriptional Regulation

1 INTRODUÇÃO

1.1 A IMPORTÂNCIA DO GÊNERO MICOPLASMA

O primeiro cultivo de micoplasma, um agente da pleuropneumonia bovina, foi relatado em 1898. A partir da sua descoberta foi considerado um organismo viral, até que em meados dos anos 1960s, com os primeiros dados de análises genômicas através de hibridização de DNA, foi incluído no vasto grupo das bactérias (RAZIN & HAYFLICK, 2010).

O gênero *Mycoplasma* compreende um grupo de mais de 200 espécies de eubactérias, filogeneticamente relacionadas a bactérias gram-positivas que se destacam principalmente por seu tamanho diminuto e ausência de parede celular. A ausência total de parede celular explica algumas de suas características únicas: sensibilidade a detergentes e choques osmóticos, resistência a penicilina e formação de uma morfologia de colônia peculiar. Este gênero se caracteriza por ser parasita obrigatório de uma ampla gama de organismos incluindo humanos, plantas e animais (RAZIN & HAYFLICK, 2010).

São os menores micro-organismos autorreplicativos e, portanto, amplamente dependentes de fatores do meio de crescimento, como aminoácidos, purinas, pirimidinas e componentes de membrana (RAZIN, 1992). Normalmente apresentam especificidade ao hospedeiro, possivelmente devido a suas necessidades nutricionais, um resultado direto da redução do genoma que provavelmente ocorreu como consequência da complementaridade metabólica de seus hospedeiros (RAZIN, 1992). Os habitats preferenciais de micoplasmas humanos e animais são as superfícies mucosas do trato respiratório e urogenital, olhos e glândulas mamárias.

O genoma limitado dos micoplasmas (580–1350 kb), com um alto conteúdo de adenina e timina (A + T em torno de 70%) (SIMIONATTO *et al.*, 2013) não permite que tenham uma ampla gama de vias metabólicas associadas. As existentes, normalmente estão integradas com a geração de energia, ao invés de prover substratos para outras vias sintéticas. Todos os micoplasmas analisados até agora, apresentaram sistemas respiratórios truncados, com a ausência de um ciclo do ácido tricarboxílico completo. Além disso, também não possuem quinonas e citocromos, descartando uma via altamente eficiente de fosforilação oxidativa, bem como uma fonte de geração de ATP (POLLACK *et al.*, 1983).

As vias existentes de geração de energia produzem quantidades relativamente baixas de ATP e altas de produtos metabólicos finais. A presença desse mecanismo pouco eficiente faz com que essas bactérias esgotem os tecidos infectados, a partir da utilização de todos os substratos específicos a serem metabolizados que o micro-organismo necessita para o seu bom desenvolvimento.

Entre os micoplasmas, muitas espécies são responsáveis por doenças em humanos, animais e plantas. Entretanto, há uma parcela desses organismos que integram a microbiota natural do hospedeiro, sem causar qualquer quadro infecioso (RAZIN *et al.*, 1998). No caso dos patogênicos, as infecções raramente são do tipo fulminante, mas sim seguem um curso crônico e podem se agravar pela invasão de organismos oportunistas. O resultado da contaminação desses organismos em humanos e animais está mais relacionado a danos ao sistema imune e respostas inflamatórias do que uma toxicidade direta causada pelos componentes da própria célula (RAZIN *et al.*, 1998). Eles tendem a aderir a superfícies epiteliais do trato respiratório ou urogenital e raramente invadem tecidos, podendo então ser considerados parasitas superficiais. Um pré-requisito importante para a colonização e estabelecimento do processo infecioso é um eficiente mecanismo de adesão da bactéria a suas células alvo.

Mycoplasma hyopneumoniae, *Mycoplasma flocculare* e *Mycoplasma hyorhinis* são as espécies mais relevantes que se tem identificado no sistema respiratório de suínos (MARE & SWITZER, 1965; MEYLING & FRIIS, 1972). *M. hyopneumoniae* é o agente etiológico da pneumonia enzoótica suína, enquanto que *M. flocculare*, mesmo sendo amplamente encontrado em rebanhos, não tem relação direta com nenhuma enfermidade (KOBISCH & FRIIS, 1996). Já o *M. hyorhinis* resulta principalmente no desenvolvimento de peritonite, pericardite e artrite. Normalmente ocorre concomitante a outras infecções causadas por patógenos suínos (FRIIS & FEENSTRA, 1994; MORITA *et al.*, 1995; KOBAYASHI *et al.*, 1996; KAWASHIMA *et al.*, 2007).

Análises de sequências de rRNA 16S e comparações genômicas, entre *M. hyopneumoniae* e *M. flocculare* revelaram que as duas espécies são filogeneticamente associadas (STEMKE *et al.*, 1992; SIQUEIRA *et al.*, 2013). Enquanto o primeiro pode aderir aos cílios de células do trato epitelial e causar prejuízos ao organismo, o segundo também adere, porém não resulta em danos importantes ao hospedeiro (YOUNG *et al.*, 2000). Seis linhagens de *M. hyopneumoniae* já foram sequenciadas (MINION *et al.*, 2004;

VASCONCELOS *et al.*, 2005; LIU *et al.*, 2011; LIU *et al.*, 2013; SIQUEIRA *et al.*, 2013), enquanto que em *M. flocculare* duas estão disponíveis (SIQUEIRA *et al.*, 2013; CALCUTT *et al.*, 2015). Entre as diferentes linhagens já identificadas de *M. hyopneumoniae* existe uma variabilidade de patogenicidade entre elas. As linhagens 232, 7448 e 7422 causam doenças quando os suínos são infectadas (MINION *et al.*, 2004; VASCONCELOS *et al.*, 2005; SIQUEIRA *et al.*, 2013). Já a linhagem J não é capaz de causar pneumonia em animais (ZIELINSKI & ROSS, 1990). Estudos indicam que essa linhagem perdeu a capacidade de colonizar o epitélio respiratório suíno, pois apresenta uma baixa atividade de aderência, impossibilitando o estabelecimento do quadro infeccioso e afetando o potencial de patogenicidade dessa bactéria (ZHANG *et al.*, 1995).

1.2 PNEUMONIA ENZOÓTICA SUÍNA E SUAS CONSEQUÊNCIAS.

M. hyopneumoniae como já descrito, é o principal agente etiológico da pneumonia enzoótica suína, uma doença respiratória crônica, que atinge até 80% dos suínos ao redor mundo (THACKER *et al.*, 1999). O sinal clínico primário associado à infecção por *M. hyopneumoniae* é uma tosse esporádica e seca. Outros sinais, como febre ou crescimento deficiente, estão ligados a invasores secundários, especialmente *Pasteurella multocida* (THACKER *et al.*, 1999). A pneumonia enzoótica suína é uma doença com altos índices de morbidade, porém baixa mortalidade. A infecção ocorre em praticamente todas as áreas de produção mundial, causando significativas perdas econômicas principalmente devido aos maiores gastos com utilização de medicamentos, além da diminuição do desempenho do animal (MAES *et al.*, 2008; SIMIONATTO *et al.*, 2013).

O contágio por *M. hyopneumoniae* ocorre principalmente por transferência horizontal de suínos infectados para não infectados. Também é verticalmente transmitida através de fêmeas contaminadas para sua prole, inclusive pelo aleitamento (NATHUES *et al.*, 2013). O estabelecimento da infecção por *M. hyopneumoniae* ocorre com a sua aderência ao epitélio ciliado respiratório, que inicialmente causa ciliostase, destruição dos cílios e, possivelmente, morte do tecido epitelial (DEBEY & ROSS, 1994). Após a adesão bacteriana ao epitélio respiratório, ocorre a estimulação de uma reação inflamatória prolongada, seguida da supressão e modulação das respostas imunes inatas e adaptativas

(THACKER & MINION, 2012), tornando os animais ainda mais suscetíveis a infecções secundárias, que podem agravar o quadro clínico.

A infecção causada por *M. hyopneumoniae* pode ser controlada através da combinação de vários fatores tais como a otimização de práticas de gestão, medidas de biossegurança e condições de manejo, além do uso de antibióticos e vacinação (SIMIONATTO *et al.*, 2013). Entre os antibióticos, tetraciclinas e macrolídeos são os mais frequentemente usados no tratamento da doença (VICCA *et al.*, 2004). A utilização dos medicamentos no controle das infecções de *M. hyopneumoniae*, muitas vezes oferece apenas uma proteção parcial e assim que o tratamento é interrompido, a doença é novamente estabelecida (THACKER & MINION, 2012). Além disso, a medicação preventiva ou estratégica com antimicrobianos deve ser minimizada tanto quanto possível para evitar o desenvolvimento de resistência medicamentosa (SIMIONATTO *et al.*, 2013).

Outro método de controle que vem sendo utilizado é através da vacinação dos animais. Todas as vacinas comercialmente disponíveis são preparadas a partir de culturas de micoplasmas atenuadas ou mortas – bacterinas (SIMIONATTO *et al.*, 2013). A eficácia da imunização tem sido demonstrada em diferentes estudos que apontam efeitos variáveis entre rebanhos. Essa instabilidade pode ser resultante de diferentes fatores como o nível e o tempo da infecção, fatores complicadores e a variabilidade entre diferentes isolados de *M. hyopneumoniae* (VILLARREAL *et al.*, 2011). Os dados apresentados na literatura mostram que um único método sozinho ainda não foi capaz de eliminar por completo a infecção de forma rápida e eficiente, evidenciando a importância do entendimento de mecanismos moleculares que possam identificar novos potenciais alvos no controle de *M. hyopneumoniae.*

1.3 ORGANIZAÇÃO GENÔMICA DE *M. hyopneumoniae*.

M. hyopneumoniae linhagem 7448 teve seu genoma completamente sequenciado em 2005 por VASCONCELOS *et al.* (2005) possibilitando avanços importantes no entendimento da biologia molecular desse organismo. Com um genoma total de 920.079 pb e um conteúdo GC de 28%, *M. hyopneumoniae* possui um total de 681 CDS, representando um índice de 88% de sequência genômica associada a regiões codificantes. As regiões codificantes (CDS) possuem em média 1.190 pb de tamanho, codificando para

421 proteínas com função conhecida, representando um pouco mais da metade do total, sendo o restante classificadas como proteínas hipotéticas. Foram já caracterizados 30 tRNAs, 3 rRNAs e 13 sequências de inserções (VASCONCELOS *et al.*, 2005).

Com o genoma revelado, começaram os esforços para entender processos essenciais dessa célula bacteriana. Estudos iniciais revelaram a organização transcricional de genes de *M. hyopneumoniae* a partir de abordagens tanto *in silico* como *in vitro*. Como resultados gerais, foi possível estabelecer que *M. hyopneumoniae* linhagem 7448 preferencialmente transcreve seus genes em RNAs policistrônicos, estando assim, organizados em unidades transcricionais. Aproximadamente 95% dos genes são transcritos em unidades transcricionais policistrônicas, enquanto que os outros 5% são transcritos singularmente, chamadas então unidades transcricionais monocistrônicas. Isso possibilita o uso mínimo de uma maquinaria de transcrição, uma vez que para muitos genes apenas um aparato transcricional é montado, representando uma significativa economia de energia, em um genoma que já é bastante limitado e dependente de fatores externos (SIQUEIRA *et al.*, 2011).

Elementos que estariam envolvidos no início transcrição, como por exemplo sequências promotoras já foram descritos em alguns micoplasmas (WEINER *et al.*, 2000; WEBER *et al.*, 2012; HALBEDEL *et al.*, 2007; LLORENS-RICO, *et al.*, 2015).

No estudo conduzido por WEBER *et al.* (2012), inicialmente foi investigado, experimentalmente o sítio de início da transcrição de 23 genes de *M. hyopneumoniae* linhagem 7448. Os resultados mostraram um padrão que se assemelha a região -10 de reconhecimento de σ^{70} de *Escherichia coli*. Nenhum elemento comum na região promotora -35 foi estabelecido. Os padrões de sequência identificados nas análises iniciais foram utilizados para a produção de matrizes de pontuação posição-específica, permitindo a identificação de promotores a montante de todas as CDSs de *M. hyopneumoniae* linhagem 7448. Nessa abordagem *in silico*, foi identificado que aproximadamente 26% das CDS possuía pelo menos um promotor na sua região intergênica. Os índices aumentaram para 54% quando foram avaliadas regiões intergênicas de genes orientados em sentidos opostos. Em genes orientados em tandem somente 18% possuiam uma região promotora a montante da CDS. Esses dados reforçam os resultados sobre a organização em forma de unidades transcricionais, onde a transcrição pode ser guiada por promotores que não estão próximos das respectivas regiões intergênicas. Outro dado intrigante foi a identificação de

CDS possuindo mais de um promotor. A existência de múltiplas sequências promotoras pode estar envolvida, por exemplo, na transcrição basal dos genes, ou estar servindo como sítio competitivo de ligação da RNA polimerase (WEBER et al., 2012). A partir das sequências promotoras já descritas, outro trabalho, trazendo uma abordagem mais refinada e direcionada reforça a presença dessas sequências em 86% das unidades transcricionais policistrônicas e monocistrônicas de *M. hyopneumoniae* (SIQUEIRA *et al.*, 2014).

A análise dos dados revela que o genoma de *M. hyopneumoniae* apresenta uma organização das CDSs em forma de *cluster* já estabelecida, bem como a presença de regiões promotoras, essenciais para o início da transcrição. Posteriormente, mecanismos que controlam o fim da transcrição foram elucidados através da caracterização de terminadores intrínsecos (SIQUEIRA *et al.,* 2011; WEBER *et al*., 2012; SIQUEIRA *et al.*, 2014; FRITSCH *et al.,* 2015).

Terminadores Rho-independentes foram identificados no genoma de *M. hyopneumoniae* linhagem 7448 em 74% das unidades transcricionais policistrônicas e em 82% das unidades transcricionais monocistrônicas, a partir de uma abordagem *in silico*, que utilizou diversos algoritmos para a predição, associada a dados de expressão gênica. As sequências gênicas que constituem os terminadores formam estruturas secundarias na forma de grampos (*hairpins)* capazes de promoverem o término da transcrição. Os estudos também permitiram propor um padrão de formação de estrutura secundária durante o processo de término da transcrição (FRITSCH *et al.,* 2015). Com a descrição da organização transcricional e a identificação de sequências promotoras e terminadoras, foi possível estabelecer os elementos básicos que regem o processo da transcrição em *M. hyopneumoniae* linhagem 7448.

Normalmente micoplasmas apresentam um habitat restrito aos seus hospedeiros, e parecem ser responsáveis apenas por mecanismos básicos da célula (GARDNER & MINION, 2010). Entretanto, processos biológicos que expliquem diferentes fatores de patogenicidade entre linhagens de uma mesma espécie assim como a presença de respostas distintas dos organismos submetidos a condições adversas ainda estão pouco compreendidos. Um alvo que vem se tornando bastante interessante para a busca de elementos regulatórios são as regiões intergênicas, que apesar de escassas em genomas procarióticos, já tem sua importância demonstrada em diversos organismos, inclusive em *M. hyopneumoniae* (GARDNER & MINION, 2010; SIQUEIRA *et al.,* 2011).

1.4 ELEMENTOS REPETITIVOS: ASPECTOS GERAIS.

Sequências repetidas de DNA são abundantes em uma diversidade de espécies que variam desde bactérias a mamíferos (TREANGEN & SALZBERG, 2012). O desenvolvimento de técnicas mais rápidas e precisas de genotipagem possibilitou avanços nos estudos genômicos comparativos e a identificação de diferentes padrões de sequências. Quando comparados com eucariotos, bactérias e archaeas possuem genomas pequenos e compactos, tipicamente contendo entre 85-90% de sequências codificantes para proteínas ou RNAs estáveis. A maioria do genoma remanescente possivelmente corresponde a regiões regulatórias. Uma vez que poucas regiões do cromossomo procariótico não possuem características funcionais, poderia se esperar que elementos repetidos fossem evitados, porém, se observa, a existência de centenas ou até milhares de cópias desses elementos em alguns organismos. A disponibilidade de ferramentas de predição de DNA repetitivo permitiu que esses elementos pudessem ser identificados *in silico* uma vez que a sequência gênomica do organismo de interesse esteja disponível (TREANGEN *et al.*, 2009).

Uma repetição pode ser definida como uma subsequência de um dado genoma que se assemelha a outra subsequência no mesmo genoma. Tais repetições não surgem por processos puramente estocásticos, uma vez que normalmente se relacionam com fenômenos biológicos relevantes (TREANGEN *et al.*, 2009). As repetições genéricas podem surgir por transferência horizontal de genes; recombinação homóloga entre repetições menores e degeneradas ou recombinação entre potenciais repetições próximas e pequenas (NIAUDET *et al.*, 1985).

Em 1983, um dos primeiros programas publicamente disponíveis para a busca de repetições em sequências de DNA se tornou acessível aos pesquisadores (MARTINEZ, 1983). Posteriormente, muitos algoritmos foram desenvolvidos para a busca de repetições exatas e degeneradas. A procura pode ocorrer a partir de duas abordagens. Quando a sequência da repetição é previamente conhecida, buscas por similaridade podem ser feitas, uma vez que se têm muitos bancos de dados especializados disponíveis. Porém, se as sequências repetitivas são desconhecidas se torna estritamente necessário o uso de ferramentas do tipo *ab initio*, que identificam, classificam e associam a presença de elementos que se assemelham a alguma vizinhança genética ou funcional no cromossomo.

Embora muitos progressos tenham sido feitos, a tarefa de detecção de famílias de repetições no genoma pelo método *ab initio* continua sendo um desafio computacional, principalmente pela diversidade das características das repetições, as dificuldades de permitir espaços e substituições nos alinhamentos e uma análise estatística eficiente.

As repetições podem ser encontradas em regiões codificantes, regiões intergênicas e sequências de inserções. Quando presentes em regiões codificantes estão mais relacionadas a mudanças na fase de leitura das proteínas, resultando em modificações na sua estrutura, podendo gerar proteínas truncadas e não funcionais. Se a fase de leitura não for alterada, a presença de uma repetição pode ainda modificar ou suprimir a função proteica através de dobramentos estruturais falhos (MARCOTTE *et al.*, 1999). Mesmo que muitos eventos repetitivos em regiões codificantes sejam de fundo deletério, muitas proteínas podem se ajustar a essas condições, permitindo adquirir características vantajosas, como por exemplo, a duplicação de domínios importantes (VOGEL *et al.*, 2004). Quando encontradas em regiões intergênicas, as repetições de DNA, provavelmente representam uma necessidade regulatória ou estrutural ao cromossomo bacteriano. Relatos de repetições afetando o início da transcrição, através da alteração da afinidade de ligação de proteínas regulatórias (MARTIN *et al.*, 2005; METRUCCIO *et al.*, 2009) ou interferindo na distância das sequências promotoras (YOGEV *et al.*, 1991; VAN DER ENDE *et al.*, 2000) já foram descritos.

As sequências de inserção, já muito bem estudadas, podem afetar a expressão de genes próximos, tanto por interrupção/silenciamento, ou até mesmo por ativação. Esses elementos também provêm estruturas para duplicações, deleções, rearranjos e incorporação de DNA exógeno (NAAS *et al.*, 1995). Existem muitos processos biológicos em que as repetições de DNA podem exercer um papel regulatório e fundamental. Para o melhor entendimento da relevância dessas sequências, as principais funções já descritas são relatadas a seguir.

Os elementos repetitivos podem atuar como “reservatórios evolutivos”, visto que são potenciais fontes de eventos de recombinação e transferência lateral. A presença deles pode proporcionar a um sistema, uma diversidade de combinações de sequências proteicas e ser uma fonte de variação genética (PALMER & BRAYTON, 2007). Um exemplo interessante de geração de variabilidade genética por recombinação homóloga é a variação antigênica em *Borrelia* spp. Estes organismos possuem um grande número de plasmídeos que contém a vasta maioria dos elementos repetitivos encontrados em todo o genoma (CASJENS, 1999). Em espécies patogênicas, a variação antigênica está sob forte seleção e ocorre através de processos de recombinação entre as repetições presentes nos plasmídeos.

Os elementos repetitivos podem também causar um silenciamento transiente dos genes, mantendo uma maneira reversível de ganho e perda da função que ocorre naturalmente por controle da expressão gênica. As funções não são perdidas, somente a expressão é ativada ou não, dependendo de condições especificas. Células microbianas estão constantemente mudando seu repertório de proteínas na superfície, sendo assim capazes de enganar o sistema imune dos hospedeiros. Um importante mecanismo responsável por esse tipo de controle é a variação de fase (MOXON *et al.*, 1994), onde o DNA repetitivo se mostra estritamente relacionado, principalmente quando encontrado em regiões regulatórias. Repetições são bastante instáveis e como consequência a alteração de cópias de um elemento em tandem, por exemplo, pode estar envolvido na variação de fase de antígenos importantes (HENDERSON *et al.*, 1999). Essa variação é herdável e reversível. Através de contrações e expansões subsequentes do número de cópias de uma unidade repetida, é possível que um gene oscile entre a sua expressão ou o silenciamento. Resumidamente, as repetições podem ser consideradas oportunidades, uma vez que permitem a geração de novas funções a partir de outras já pré-existentes, além de possibilitar ajustes pré-evolutivos (JACOB, 1977). Por outro lado, podem representar também aspectos negativos, uma vez que desafiam a integridade e organização do cromossomo em que estão inseridos.

1.5 ELEMENTOS TANDEM E PALINDROMES: RELEVÂNCIA EM PROCESSOS REGULATÓRIOS.

Como já discutido na seção anterior, as repetições podem estar envolvidas em diversos processos. Em bactérias elas podem ser divididas como elementos de baixa e alta complexidade. Repetições de baixa complexidade podem ser compostas de oligonucleotídeos simples repetidos de uma maneira sequencial (tandem) através do genoma, enquanto que sequências maiores incluem repetições mais complexas, possuindo ou não *gaps* entre as cópias, tais como palindromes.

Repetições em tandem podem ser definidas como unidades repetidas, que ocorrem de maneira adjacente umas as outras. Normalmente são subdivididas em duas subcategorias:

microssatélites, ou sequências de repetições simples (SSR) e minissatélites. As SSRs são repetições em tandem curtas, como uma unidade de repetição variando entre 1 e 10 nucleotídeos. Os minissatélites, por outro lado, apresentam unidades repetidas maiores que 10 nucleotídeos (GEMAYEL *et al.*, 2012). Estas repetições podem desempenhar muitos papéis nos genomas procarióticos. Um bom exemplo são as SSRs que já foram descritas em muitos organismos como elementos capazes de influenciar a região promotora da transcrição e promover a troca reversível de fenótipos (MOXON *et al.*, 1994).

Os elementos palindrômicos são sequências idênticas quando lidas na direção 5' – 3' ou 3' - 5' do genoma. Apresentando ou não *gaps* na sua estrutura, podem formar grampos no RNA, os quais tem a capacidade de atuar como terminadores da transcrição, ribocomutadores, ou promover a formação de estruturas cruciformes, que influenciam a replicação, a regulação da expressão gênica, a recombinação, entre outros mecanismos (BRAZDA *et al.*, 2011). A formação de estruturas cruciformes é altamente dependente da sequência nucleotídica e requer palindromes, perfeitas ou imperfeitas de pelo menos 6 nucleotídeos (MIKHEIKIN *et al.*, 2006). A super-representação de palindromes, que ocorre de uma maneira não randômica no DNA de todos os organismos, tem acontecido mais notoriamente próxima a regiões promotoras e de início da replicação (WERBOWY *et al.*, 2009). Essas estruturas podem afetar o nível de torção no DNA, a posição dos nucleossomos *in vivo* e a formação de estruturas secundárias (PETER *et al.*, 2004). Além disso, várias proteínas já demonstram interagir com estruturas cruciformes, evidenciando assim ainda mais a importância desses elementos no genoma em que se encontram (MANI *et al.*, 2009)

1.6 REPETIÇÕES DE DNA EM MICOPLASMAS

Estudos anteriores envolvendo micoplasmas já realizaram buscas por padrões de DNA repetitivos específicos e conseguiram estabelecer uma relação funcional com a sua presença. Os trabalhos mais relevantes serão exemplificados abaixo, dando sempre ênfase ao papel regulatório desempenhado pelo elemento:

- *Mycoplasma genitalium:* foram identificadas repetições em tandem, do tipo SSR e a sua presença foi associada a sítios hipervariáveis no operon MgPa, que codifica duas proteínas adesinas, presentes na superfície celular e importantes para a aderência e estabelecimento da infecção no hospedeiro (MA *et al.*, 2010).

- *M. hyorhinis:* genes codificantes para lipoproteínas superficiais de membrana (*vlpA*, *vlpB* e *vlpC*) apresentam dois diferentes polimorfismos associados a DNA repetitivo, que estão envolvidos na variação antigênica da bactéria (YOGEV *et al.*, 1991). Nesse organismo, os elementos repetitivos estão envolvidos em dois níveis distintos de regulação. O primeiro envolve a perda ou o ganho de repetições nas regiões codificantes, que resulta na variação do tamanho e estrutura proteica. Além disso, em um segundo nível, a expressão gênica é controlada através da variação do número de cópias de unidades repetidas entre as regiões -10 e -35 do promotor, afetando a eficiência da transcrição. Os diferentes papéis regulatórios desempenhados pelas repetições dão origem a uma adaptabilidade fenotípica.
- *Mycoplasma fermentans*: um mononucleotídeo de adeninas foi observado como um elemento intragênico variável (THEISS & WISE, 1997) onde a alteração do número de cópias de adeninas, estava associado a diversidade de capacidade de produção de um substrato ligante à lipoproteína entre diferentes linhagens.
- *Mycoplasma bovis*: várias repetições intragênias também já foram identificadas em lipoproteínas superficiais de membranas (LYSNYANSKY *et al.*, 1996) estando associadas com a variação fenotípica desses antígenos, explicada pela presença de polimorfismos de DNA nas respectivas regiões codificadoras.
- *Mycoplasma synoviae:* a presença de repetições no DNA influenciou a expressão de uma hemaglutinina, através do mecanismo de variação de fase (NOORMOHAMMADI *et al.*, 1997).
- *Mycoplasma pneumoniae*: foram encontradas múltiplas cópias de quatro elementos repetitivos grandes, que mostraram estar envolvidos na criação de variações de sequências entre linhagens individuais, resultando em diversidade bacteriana (MUSATOVOVA *et al.*, 2012).
- *Mycoplasma gallisepticum:* a repetição de trinucleotídeos GAA regula o sistema M9/pMGA (codificante para adesinas associadas a hemaglutinação). Dependendo do número de cópias de GAA nas regiões intergênicas a expressão gênica é ativada (LIU *et al.*, 2000).

- *Mycoplasma hominis*: a diversidade de repetições em tandem de 121 aminoácidos na região central de um antígeno variável associado a aderência traz outra situação de regulação associada a DNA repetitivo (ZHANG & WISE, 1996).
- *Mycoplasma pulmonis*: Presença de elementos repetitivos em proteínas Vsa (antígenos superficiais) altamente variáveis, estão relacionadas a invasão do micro-organismo ao hospedeiro. Dependendo do número de cópias da repetição na região C-terminal das proteínas o organismo adere as células e estabelece o processo infeccioso (Simmons *et al.*, 2004).
- *Mycoplasma mycoides*: Repetições em tandem longas e também do tipo SSR já foram identificadas no genoma do organismos, causando principalmente um variabilidade genética que auxilia o entendimento da origem evolutiva do patógeno (Mrazek, 2006; Gosney *et al.*, 2011).
- *M. hyopneumoniae:* Repetições encontradas na região C-terminal da proteína adesina P97 são fonte de variabilidade genética e potenciais alvos para o desenvolvimento de vacinas (Barate *et al.*, 2014). Algumas repetições do tipo SSR também já foram elucidadas nogenoma (Mrazek, 2006).

Como visto acima, existem diversas descrições de funcionalidade associada a elementos tandem em diferentes micoplasmas, porém o mesmo não ocorre quando se trata de repetições palindromes. Poucos relatos associados à função desses elementos são encontrados, sendo o caso mais interessante relativo a sequências palindrômicas apresentando *gaps*, identificadas em *M. hyopneumoniae*. Elas são descritas como sítios de ligação da chaperona HrcA, importante proteína na resposta a choques térmicos (CHANG *et al.*, 2008).

Devido a importância da enfermidade causada por *M. hyopneumoniae,* é fundamental o entendimento de processos regulatórios que regem os principais mecanismos da célula. Mesmo que esses processos ainda sejam pouco compreendidos, a disponibilidade das sequências genômicas de várias linhagens de *M. hyopneumoniae*, que inclusive diferem em níveis de patogenicidade, se torna um fator facilitador para estudos comparativos. As repetições de DNA que amplamente representam elementos regulatórios são interessantes pontos de investigação, e podem explicar respostas biológicas ainda não elucidadas.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Identificar e caracterizar elementos de DNA repetitivos presentes em regiões intergênicas do genoma de *M. hyopneumoniae* linhagem 7448 através da utilização de ferramentas computacionais e abordagens experimentais, inferindo seu possível papel regulatório, principalmente relativo à regulação da transcrição dessa bactéria.

2.2 OBJETIVOS ESPECÍFICOS

- Determinar os diferentes tipos de elementos repetitivos a serem analisados.
- Definir os padrões de sequências dos elementos.
- Classificar os dados oriundos das análises computacionais, relativos à organização transcricional e outros possíveis elementos regulatórios.
- Realizar análises comparativas dos elementos encontrados com genomas de outros micoplasmas relacionados.
- Realizar análise comparativa da localização dos elementos repetitivos em genes já descritos com expressão diferencial.
- Realizar ensaios para validação experimental da predição *in silico.*

3 MANUSCRITO

Repetitive elements in *Mycoplasma hyopneumoniae* transcriptional regulation.

Manuscrito a ser submetido ao periódico ‘DNA research’

O material suplementar referente ao manuscrito está apresentado na sessão ‘Anexos.’

Brief Title:

Repetitive elements in *Mycoplasma hyopneumoniae* transcriptional regulation.

Running Title:

Tandem and palindrome repetitive elements in intergenic regions of *Mycoplasma hyopneumoniae*: association with transcriptional regulation.

Amanda Malvessi Cattani[1], Franciele Maboni Siqueira[1], Rafael Lucas Muniz Guedes[2], Irene Silveira Schrank[1, 3*]

1. Centro de Biotecnologia, Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.
2. Laboratório de Bioinformática, Laboratório Nacional de Computação Científica (LNCC), Petrópolis, RJ, Brazil.
3. Centro de Biotecnologia, Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

*To whom correspondence should be addressed. Tel. +55 51-3308-6055. Fax. +55 51-3308-7309.
E-mail: irene@cbiot.ufrgs.br

Abstract

Transcriptional regulation, a multiple-step process, is still poorly understood in the important pig pathogen *Mycoplasma hyopneumoniae*. Basic motifs like promoters and terminators have been already described, but no other cis-regulatory elements were found. DNA repeat sequences have been shown as an interesting potential source of cis-regulatory elements in. In this work, a genome-wide search for tandem and palindrome repetitive elements was performed in intergenic regions of all coding sequences from *M. hyopneumoniae* strain 7448 . Computational analysis demonstrated the presence of 144 tandem repeats and 1,171 palindrome elements. The DNA repeat sequences were distributed within the 5' upstream region of 86% of transcriptional units of *M. hyopneumoniae* strain 7448. Comparative analysis between distinct repetitive sequences found in related mycoplasma genomes demonstrated different percentages of conservation among pathogenic and non-pathogenic strains. qPCR assays revealed differential expression among genes showing variable numbers of repetitive elements. In addition, repeats found in 206 genes already described to be differentially regulated in different culture conditions of *M. hyopneumoniae* strain 232 showed almost 80% of conservation in relation to *M. hyopneumoniae* strain 7448 repeats. Altogether, these findings suggest a potential regulatory role of tandem and palindrome DNA repeats in *M. hyopneumoniae* transcriptional profile.

Key-words: Tandem, Palindrome, DNA repeats, Transcriptional regulation

1.Introduction

Mycoplasma hyopneumoniae is a diminute bacterium, characterized by a small genome with a low GC content. It is commonly associated with mycoplasmal pneumonia in pigs[1] and infected animals are affected by a sporadic, dry and non-productive cough, retarded growth rate and inefficient utilization of feed.[2] Until now, the genomes of six strains of *M. hyopneumoniae* have been sequenced[3,4,5,6,7] and their availability had been enabled genome-wide association studies.[3, 8] Transcription, a multiple-step mechanism, is finely regulated in all forms of life and is still poorly understood in *M. hyopneumoniae*. The occurrence of transcription units[3,9], promoters[10,11] and terminators[12] had already been described in this species, but the existence of other regulatory sequences remains to be elucidated.

DNA repeats, wide spread and characterized in eukaryote genomes, can also play an import role in genomic regulation of prokaryotes.[13] It is usually hypothesized that repeats arise by successive duplications and several causal mechanisms, like homologous recombination, slipped-strand mispairing of DNA polymerase or by events of genomic transposition.[14] Bacterial repeats are commonly classified as low complexity repeats and longer repeats.[15] Low complexity repeats can be composed of simple oligonucleotides repeated in a head-to-tail configuration (tandem) through the genome, while longer repeats may include complex and spaced repeats, such as palindromes.

DNA repeats can be found within coding-sequences (CDS), in intergenic regions, or in transposable elements.[16] Once found in intergenic regions, these elements could both reflect regulatory or structural requirements of the bacterial chromosome. Intergenic regions repetitive elements have been related to pathogenicity through phase variation in *Neisseria* species, *Haemophilus parainfluenzae*, and *Moraxella catarrhalis.*[17] Many other phenotypes associated with repetitive genotypes were revised by Belkum et al.[13] involving microbial evolution, pathogenesis and molecular epidemiology. In *Mycoplasmas* spp. the presence of some repeats was described in *Mycoplasma genitalium,*[18, 19,20] *Mycoplasma gallisepticum,*[20,21,22] *Mycoplasma bovis,*[23] *Mycoplasma hyorhinis,*[24] *Mycoplasma fermentans,*[25] *Mycoplasma synoviae,*[20, 26] *Mycoplasma pneumoniae*,[20, 27] *Mycoplasma hominis,*[28] *Mycoplasma penetrans,*[20] *Mycoplasma pulmonis,*[20, 29] *Mycoplasma mobile,*[20] *Mycoplasma mycoides,*[20, 30] and also in *M. hyopneumoniae*[20, 31] genomes. The repeats found in *M. genitalium* genome represent hypervariable sites in MgPa operon that codes two adhesin proteins, which represent surface proteins required for the development of the terminal-organelle structure and attachment of the organism to host epithelial cells. Such adhesin variation may allow this organism to evade the host immune response and to adapt to diverse host microenvironments, thus establishing persistent infection.[18] In *M. gallisepticum* the GAA trinucleotide repeat region regulate the M9/pMGA gene expression (that encodes adhesin(s) associated with hemagutination). Depending on the copy number of GAA in the intergenic regions, gene expression can be inhibited.[21] In *M. bovis* repetitive elements were involved in a family of phase-and size-variable membrane surface lipoprotein antigens.[23] Even as some isolated repetitive elements were already elucidated, a genome-wide approach correlating DNA repeats with its possible genomic regulatory consequences in *M. hyopneumoniae* are still not available. Therefore, in this work, an *in silico* prediction associated with experimental validation was performed, aiming to verify and associate the presence of tandem and palindrome repeats as transcriptional regulatory sites.

2. Material and Methods

2.1. Search strategy

Two types of DNA repeats were selected to be investigated in this work, tandem and palindrome elements, based on the most common repeats found in prokaryote genomes already described, mainly in *Mycoplasma* sp.[32]. Initially, tandem sequences were divided into: i) Simple Sequence Repeats (SSR); ii) Simple Sequence Repeats of Mononucleotide (SSRM) and iii) Tandem (longer sequences). Palindromic elements were searched using two approaches: allowing gaps (PALG) or not allowing gaps (PAL). In-house PERL scripts were used to extract intergenic regions up to 500 bp upstream from the start codon of all *M. hyopneumoniae* 7448 strain CDSs which were used as input for the software's prediction. For each type of repeat sequence, two independent software packages were used, as detailed below.

In this work, SSR is defined as a sequence containing 2 to 10 repeated nucleotides as described by Huang et al.[32] The algorithms used to search for SSR were SSRLocator[33] and Simple Sequence Repeat Identification Tool (SSRIT).[34] SSRM is defined as a mononucleotide repeated at least 8 times in a head-to-tail way. Software packages used to predict SSRM were Tandem Repeat Finder (TRF)[35] and SSRIT. Tandem is defined as repeats with 11 to 50 nucleotides repeated at least twice. TRF and etandem from The European Molecular Biology Open Software Suite (EMBOSS)[36] were used for the prediction of this type of repeat. PALG is an element of 9 to 20 nucleotides containing a gap of 5 to 15 bases. Palindrome from EMBOSS and Palindrome Search[37] were the algorithms used in this approach. PAL is a repeat of 9 to 15 bases with no gaps. The same software were used to predict PAL and PALG, and a maximum of 2 mismatches were allowed in the search. Output data of TRF and etandem were filtered to identity of ≥ 95% within copies of a tandem repeat and elements predicted by Palindrome and Palindrome Search with more than 300 bp distant from start codon were excluded. After the individual search of each element, results were grouped and all overlaps were eliminated following the hierarchy: PAL, SSRM, tandem, SSR and PALG.

All repetitive elements found in non-coding region of *M. hyopneumoniae* strain 7448 were classified in relation to its position. Moreover, downstream CDS identification, ATG distance, association with promoter sequences and transcription unit distribution of palindrome and tandem elements were performed. A search for common motifs among repeats classes were performed through the web server MEME SUITE,[38] using default parameters A complete workflow of the search strategy was represented in Fig. 1.

2.2 Comparison analysis

To validate the relevance of repetitive elements found, comparison analysis was performed with three different approaches. Initially, all tandem repeats (SSR, SSRM and Tandem) found upstream the first gene of *M. hyopneumoniae* strain 7448 transcription units were compared against three other mycoplasma genomes: *M. hyopneumoniae* strain 7422 – pathogenic; *M. hyopneumoniae* strain J – non-pathogenic; and *Mycoplasma flocculare* ATCC 27716– phylogenetically related to *M. hyopneumoniae* and non-pathogenic. Therefore, a BLAST[39] search was performed with the first CDS of each transcription unit against the genomes of the three organisms above mentioned. The corresponding tandem motif was manually localized in the 5´upstream region of each ortholog gene. Three classifications could be assigned to each repeat: CONSERVED (C) - repetitive element was exactly the same in sequence and number; NON-CONSERVED (NC) - copy number of tandem element was different; and ABSENT (A) - repetitive element was not found in the ortholog gene. Schematic pipeline of tandem comparison can be seen in Fig. 2.

The second comparison analysis was performed in genes coding for adhesins.[40] Palindrome repeats (PAL and PALG) already predicted upstream of the start codon of adhesins coding genes were compared among the same mycoplasma genomes previously described. However, the classification strategy, in this case, was different than the tandem repeats strategy. Due to the large quantity of palindrome repetitive elements in intergenic regions of orthologs genes, the elements were identified by Patter Locator software.[41] The repeat search was done directly upstream the region of the ortholog gene or in the upstream region of the corresponding transcription unit. A percentage of conservation level of palindrome repeats between mycoplasma genomes was calculated dividing the number of repeats found in the ortholog by the number of repeats found in *M. hyopneumoniae* strain 7448 respective gene (Fig. 3A).

The third approach was based on previous studies of *M. hyopneumoniae* strain 232, which by microarray assays have identified genes differentially expressed in some adverse situations such as exposure to norepinephrine,[42] hydrogen peroxide,[43] *in vitro* infection,[44] iron depletion [45] and heat shock[46]. A BLAST search was performed to find the *M. hyopneumoniae* strain 7448 orthologs for the *M. hyopneumoniae* strain 232 differentially expressed genes. Comparison of palindromes repeats found in *M. hyopneumoniae* strain 7448 against *M. hyopneumoniae* strain 232 was done as described for the adhesin genes search strategy (Fig. 3B). As palindromes elements can form secondary structure, the ΔG was evaluated by QuickFold algorithm,[47] using default parameters.

2.3 qPCR experiments

Culture conditions, RNA isolation and the cDNA synthesis by Reverse Transcription (RT) were performed as Siqueira et al.[48] Target and primers descriptions for *M. hyopneumoniae* strain 7448, *M.*

hyopneumoniae strain J and *M. flocculare* ATCC 27716 reactions are available in Supplementary Table S1. Primers were designed using *Vector NTI Advance 10* (Invitrogen, USA).

Quantitative PCR (qPCR) assay was performed using 1:50 cDNA as template and Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen, USA) on 7500 Real-Time PCR Systems (Applied Biosystems, USA). The qPCR reactions were carried out at 90°C for 2 min and 95°C for 10 min followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. A melting curve analysis for each primer's pair was done to verify primer's efficiency.

Relative expression of each gene was evaluated in *M. hyopneumoniae* strain 7448, *M. hyopneumoniae* strain J and *M. flocculare* ATCC 27716 RNAs. Relative expression of mRNA was calculated by the $2^{-\Delta Ct}$ method.[49] The threshold cycle (CT) values were normalized to the reference gene *lon* (MHP7448_0524). Reference gene was determined in a specific assay where expression profiles of several genes were evaluated in the three mycoplasma RNAs tested (*M. hyopneumoniae* strain 7448, *M. hyopneumoniae* strain J and *M. flocculare* ATCC 27716). The gene that showed no differential expression in all mycoplasmas genomes tested was used to relative expression calculation. Three technical and two biological replicates were done to each target evaluated. Statistical analyses were performed using GraphPad Prism 6 software by One-way ANOVA followed by Tukey's multi comparison test ($P < 0.05$).

3. Results

3.1 *in silico* prediction

Computational analysis predicted the presence of 340 tandem repeats and 1,879 palindromic elements in the genome of *M. hyopneumoniae* strain 7448. Among the tandem repeats, 272 SSRM, 55 SSR and 13 tandem were found. Palindrome elements were divided into 689 PAL and 1,190 PALG (Table 1). These results were further analyzed and all overlaps sequences were excluded. Detailed analysis of palindrome elements demonstrated that PAL elements, in most of the cases, overlap partially with PALG (Table 1) and therefore PAL overlapped elements were excluded from further analysis. The decision for the exclusion was based on the capability of the selected PALG elements to form stronger secondary structure representing an interesting physical modification in the DNA molecule.[50] As result, 59% of previously located repeats were maintained in the *M. hyopneumoniae* strain 7448 genome, distributed as 144 tandems (111 SSRM, 29 SSR and 4 tandems) and 1,171 palindrome repeats (73 PAL and 1,098 PALG) (Table 1; Supplementary Table S2). The presence of the tandem repeats and palindromic elements was confirmed by utilization of two independent software´s. The best results were found with the SSR tandem repeats wherein almost 90% of elements found were predicted by both algorithms tested (Table 1).

The distribution of repetitive elements among 5´upstream region of *M. hyopneumoniae* strain 7448 CDSs demonstrated an average of 3 PALG, 1 PAL, 1 SSRM, 1 SSR and 1 Tandem per 5´upstream region of an unique CDS (Table 1). In accordance with a high density of repetitive elements per CDS, many of them have different combinations of palindrome and tandem repeats (Supplementary Table S2). Comparative analysis of nucleotide sequence in each class of repetitive element was unable to determine a common motif for tandem repeat or palindrome sequence, as they were diverse in nucleotide composition and length (Supplementary Table S2).

3.2 Repeats Classification

All 731 CDSs and 162 transcriptional units (polycistronic and monocistronic) of *M. hyopneumoniae* strain 7448 were classified depending on the putative regulatory element found in the upstream region from the respective start codon. Analyses of all CDSs revealed that 498 of them (68%) have a putative promoter sequence (according to Weber et al.[11]) or a repeat sequence (present analysis) in the 5´upstream region. In the majority of the analyzed CDS (73%) at least one of all-possible putative regulatory element combinations was found. The most representative element in the CDS 5´upstream region are palindrome repeats plus promoter sequences (35%), followed by palindrome repeats only (24%) and the combination of all putative sequences (tandem, palindrome and promoter) representing 14% (Fig. 4A).

The distribution of tandem and palindrome elements showed that individually, PALG were present in the majority of the CDS 5´upstream region investigated (49%). The combination of PALG plus SSRM repeats was seen in 10% of CDS 5´upstream region. SSRM plus PAL elements were seen in 8% of CDS 5´upstream region. None of the CDS had a unique tandem repeat or a combination of all repeats in a single 5´upstream region (Fig. 4B).

Considering all 162 transcriptional units of *M. hyopneumoniae* strain 7448, 149 (92%) of them show a putative regulatory element (repeat or promoter) in the respective 5´upstream region from the start gene of the transcription unit. Among the 149 transcription units, 56% showed a palindrome element associated with a putative promoter; 21% showed the presence of the three elements (tandem, palindrome and promoter), and 10% contain only a palindrome sequence. The three situations mentioned above represented 87% of all possible cases of elements in the regulatory regions of the transcriptional units (Fig. 4C). The classification of palindrome sequences and tandem repeats in the 5´region of the transcription units resulted in: 50% of transcription units with PALG, 13% with PALG plus SSRM and 12% had PAL plus PALG. None of transcription unit 5’upstream region showed the presence of an exclusive tandem or SSR element, and finally, none have the four elements together (Fig. 4D).

3.3 Comparison analysis: conservation of repeat elements in orthologous CDSs

Aiming to understand the prevalence in distribution of the repeat elements identified in *M. hyopneumoniae* strain7448 among related species, a comparative analysis was performed with other Mycoplasma genomes. In this work, the presence of repeat sequence in 5´upstream region of orthologous genes in related strains as *M. hyopneumoniae* strain 7422 (pathogenic) and *M. hyopneumoniae* strain J (non-pathogenic) and also with related non-pathogenic mycoplasma specie – *M. flocculare* - was explored (see Fig. 2 for the pipeline analysis).

A total of 45 genes (Supplementary Table S3) that present a tandem (SSR, SSRM or Tandem) repeat in 5' upstream region and were positioned as the first gene of a transcriptional unit were selected for comparison analysis. Concerning the 45 *M. hyopneumoniae* strain 7448 investigated genes through BLAST approach, 97%, 90% and 67% of them display an ortholog gene in the genome of *M. hyopneumoniae* strain 7422 , *M. hyopneumoniae* strain J and *M. flocculare*, respectively. Repeat comparison analysis was performed only among the ortholog genes found in all genomes and localized in the upstream region of transcription units' first gene. Therefore, 4 tandem repeats, 13 SSR and 44 SSRM elements were selected for further comparative studies.

Detailed analysis revealed that conservation in repeat sequences among the ortholog genes was higher between the same species (*M. hyopneumoniae* strain 7422 or *M. hyopneumoniae* strain J) than with *M. flocculare* sequences (Fig. 5A; Supplementary Table S3). However, differences were found between *M. hyopneumoniae* strain 7448 versus *M. hyopneumoniae* strain 7422 and *M. hyopneumoniae* strain 7448 versus *M. hyopneumoniae* strain J when each repetitive element was analyzed. Higher differences could be seen in SSRM elements, which are 60% conserved in *M. hyopneumoniae* strain 7422, 30% in *M. hyopneumoniae* strain J and had none element conserved in *M. flocculare*. Interestingly, although the number of repetitive elements classified as "non-conserved" maintain the same pattern previously observed for strains of *M. hyopneumoniae*, it increased during analysis between *M. hyopneumoniae* strain 7448 and *M. flocculare* ortholog genes. Tandem and SSR showed lower values of divergence in copy number compared to SSRM (see Fig. 5B).

As expected, the number of orthologs that do not share repetitive element with *M. hyopneumoniae* strain 7448 were higher in *M. flocculare* genome, mainly within the tandem repeats class, that were totality absent. Lower absence values were seen in SSR repeats compared to tandem repeat class and none SSRM were absent in the tested situation (see Fig. 5C).

To understand the relation between repetitive elements presence and pathogenicity a comparative analysis was further performed considering the following: repeats must be identical in *M. hyopneumoniae* strain 7448 and *M. hyopneumoniae* strain 7422 ortholog genes and only those elements that fitted this feature were analyzed in relation to *M. hyopneumoniae* strain J or *M. flocculare* orthologs. These comparisons were named Pathogenic-non-Pathogenic species (P-nP(sp)) and Pathogenic-non-Pathogenic genera (P-nP(gn)), respectively. Among *M. hyopneumoniae* strains two tandem repeats types, 12 SSR and 39 SSRM could be analyzed. Higher conservation could be observed in P-nP(sp), and a reversed scenario was seen in P-nP(gn) where non-conserved and absent tandem repeats were found (Fig. 5; Supplementary Table S3). These results demonstrated that although the number of tandem repeats is almost the same in the ortholog genes of *M. hyopneumoniae* strains the number of SSR and SSRM decrease. Moreover, the numbers of orthologs that share repeat conservation among the pathogenic *M. hyopneumoniae* strains and *M. flocculare* orthologs decrease drastically (see Fig. 5).

In summary, comparison of *M. hyopneumoniae* strain 7448 ortholog genes revealed that 59% of repeats were conserved, 36% were non-conserved and 6% were absent in *M. hyopneumoniae* strain 7422 orthologs. Considering *M. hyopneumoniae* strain J, the values change to 46% of repeat conservation, 46% of non-conservation and 8% of absence in the ortholog genes. *M. flocculare* orthologs demonstrated only 6% of repeat conservation, 44% of non-conservation and 75% of absence. The comparison among the ortholog genes relating to pathogenicity P-nP(sp) showed 40% of repeat conservation and 9% and 6% of non-conservation and absence, respectively. The P-nP(gn) analysis revealed that only 6% of repeats are conserved and 56% were non-conserved or absent among orthologs genes (Fig. 5D; Supplementary Table S3).

Repetitive elements in the upstream regions of the adhesin coding genes from the three *M. hyopneumoniae* strains and *M. flocculare* were searched using the same approach described above. Sixteen *M. hyopneumoniae* adhesins coding genes[40] were selected for comparison analysis. In all genes, palindrome repeats (PAL and PALG) were found with an average of 5 repeat elements per gene. In the majority of the genes coding for adhesins (12 genes) palindrome elements were located directly in the CDS respective 5’ upstream region. However, this situation was not applied to MHP7448_0005, MHP7448_0006, MHP7448_0361, and MHP7448_0362 CDSs. Therefore, a search for repetitive elements was performed in the upstream region of TU 01 (containing MHP7448_0005 and MHP7448_0006) and TU 63 (containing MHP7448_361 and MHP7448_362). Using this approach PAL and PALG elements could be found. Comparative analysis revealed that all *M. hyopneumoniae* strain 7422 and *M. hyopneumoniae* strain J adhesin ortholog genes were present into the respective genome and a palindrome repeats conservation level of 87% and 69%, respectively, were established. Comparison between *M. hyopneumoniae* strain 7448 and *M. flocculare* demonstrated only 27% of conserved elements in 5’ upstream region of the ortholog genes (Table 2; Supplementary Table S4).

3.4 Experimental validation

The *in silico* comparison analysis showed differences related to copy number or repeats presence in the upstream regions of CDSs from different mycoplasma genomes. Therefore, 12 CDSs with different tandem repeats compositions at 5’ upstream region were selected to perform experimental procedures (3 CDSs with tandem repeats, 4 with SSR repeats, and 5 with SSRM repeats) as well as 9 genes encoding adhesins with palindromic sequences at the upstream region (Table 3).

Comparative transcription quantification was performed considering the conservation level of SSR and SSRM elements in genes from *M. hyopneumoniae* strain J and *M. flocculare,* which are orthologs to *M. hyopneumoniae* strain 7448. Analysis of these results revealed that, when repeats were classified as conserved in one of the mycoplasma genomes in relationship to *M. hyopneumoniae* strain 7448, non-significant differences in the basal transcription level was observed among orthologs genes. However, in genes whose DNA repeats were classified as non-conserved or absent in relation to *M. hyopneumoniae* strain 7448, significant differences in the transcriptional levels of these genes was found (Table 3). Interestingly, genes with tandem repeats in the upstream region displayed a different expression profile in relation to those with SSR and SSRM repeats, as only conserved elements demonstrated distinct expression profile among *M. hyopneumoniae* strain 7448, *M. hyopneumoniae* strain J and *M. flocculare* bacteria (Table 3).

Detailed analyses of conserved SSR element found in the 5’ upstream region of MHP7448_0623 ortholog from *M. hyopneumoniae* strain J and *M. flocculare* showed no significant differences in basal gene transcription levels in relation to *M. hyopneumoniae* strain 7448 (Supplementary Fig. S1). However, the gene *glyA* among the mycoplasma orthologs display a non-conserved SSRM in 5’ upstream region and exhibited differential expression in the tested condition (Supplementary Fig. S1). The MHP7448_0087 orthologs in both *M. hyopneumoniae* strain J and *M. flocculare* were classified as SSR absent and different gene transcription levels comparing to *M. hyopneumoniae* strain 7448 was observed (Supplementary Fig. S1).

In relation to genes that encoded adhesins in *M. hyopneumoniae* strain 7448, MHP7448_0005 and MHP7448_0006, with 100% of conservation in the palindrome elements at 5’ upstream region of *M. hyopneumoniae* strain J orthologs, significant differences at gene relative expression in relation to *M. hyopneumoniae* strain 7448 was not observed. However, when relative conservation dropped to 60% in 5´upstream region of orthologs of both genes from *M. flocculare*, basal transcript level varied significantly. Orthologs of MHP7448_0271 in *M. hyopneumoniae* strain J and *M. flocculare* genome without palindrome element resulted in significant differences in basal gene transcription level comparing to *M. hyopneumoniae* strain 7448 gene. However, MHP7448_0272 gene showed differential expression level even when few elements were lost in 5´upstream region orthologs of *M. hyopneumoniae* strain J and *M. flocculare* in relation to *M. hyopneumoniae* strain 7448 (Table 3 and Supplementary Fig. S2). Detailed results related to adhesin coding genes with palindrome elements that were experimentally analyzed were described in Supplementary Table S4. In summary, genes that had differences in repetitive DNA elements among the 5’ upstream regions of two strains of *M. hyopneumoniae* (7448 and J) and *M.flocculare* genomes varied in gene expression in the tested conditions. All these findings suggest a putative regulatory influence at gene expression level of both tandem and palindromes repeats.

3.5 Repeats presence in CDS differentially regulated

Aiming to understand the possible role of repetitive elements in transcription regulation a search for the presence of palindromic elements was performed in genes with differential expression profile. A comparison study was done in the upstream regions of genes already described as potentially regulated in different situations.[42,43,44,45,46] A total of 243 *M. hyopneumoniae* strain 232 differentially expressed genes were compared against *M. hyopneumoniae* strain 7448 genome (Supplementary Table S5 and S6). In 206 (85%) of the differentially expressed genes, an ortholog could be found through BLAST approach in the *M. hyopneumoniae* strain 7448 genome. An average of 4 palindrome elements was identified in these genes from *M. hyopneumoniae* strain 7448 genome (Supplementary Table S5 and S6). Comparison with ortholog genes from *M. hyopneumoniae* strain 232 showed that 103 (61%) genes have exactly the same element found in *M. hyopneumoniae* strain 7448 and only 17 (10%) do not have any corresponding element (Fig. 6; Supplementary Table S5 and S6). In conclusion, almost 80% of palindromic elements were conserved among differentially expressed genes, reinforcing the possible importance of PALG and PAL repeats in transcriptional regulation. The ΔG values of PAL and PALG found in differentially expressed genes were evaluated and at least one of them always demonstrated a good potential to form secondary structures (Supplementary Table S5 and S6).

4. Discussion

Prokaryote genomes are extremely diverse in terms of nucleotide composition and presence of distinct patterns of repeat sequences that could affect physical properties of DNA molecule.[32] In this work a global analysis of tandem and palindrome repetitive elements found in non-coding sequences of *M. hyopneumoniae* strain 7448 was reported. The *in silico* analysis revealed that the majority of the repeats found were classified as palindromes elements (1,171 elements) compared to tandem repeats (144 elements). Similar results were also reported by Huang et al. [32] that investigated tandem and palindromes repeats in protein-coding sequence and intergenic regions through a global analysis with more than 1,000 genomes, including *Mycoplasma agalactiae, M. bovis, M. fermentans* and *Mycoplasma mycoides.* Further analysis of the repetitive elements identified in our work demonstrated that similar motifs between tandem or palindromes repetitive elements could not be established. The pattern frequently observed was the presence of AT rich sequences in all elements investigated. In SSRM, for example, mononucleotides containing only Adenines (A) or Thymine (T) repeated 8 to 25 times were observed, in accordance with AT rich genome of *M. hyopneumoniae* strain 7448.[5]

Combinations of palindrome elements, tandem repeats and promoter sequences could be detected in 92% of all *M. hyopneumoniae* strain 7448 transcriptional units demonstrating that different mechanisms of regulation can be considered (see Supplementary Table S2). Previous work demonstrated that copy number of tandem repeats next to the promoter can modulate RNA polymerase action, by spacing the promoter region in distinct way, affecting gene expression.[51] Moreover, palindromes can form cruciform structures, and mediate promoter sequence availability or create physical barriers, that could be broken in a regulatory way.[50]

DNA repeats have already been described involved in phase variation resulting in diversity of pathogenic phenotypes and others bacterial biological process.[52] Therefore, aiming to establish the role of the repetitive elements found in *M. hyopneumoniae* strain 7448 a comparative investigation of tandem elements was performed in relation to mycoplasmas genomes. Although subtle differences were observed, in general, the level of repeat conservation was higher between the two pathogenic *M. hyopneumoniae* strains (7448 and 7422), compared to non-pathogenic mycoplasmas (*M. hyopneumoniae* strain 7448 versus *M. hyopneumoniae* strain J and *M. hyopneumoniae* strain 7448 versus *M. flocculare*) as shown in Fig. 5. Adhesin proteins are known to be essential in mycoplasmas' host infection and for the establishment of the disease.[40] To evaluate the correlation of the palindrome repeats associated with these genes, another comparison analysis was done. A comparative analysis of the PAL and PALG elements among the ortholog adhesin genes in the three *M. hyopneumoniae* strains and *M. flocculare* was performed with similar results as previous demonstrated for the tandem repeats. Adhesin coding genes show higher difference in the number of palindrome in non-pathogenic strains (*M. hyopneumoniae* strain J and *M. flocculare*) as compared between the two *M. hyopneumoniae* pathogenic strains (7448 and 7422). An example was seen in MHP7448_0107, MHP7448_0271 and MHP7448_497 genes that showed 100% conservation level in *M. hyopneumoniae* strain 7422 and an expressive reduction of palindrome elements in *M. hyopneumoniae* strain J and *M. flocculare*. Previous reports have described that tandem repeats, more precisely SSR, were able to influence adhesin genes expression[18,21] but differences in palindromes repeats had not yet been reported.

Aiming to understand the role of the repetitive elements found among the 5´ upstream region in different CDSs an experimental analysis was performed with ortholog genes from two *M. hyopneumoniae* strains (7448 and J) and *M. flocculare*. Detailed analysis of the comparative studies of the presence of tandem repeats (SSR, SSRM and tandem) in 45 different genes and palindrome elements found in 16 adhesins coding genes were the base of the experimental investigation. In general, results of qPCR assay revealed that conserved elements among distinct mycoplasmas seem to have no influence in gene expression, as observed in MHP7448_0623 gene (Table 3). Whereas when non-conserved or absent elements were investigated a relation between differences in repetitive elements found in 5' upstream region and gene expression could be suggested (Table 3). Interestingly, a SSR found in 5' upstream region of MHP7448_0087 (GTP-binding protein) gene was conserved in all mycoplasmas pathogenic strains (7448 and 7422) and absent in non-pathogenic mycoplasmas (*M. hyopneumoniae* strain J and *M. flocculare*). The same pattern was present in SSRM found in 5' upstream region of *glyA* (glycine hydroxymethyltransferase) and *sipS* (signal peptidase I) genes that are conserved in pathogenic mycoplasmas strains and non-conserved in non-pathogenic mycoplasmas (Supplementary Table S3). Differences in gene expression were observed among tested mycoplasmas RNAs of *M. hyopneumoniae* strain J and *M. flocculare* in relation to *M. hyopneumoniae* strain 7448 in

MHP7448_0087, *glyA* and *sipS* genes (Table 3), suggesting that some of DNA repeats analyzed may be involved in mycoplasma pathogenicity.

The adhesin coding genes are important virulence factors and seem to be regulated depending on the presence or absence of palindrome repeats. Therefore, it's possible to suggest a relationship between the differences in gene transcriptional level and the divergence of DNA repeats found in 5' upstream region among mycoplasma ortholog genes, as observed in genes coding for adhesins proteins P102 (MHP7448_0271) and P97 (MHP7448_0272) and genes coding for proteins in MgPa operon (MHP7448_0005 and MHP7448_0006). Similar results were previously demonstrated for genes coding adhesin in *M. genitalium* (MgPa adhesin proteins)[18] and in *M. gallisepticum* (M9/pMGA adhesin proteins*)*[21] whereas repetitive DNA could influence in transcription regulation.

Previous studies have demonstrated differential expression in genes from *M. hyopneumoniae* 232 growing on several specific culture conditions.[42,43,44,45,46] A detailed analysis was performed to better understand the differences in transcriptional level of genes that diverge in the presence of repetitive elements at 5' upstream region. All *M. hyopneumoniae* strain 7448 orthologs from differentially expressed genes of the *M. hyopneumoniae* strain 232 were selected and analyzed for the repeats conservation. This study revealed that almost all differentially expressed genes had palindrome elements at 5' upstream region. Ortholog gene comparison between the two *M. hyopneumoniae* strains (7448 and 232) showed that the PAL and PALG repeats were approximately 80% conserved in *M. hyopneumoniae* strain 232, reinforcing the putative regulatory role of these repetitive elements.

In this work, the presence of SSR, SSRM, tandem, PAL and PALG DNA repeats found in 5' upstream regions of *M. hyopneumoniae* strain 7448 CDSs was described. Relevance in transcriptional regulation of DNA repeats found could be established through comparison analysis, demonstrating that repeats could be perpetuated among related mycoplasmas, and some of them could be involved in pathogenicity. Experimental assays revealed differential expression in genes differing in repetitive elements in the 5' upstream region, reinforcing the putative regulatory role of palindrome and tandem repeats. Previously described differentially expressed genes show palindrome elements in the upstream region of the start codon and are conserved between two strains of *M. hyopneumoniae*. All these finding suggests the importance of repetitive DNA elements in *M. hyopneumoniae*, although further analysis should be done to confirm the regulatory potential role in transcription regulation, reported in this pioneering study.

5. Funding

This work was funded by Coordenação de Apoio de Pessoal de Nível Superior – CAPES. CAPES-Biologia Computacional (Proccess number: 23038.010043/2013-02) and MCTI/CNPq/ Universal (Proccess number: 445228/2014-8).

6. References

1. Thacker, E. L. 2004, Diagnosis of *Mycoplasma hyopneumoniae*, *Anim Health Res Rev.*, *5* (2), 317-20.
2. Kuhnert, P., Overesch, G. 2014, Molecular epidemiology of *Mycoplasma hyopneumoniae* from outbreaks of enzootic pneumonia in domestic pig and the role of wild boar, *Vet Microbiol.*, *174* (1-2), 261-6.
3. Siqueira, F. M., Thompson, C. E., Virginio, V. G., et al. 2013, New insights on the biology of swine respiratory tract mycoplasmas from a comparative genome analysis, *BMC Genomics*, *14*, 175.
4. Minion, F. C., Lefkowitz, E. J., Madsen, M. L., Cleary, B. J., Swartzell, S. M., Mahairas, G. G. 2004, The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis, *J Bacteriol.*, *186* (21), 7123-33.
5. Vasconcelos, A. T., Ferreira, H. B., Bizarro, C. V., et al. 2005, Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*, *J Bacteriol*., 187, 5568-77.
6. Liu, W., Feng, Z., Fang, L., et al. 2011, Complete genome sequence of *Mycoplasma hyopneumoniae* strain 168, *J Bacteriol.*, *193* (4), 1016-7.
7. Liu, W., Xiao, S., Li, M., et al. 2013, Comparative genomic analyses of *Mycoplasma hyopneumoniae* pathogenic 168 strain and its high-passaged attenuated strain, *BMC Genomics*, *14*, 80.
8. Liu, W., Fang, L., Li, M., et al. 2012, Comparative genomics of Mycoplasma: analysis of conserved essential genes and diversity of the pan-genome, *PLoS One*, *7* (4), 35698.
9. Siqueira, F. M., Gerber, A. L., Guedes, R. L., Almeida, L. G., Schrank, I. S., Vasconcelos, A. T., Zaha, A. 2014, Unravelling the transcriptome profile of the Swine respiratory tract mycoplasmas, *PLoS One*, *9* (10), 110327.
10. Siqueira, F. M., de Souto Weber, S., Cattani, A. M., Schrank, I. S. 2014, Genome organization in *Mycoplasma hyopneumoniae*: identification of promoter-like sequences, *Mol Biol Rep.*, *41* (8), 5395-402.
11. Weber S de, S., Sant'Anna, F. H., Schrank, I. S. 2012, Unveiling *Mycoplasma hyopneumoniae* promoters: sequence definition and genomic distribution, *DNA Res.*, *19* (2), 103-15.
12. Fritsch, T. E., Siqueira, F. M., Schrank, I. S. 2015, Intrinsic terminators in *Mycoplasma hyopneumoniae* transcription, *BMC Genomics*, *16*, 273.
13. van Belkum, A., Scherer, S., van Alphen, L., Verbrugh, H. 1998, Short-sequence DNA repeats in prokaryotic genomes, *Microbiol Mol Biol Rev.*, *62* (2), 275-93.
14. Treangen, T. J., Abraham, A. L., Touchon, M., Rocha, E. P. 2009, Genesis, effects and fates of repeats in prokaryotic genomes, *FEMS Microbiol Rev.*, *33* (3), 539-71.
15. Achaz, G., Rocha, E. P., Netter, P., Coissac, E. 2002, Origin and fate of repeats in bacteria, *Nucleic Acids Res.*, *30* (13), 2987-94.
16. Rocha, E. P., Danchin, A., Viari, A. 1999, Analysis of long repeats in bacterial genomes reveals alternative evolutionary mechanisms in *Bacillus subtilis* and other competent prokaryotes, *Mol Biol Evol.*, *16* (9), 1219-30.
17. Peak, I. R., Jennings, M. P., Hood, D. W., Bisercic, M., Moxon, E. R. 1996, Tetrameric repeat units associated with virulence factor phase variation in *Haemophilus* also occur in *Neisseria* spp. and *Moraxella catarrhalis*, *FEMS Microbiol Lett.*, *137* (1), 109-14.
18. Ma, L., Jensen, J. S., Mancuso, M., Hamasuna, R., Jia, Q., McGowin, C. L., Martin, D. H. 2010, Genetic variation in the complete MgPa operon and its repetitive chromosomal elements in clinical strains of *Mycoplasma genitalium*, *PLoS One*, *5* (12), e15660.
19. Ma, L., Jensen, J. S., Mancuso, M., Hamasuna, R., Jia, Q., McGowin, C. L., Martin, D. H. 2012, Variability of trinucleotide tandem repeats in the MgPa operon and its repetitive chromosomal elements in *Mycoplasma genitalium*, *J Med Microbiol.*, *61* (2), 191-7,
20. Mrazek, J. 2006, Analysis of distribution indicates diverse functions of simple sequence repeats in Mycoplasma genomes, *Mol Biol Evol.*, *23* (7), 1370-85.
21. Liu, L., Dybvig, K., Panangala, V. S., van Santen, V. L., French, C. T. 2000, GAA trinucleotide repeat region regulates M9/pMGA gene expression in *Mycoplasma gallisepticum*, *Infect Immun.*, *68* (2), 871-6.
22. Liu, L., Panangala, V. S., Dybvig, K. 2002, Trinucleotide GAA repeats dictate pMGA gene expression in *Mycoplasma gallisepticum* by affecting spacing between flanking regions, *J Bacteriol.*, *184* (5), 1335-9.

23. Lysnyansky, I., Rosengarten, R., Yogev, D. 1996, Phenotypic switching of variable surface lipoproteins in *Mycoplasma bovis* involves high-frequency chromosomal rearrangements, *J Bacteriol.*, *178* (18), 5395-401.
24. Yogev, D., Rosengarten, R., Watson-McKown, R., Wise, K. S. 1991, Molecular basis of Mycoplasma surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5' regulatory sequences, *Embo j.*, *10* (13), 4069-79.
25. Theiss, P., Wise, K. S. 1997, Localized frameshift mutation generates selective, high-frequency phase variation of a surface lipoprotein encoded by a mycoplasma ABC transporter operon, *J Bacteriol.*, *179* (12), 4013-22.
26. Noormohammadi, A. H., Markham, P. F., Whithear, K. G., Walker, I. D., Gurevich, V. A., Ley, D. H., Browning, G. F. 1997, *Mycoplasma synoviae* has two distinct phase-variable major membrane antigens, one of which is a putative hemagglutinin, *Infect Immun.*, *65* (7), 2542-7.
27. Musatovova, O., Kannan, T. R., Baseman, J. B. 2012, *Mycoplasma pneumoniae* large DNA repetitive elements RepMP1 show type specific organization among strains. *PLoS One*, *7* (10), 47625.
28. Zhang, Q., Wise, K. S. 1996, Molecular basis of size and antigenic variation of a *Mycoplasma hominis* adhesin encoded by divergent vaa genes, *Infect Immun.*, *64* (7), 2737-44.
29. Simmons, W. L., Denison, A. M., Dybvig, K. 2004, Resistance of *Mycoplasma pulmonis* to complement lysis is dependent on the number of Vsa tandem repeats: shield hypothesis, *Infect Immun.*, *72* (12), 6846-51.
30. Gosney, F., Corro, M., Iob, L., McAuliffe, L., Nicholas, R. A. 2011, Variable number tandem repeat (VNTR) typing of strains of *Mycoplasma mycoides* subspecies mycoides small colony isolated from the north-eastern regions of Italy between 1990 and 1993, *Vet Microbiol.*, *147* (1-2), 220-2.
31. Chang, L. J., Chen, W. H., Minion, F. C., Shiuan, D. 2008, Mycoplasmas regulate the expression of heat-shock protein genes through CIRCE-HrcA interactions, *Biochem Biophys Res Commun.*, *367* (1), 213-8
32. Huang, Y., Mrazek, J. 2014, Assessing diversity of DNA structure-related sequence features in prokaryotic genomes, *DNA Res.*, *21* (3), 285-97.
33. Maia, L. C., Palmieri, D. A., de Souza, V. Q., Kopp, M. M., de Carvalho, F. I., Costa de Oliveira, A. 2008, SSR Locator: Tool for Simple Sequence Repeat Discovery Integrated with Primer Design and PCR Simulation, *Int J Plant Genomics.*, 412696.
34. Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., McCouch, S. 2001, Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential, *Genome Res.*, *11* (8), 1441-52.
35. Benson, G. 1999, Tandem repeats finder: a program to analyze DNA sequences, *Nucleic Acids Res.*, *27* (2), 573-80.
36. Rice, P., Longden, I., Bleasby, A. 2000, EMBOSS: the European Molecular Biology Open Software Suite, *Trends Genet.*, *16* (6), 276-7.
37. Freund, T., Engel, N. 2013, Finding approximate palindromes in genomic sequences.
38. Bailey, T. L., Boden, M., Buske, F. A., et al. 2009, MEME SUITE: tools for motif discovery and searching, *Nucleic Acids Res*, *37*, 202-8.
39. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. 1990, Basic local alignment search tool, *J Mol Biol.*, *215* (3), 403-10.
40. Ferreira, H. B., Castro, L. A. 2007, A preliminary survey of *M. hyopneumoniae* virulence factors based on comparative genomic analysis, *Genet. Mol. Biol.*, *30* (1), 245-255.
41. Mrazek, J., Xie, S. 2006, Pattern locator: a new tool for finding local sequence patterns in genomic DNA sequences, *Bioinformatics*, *22* (24), 3099-100.
42. Oneal, M. J., Schafer, E. R., Madsen, M. L., Minion, F. C. 2008, Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to norepinephrine, *Microbiology*, *154* (9), 2581-8.
43. Schafer, E. R., Oneal, M. J., Madsen, M. L., Minion, F. C. 2007, Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to hydrogen peroxide, *Microbiology*, *153* (11), 3785-90.
44. Madsen, M. L., Puttamreddy, S., Thacker, E. L., Carruthers, M. D., Minion, F. C. 2008, Transcriptome changes in *Mycoplasma hyopneumoniae* during infection, *Infect Immun.*, *76* (2), 658-63.
45. Madsen, M. L., Nettleton, D., Thacker, E. L., Minion, F. C. 2006,Transcriptional profiling of *Mycoplasma hyopneumoniae* during iron depletion using microarrays, *Microbiology*, *152* (4), 937-44.

46. Madsen, M. L., Nettleton, D., Thacker, E. L., Edwards, R., Minion, F. C. 2006, Transcriptional profiling of *Mycoplasma hyopneumoniae* during heat shock using microarrays, *Infect Immun.*, *74* (1), 160-6.
47. Zuker, M. 2003, Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res.*, *31* (13), 3406-15.
48. Siqueira, F. M., Schrank, A., Schrank, I. S. 2011, *Mycoplasma hyopneumoniae* transcription unit organization: genome survey and prediction, *DNA Res.*, *18* (6), 413-22.
49. Livak, K. J., Schmittgen, T. D. 2001, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods*, *25* (4), 402-8.
50. Brazda, V., Laister, R. C., Jagelska, E. B., Arrowsmith, C. 2011, Cruciform structures are a common DNA feature important for regulating biological processes, *BMC Mol Biol.*, *12*, 33.
51. Gemayel, R., Cho, J., Boeynaems, S., Verstrepen, K. J. 2012, Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences, *Genes*, *3* , 461-80.
52. Bayliss, C. D., Palmer, M. E. 2012, Evolution of simple sequence repeat-mediated phase variation in bacterial genomes, *Ann N Y Acad Sci.*, *1267*, 39-44.

Table 1: Computational analysis of tandem and palindrome repeats in *M. hyopneumoniae* strain 7448 genome.

Repeats	**Software**	**General Data**		**Refined Data**	
		N° repeats	**N° CDS**	**N° repeats**	**N° CDS**
	TRF	1	1	0	0
SSRM	SSRLocator	229	165	82	68
	TRF x SSRLocator*	42	41	29	29
	Total	**272**	**192**	**111**	**89**
	SSRIT	5	4	2	2
SSR	SSRLocator	3	3	1	1
	SSRIT x SSRLocator*	47	44	26	25
	Total	**55**	**48**	**29**	**28**
	TFR	4	4	2	2
Tandem	Etandem	8	8	1	1
	TRF x etandem*	1	1	1	1
	Total	**13**	**13**	**4**	**4**
	Palindrome	226	169	26	25
PAL	Palindrome Search	244	174	27	26
	Palindrome x Palindrome Search*	219	169	20	19
	Total	**689**	**314**	**73**	**62**
	Palindrome	843	336	779	321
PALG	Palindrome Search	168	134	149	119
	Palindrome x Palindrome Search*	179	138	170	132
	Total	**1190**	**393**	**1098**	**373**

General data represent independent software prediction. Refined data show filtered sequences, without superposition. *Prediction of individual software was combined and repeats that are common in both packages software were shown.

Table 2: Presence of repeat elements on upstream regions of adhesins coding genes.

MHP_7448		**N° ortholog element /N° elements MHP_7448**		
Gene	**Product**	**MHP_7422**	**MHP_J**	**MFL**
MHP7448_361	P29	5/5 (100%)	5/5 (100%)	3/5 (60%)
MHP7448_362	P69			
MHP7448_497	P76	5/5 (100%)	3/5 (60%)	1/5 (20%)
MHP7448_198	P97	6/6 (100%)	3/6 (50%)	-
MHP7448_108	P97	3/3 (100%)	2/3 (67%)	1/3 (33%)
MHP7448_272	P97	5/6 (83%)	5/6 (83%)	4/6 (67%)
MHP7448_199	P102	4/5 (80%)	5/5 (100%)	1/5 (20%)
MHP7448_107	P102	1/1 (100%)	1/1 (100%)	0/1 (0%)
MHP7448_271	P102	1/1 (100%)	0/1 (0%)	0/1 (0%)
MHP7448_105	P102	3/3 (100%)	1/3 (33%)	3/3 (100%)
MHP7448_663	P146	7/7 (100%)	6/7 (86%)	4/7 (57%)
MHP7448_496	P216	2/2 (100%)	1/2 (50%)	1/2 (50%)
MHP7448_373	LppS	1/2 (50%)	0/2 (0%)	1/2 (50%)
MHP7448_372	LppT	1/5 (20%)	2/5 (40%)	2/5 (40%)
MHP7448_006	MgPa	5/5(100%)	5/5(100%)	3/5(60%)
MHP7448_005	Mgpa			
Total		**49/56 (87%)**	**39/56 (69%)**	**24/56 (27%)**

No MHP7448_198 ortholog gene was found in *M. flocculare* (MFL) genome using BLAST. Abbreviations: *M. hyopneumoniae* strain 7448 = MHP_7448, *M. hyopneumoniae* strain 7422 = MHP_7422, and *M. hyopneumoniae* strain J = MHP_J.

Table 3: Repeats composition influence on mycoplasmas gene expression.

MHP_7448			**MHP_J**			**MFL**		
Repeat Analyzed	**Gene**	**Product**	**Conservation**	**Expression**	**p value**	**Conservation**	**Expression**	**p value**
Tandem_01_TE	rpsP	30S ribosomal protein S16	C	up	***	A	-	ns
Tandem_01_T	MHP7448_0397	hypothetical protein	C	up	**	A	-	ns
Tandem_06_E	MHP7448_0197	hypothetical protein	NC	-	ns	A	-	ns
SSR_49_LI	MHP7448_0485	hypothetical protein	NC	-	ns	NC	up	**
SSR_30_LI	MHP7448_0484	hypothetical protein	NC	-	ns	NC	up	***
SSR_42_LI	MHP7448_0623	ABC transporter ATP-binding - Pr1	C	-	ns	C	-	ns
SSR_05_LI	MHP7448_0087	GTP-binding protein	A	up	****	A	up	*
SSRM_07_TL	sipS	signal peptidase I	NC	do	***	NC	do	****
SSRM_69_L	glyA	glycine hydroxymethyltransferase	NC	up	*	NC	up	*
SSRM_15_TL	MHP7448_0272	P97-like	NC	up	****	NC	up	***
SSRM_10_L	MHP7448_0108	P97-like copy 2	NC	-	ns	NC	up	*
SSRM_195_L	MHP7448_0505	lipoprotein	C	-	ns	NC	up	***
PALG_E_472[a]								
PALG_E_473[a]								
PALG_E_474[a]	MHP7448_0361	P29	100%	up	***	60%	-	ns
PALG_E_975[a]	MHP7448_0362	P69		up	**		-	ns
PALG_ES_134[a]								
PALG_E_1055								
PALG_E_1056								
PALG_E_572	MHP7448_0497	P76	60%	up	***	20%	-	ns
PALG_E_573								
PALG_ES_183								
PALG_E_291								
PALG_E_859	MHP7448_0108	P97	67%	-	ns	33%	up	*
PALG_S_1184								
PALG_E_381								
PALG_E_382								
PALG_E_383	MHP7448_0272	P97	83%	up	****	67%	up	***
PALG_E_385								
PALG_E_925								
PALG_E_926								
PALG_E_863	MHP7448_0107	P102	100%	NT	NT	0%	up	*
PALG_E_924	MHP7448_0271	P102	0%	up	**	0%	up	**
PALG_E_237[b]								
PALG_E_238[b]								
PALG_E_239[b]	MHP7448_0006	MgPa	100%	-	ns	60%	up	***
PALG_E_240[b]	MHP7448_0005	Mgpa		-	ns		up	*
PAL_ES_24[b]								

Asterisks indicate statistically significant differences in levels of expression, *0.01 < P < 0.05, **0.001 < P <0.01***, P < 0.001****. ns: non-significant. Conserved (C), Non-Conserved (NC) and

Absent (A) elements were analyzed in relation to upregulation (up), downregulation (do) or same (-) expression level in *M. hyopneumoniae* strain J (MHP_J) and *M. flocculare* (MFL) in relation to *M. hyopneumoniae* strain 7448 (MHP_7448) through qPCR assays. MHP7448_0107 ortholog was not found in *M. flocculare* genome. (a): PALG elements found in the 5' upstream region of the first gene of the transcriptional unit containg MHP7448_0361 and MHP7448_0362 CDS. (b): PALG elements found in the 5' upstream region of the first gene of the transcriptional unit containing MHP7448_0005 and MHP7448_0006 CDS. NT = not tested

Figure 1: Workflow of repeats search strategy.
Up to 500 bp of 5’ intergenic regions (IR) of all *M. hyopneumoniae* strain 7448 CDSs were used for *in silico* prediction of tandem and palindrome DNA repeats.

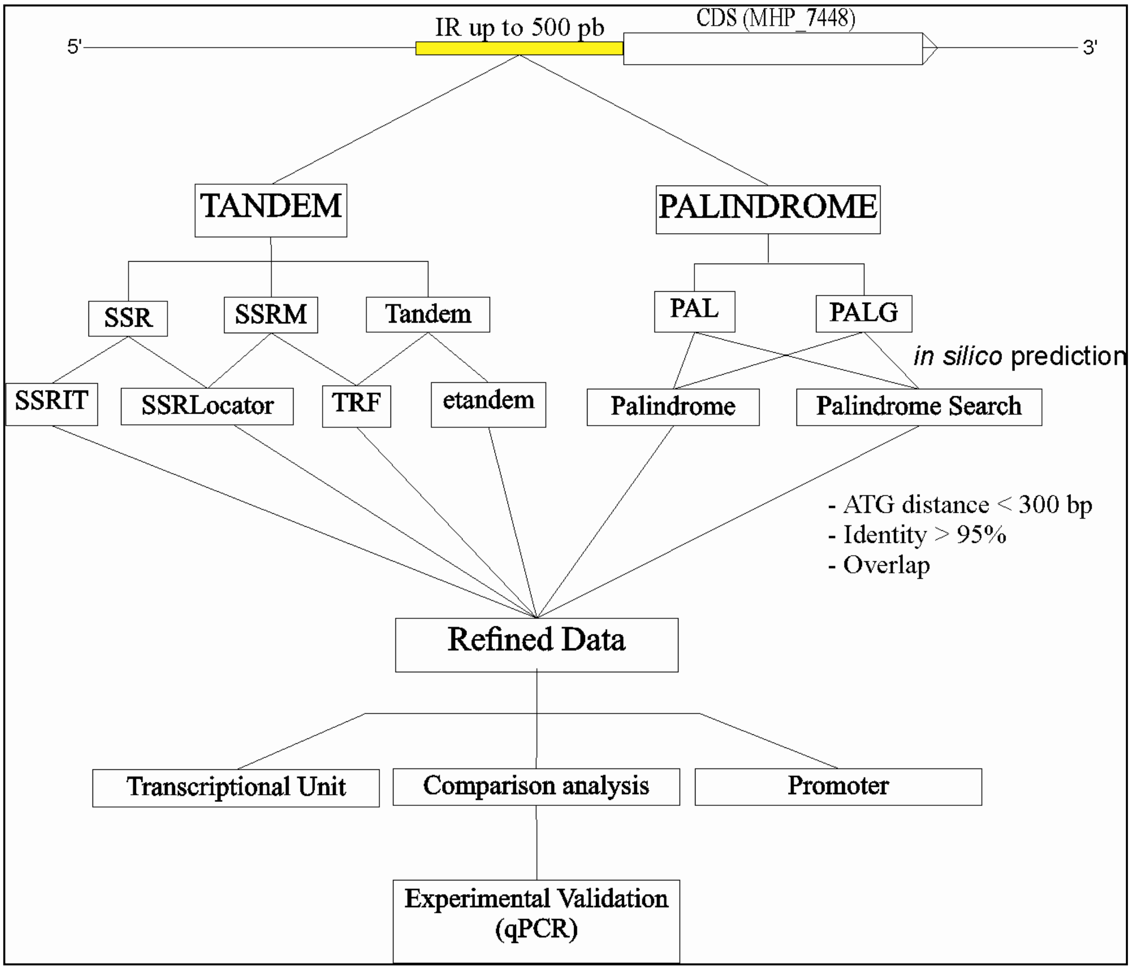


Figure 2: Pipeline of tandem repeats comparison analysis among mycoplasma genomes.
M. hyopneumoniae strain 7448 (MHP_7448), *M. hyopneumoniae* strain 7422 (MHP_7422), *M. hyopneumoniae* strain J (MHP_J) and *M. flocculare* (MFL).
An example of non-conserved SSR between MHP_7448 and MHP_J could be observed below. A $(TAT)_9$ repeat found in MHP7448 gene intergenic region was classified as non-conserved, as a $(TAT)_{11}$ repeat was localized in the respective ortholog intergenic region in MHP_J.

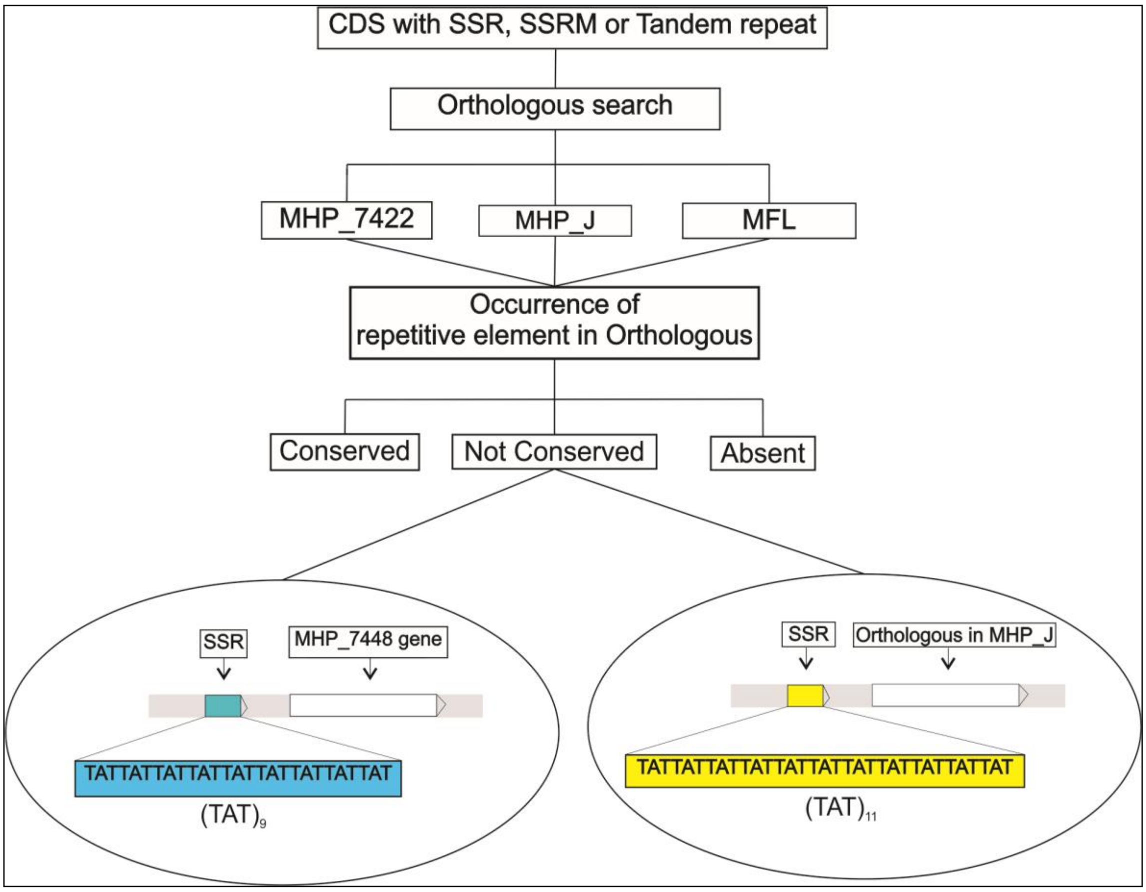


Figure 3: Workflow of repeats conservation on adhesin coding genes and differentially expressed genes from *Mycoplasma*.

A) Adhesins comparison analysis, with an example of relative conservation between MHP7448 gene that have 5 PALG repeats, compared to respective ortholog in MHP_J that have only 3. The percentage (%) of conservation between MHP_7448 and MHP_J repeats was obtained by dividing the number of repeat elements. **B)** Repeats present in differentially expressed (DE) genes found in *M. hyopneumoniae* strain 232 (MHP_232) were compared to MHP_7448 in the same way as in **A**. Abbreviations as in Fig. 2.

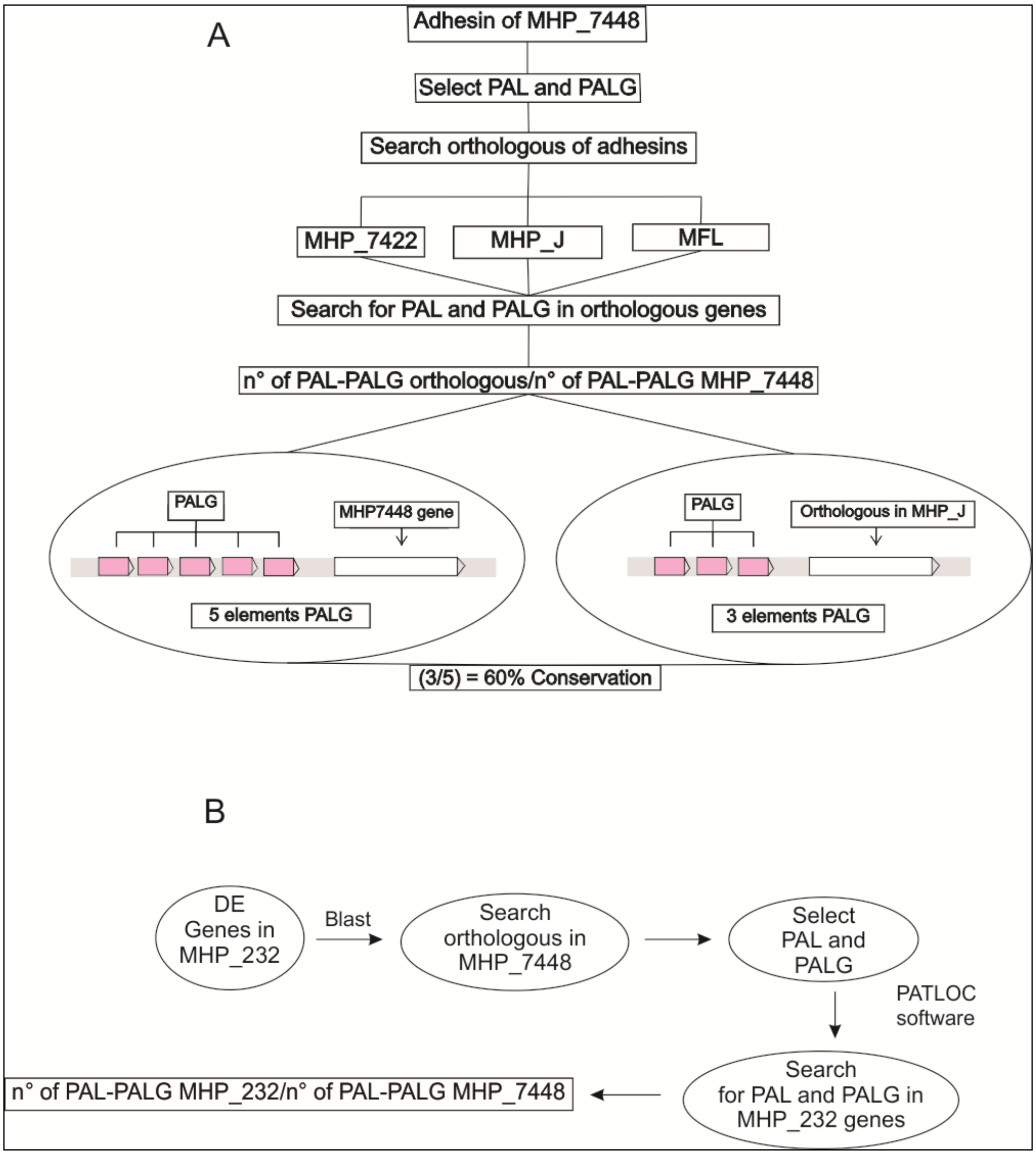


Figure 4: Repeat distribution in the intergenic regions of *M. hyopneumoniae* strain 7448.
All intergenic regions of CDS were analyzed in relation to the presence of tandem repeats, palindrome sequences or putative promoter motif (**A**). Tandem repeats and palindrome sequences were distributed in SSR, SSRM, tandem, PAL and PALG into CDS intergenic regions (**B**) The same was done to the 5’ upstream region of the first gene of each transcription units described in genome (**C**) and (**D**).

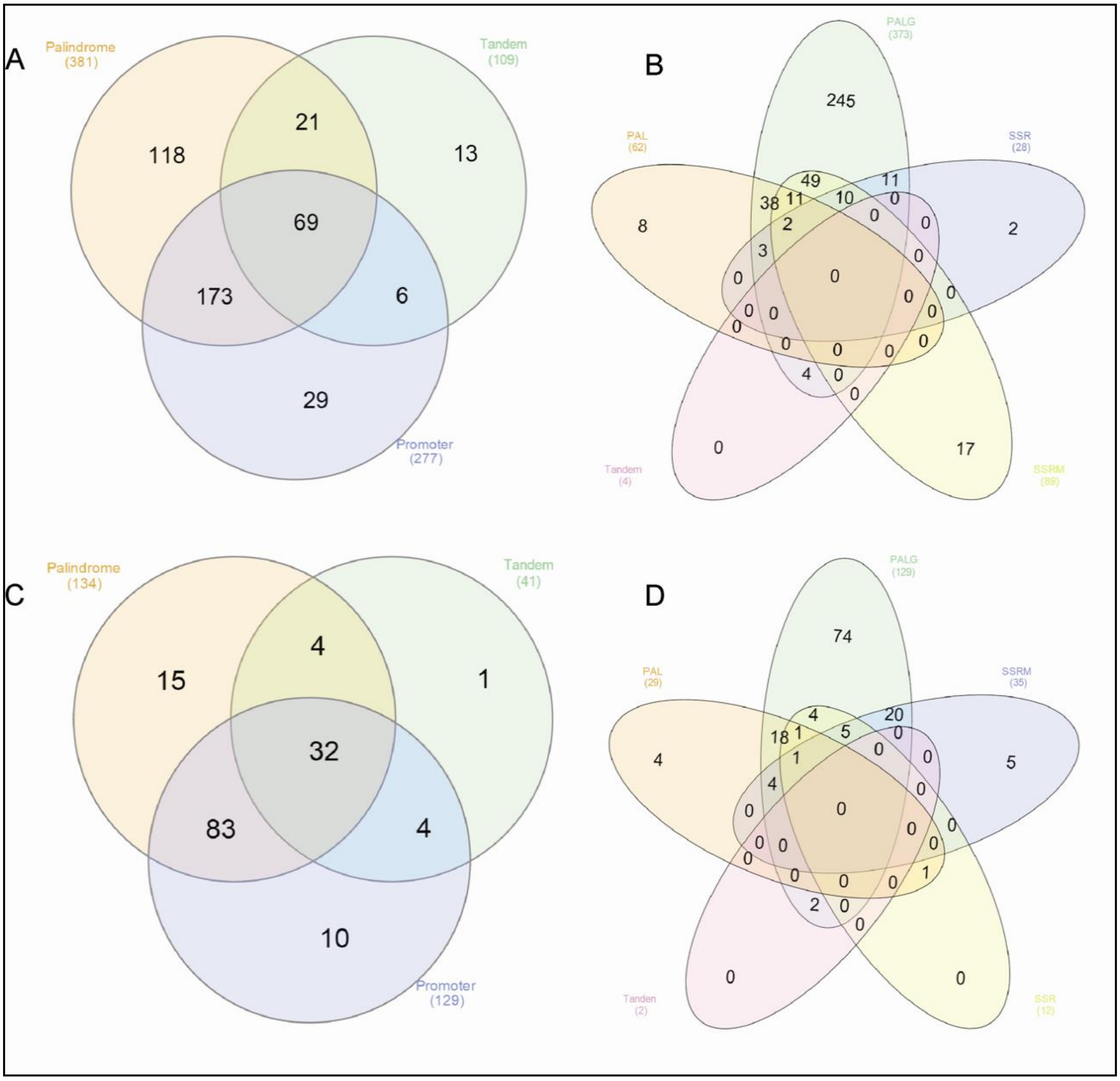


Figure 5: Conservation of repetitive elements in *Mycoplasma* genomes.

Graphics represent percent of Relative Conservation (%) of conserved (**A**), not conserved (**B**) and absent (**C**) SSR, SSRM, and tandem repeats found in *M. hyopneumoniae* strain 7448 intergenic regions in relation to strains 7422, J and *M. flocculare* (MFL) orthologs. A pathogenic versus non-pathogenic comparison was done and defined as P-nP. Repeats conserved between *M. hyopneumoniae* strain 7448 and *M. hyopneumoniae* strain 7422 were investigated in relation to MHP_J (P-nP(sp)) and MFL (P-nP(gn)). **D**) Graphic represents the average of all tandem repeats found in each tested situation.

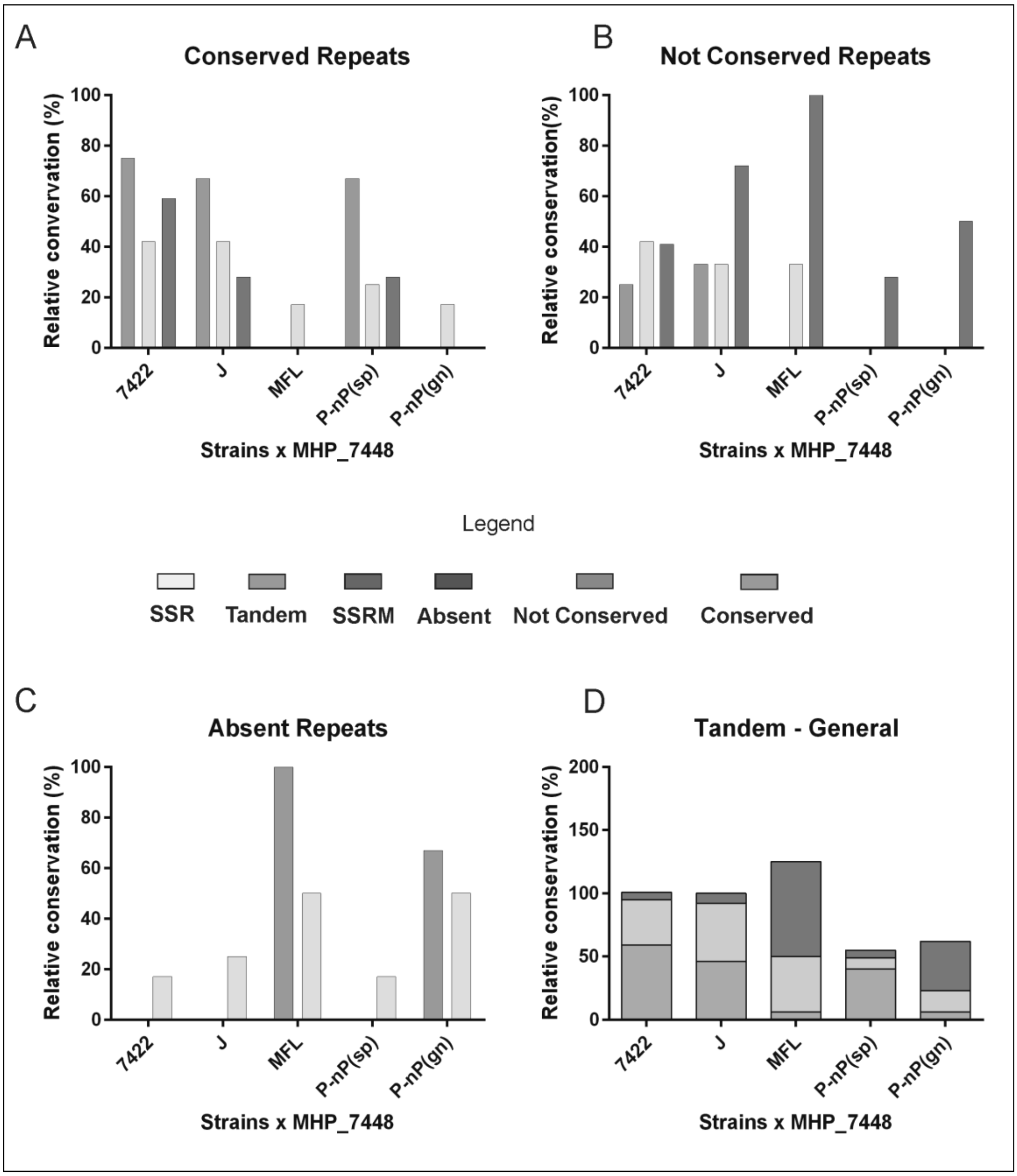
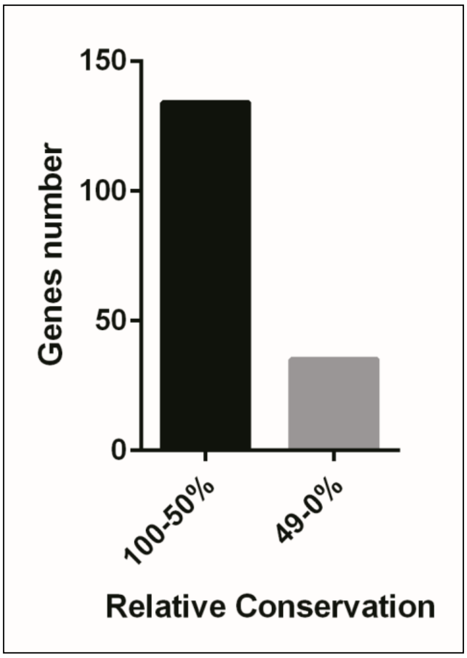


Figure 6: Repetitive element comparison of differentially expressed genes in *M. hyopneumoniae*. Comparison was performed with *M. hyopneumoniae* strain 232 versus *M. hyopneumoniae* strain 7448 orthologs. Black bar represent the number of ortholog genes that contain 50% conservation to totally conserved palindromes. Gray bar represent ortholog genes with partially or non-conserved palindromes elements.



4 DISCUSSÃO

O genoma procariótico apresenta uma grande diversidade de sequências, identificadas em diversas bactérias que já possuem seu genoma totalmente disponível (HUANG & MRAZEK, 2014), o que facilita estudos associativos entre padrões nucleotídicos encontrados e a respectiva funcionalidade dos mesmos. No presente trabalho, foi investigada a presença de repetições de DNA em regiões intergênicas (definidas como a região regulatória 5' distante até 500 pb do códon de início da tradução) do genoma de *M. hyopneumoniae* linhagem 7448, combinando tanto abordagens *in silico*, como *in vitro*. A escolha da região a ser analisada, levou em conta o objetivo de associar os elementos encontrados com um possível papel regulatório, uma vez que o conhecimento da regulação da transcrição nesse organismo é bastante limitado.

Os resultados preliminares das análises computacionais evidenciam um total de 2.210 elementos repetitivos. Após algumas etapas de filtragem e eliminação de sobreposições dos dados iniciais, o número de repetições totalizou 1.315, representando uma redução de 40%. Os 1.315 elementos estão distribuídos entre 144 repetições em tandem (SSR, SSRM e tandem). Os 1.171 elementos restantes se distribuíram entre repetições do tipo PAL e PALG, constituindo 89% do total geral. Essa alta representatividade de elementos palindromes nas regiões intergênicas de genomas bacterianos, foi também observada por WERBOWY *et al.* (2009) e HUANG & MRAZEK (2014) em um estudo englobando mais de 1.000 genomas procarióticos, incluindo micoplasmas, onde foram identificados números maiores de repetições palindromes do que elementos em tandem. Os dados encontrados corroboram os de outros autores e evidenciam a importância desses elementos em organismos bacterianos, indicando um possível nível de regulação simples e eficiente, uma vez que essas sequências formam estruturas secundárias estáveis e estão presentes em regiões próximas ao sítio de início da tradução (BRAZDA *et al.*, 2011). A formação de estruturas secundárias pode ser uma alternativa à presença de múltiplos sítios de ligação de proteínas, que despenderiam muitos esforços e gastos energéticos a um organismo com metabolismo limitado, como o *M. hyopneumoniae*.

As análises *in silico*, das diferentes classes de repetições, foram inicialmente realizadas de maneira independente e logo em seguida, os elementos foram agrupados. Nesta análise conjunta, verificou-se que muitas dessas sequências repetitivas se sobrepunham umas às outras, principalmente em relação as repetições do tipo PAL, que na maioria dos casos se encontravam sobrepostas a uma das metades de PALG. Portanto, as repetições PAL que se sobrepuseram a PALG foram eliminadas por não representarem uma repetição verdadeira. Além disso, outra característica importante que levou a seleção de repetições PALG em detrimento a elementos PAL foi a formação de estruturas secundárias mais estáveis em elementos PALG do que em repetições do tipo PAL, as quais podem representar interessantes modificações na molécula de DNA (BRAZDA *et al.*, 2011). Com base nesses critérios, 90% das repetições PALs foram eliminadas das análises posteriores.

Entre as sequências nucleotídicas dos elementos repetitivos das diferentes classes analisadas, não foi possível estabelecer motivos comuns dentro de cada grupo. No entanto, foi observado a presença de sequências ricas em AT, consequência de um genoma pobre em conteúdo GC (VASCONCELOS *et al.*, 2005). A ausência de motivos comuns foi inicialmente surpreendente, pois em um genoma reduzido, como o de *M. hyopneumoniae*, não foi prevista uma diversidade tão grande de sequências. Entretanto as repetições encontradas foram associadas com a formação de estruturas secundárias, a partir de análises de ΔG (Tabela suplementar S5 e S6). A degeneração de nucleotídeos a nível estrutural, não causa tanto efeito negativo em possíveis processos regulatórios quando comparado com sítios de ligação a proteínas regulatórias, uma vez que esse último, necessita interagir com aminoácidos específicos para que o processo seja ativado (MIKHEIKIN *et al.*, 2006).

Um estudo integrado entre localização gênica e a presença de DNA repetitivo revelou múltiplos elementos potencialmente regulatórios tanto nas regiões intergênicas de CDSs individuais, como em unidades transcricionais. Foi avaliada a presença de sequências promotoras da transcrição anteriormente descritas (WEBER *et al.*, 2012; SIQUEIRA *et al.*, 2014) juntamente com as repetições palindromes e tandem relatadas nesse trabalho. Através destas análises, foi possível concluir que em 70% das regiões intergênicas das CDSs, pelo menos um desses elementos estava presente. Na maioria dos casos identificados, as repetições do tipo palindrome, principalmente PALG, estavam associadas a presença de um promotor (Tabela Suplementar S2). Essa combinação de elementos potencialmente regulatórios é bastante interessante, pois, como já descrito, as estruturas secundárias formadas pelas sequências de repetição podem interferir na região promotora e consequentemente, reprimir ou ativar a transcrição (WERBOWY *et al.*, 2009). Em um número também expressivo de CDSs (24%) foi encontrado na região intergênica somente a presença de elementos do tipo palindrome, podendo estar relacionadas a algum processo regulatório, uma vez que nenhum promotor associado foi encontrado (Fig. 4). Em relação às unidades trancricionais, também foi observada a presença de uma repetição do tipo palindrome, principalmente PALG, associada a uma sequência promotora, na região intergênica referente ao primeiro gene da unidade transcricional. É interessante ressaltar que a presença de todos os elementos (tandem, palindrome e promotor) nas regiões intergênicas foi também observada (ver Fig. 4), sugerindo que ambos os cis-elementos podem interferir na regulação da transcrição de um único gene ou unidade trancricional, seja por interferência direta na região promotora (VINCES *et al.*, 2009) ou pela formação de estruturas secundárias (WERBOWY *et al.*, 2009).

A partir da descrição dos dados preditos pelas análises computacionais, a relevância das repetições de DNA encontradas foi avaliada a partir de estudos comparativos e ensaios experimentais. Estudos utilizando diferentes organismos relatam que repetições em tandem são mais abundantes em genes, cujos produtos, estão envolvidos na formação da estrutura da célula, como lipopolissacarideos (LPS), adesinas, pili, fimbrias e cápsula (MOXON *et al.*, 1994; JORDAN *et al.*, 2003; VERSTREPEN *et al.*, 2004; JANULCZYK *et al.*, 2010). Essas estruturas são relevantes em organismos patogênicos, uma vez que a diferente disponibilidade dessas proteínas na superfície celular pode servir como um importante mecanismo de escape da resposta do sistema imune do hospedeiro. Estes trabalhos sugerem que esta variabilidade de disponibilidade proteica está estritamente relacionada à presença de elementos repetitivos regulatórios.

Baseado na importância das repetições e sua relação com a patogenicidade de micro-organismos, uma análise comparativa dos elementos em tandem encontrados em *M. hyopneumoniae* linhagem 7448, foi realizada em relação a três outros micoplasmas. O genoma de outra linhagem (7422) de *M. hyopneumoniae* considerada patogênica foi utilizada na análise comparativa assim como uma linhagem (J) de *M. hyopneumoniae* não patogênica e uma espécie filogeneticamente relacionada a *M. hyopneumoniae* e não patogênica (*M. flocculare*) também foi utilizada. Os resultados evidenciaram uma maior conservação das sequências de repetição nas regiões regulatórias das CDSs ou unidades de transcrição entre genomas de espécies patogênicas do que não patogênicas. A relação entre a diversidade de motivos repetidos e sua influência na patogenicidade entre linhagens e espécies principalmente através do seu papel regulador na disponibilidade de proteínas antigênicas já foi descrito por MARTIN *et al*. (2005) e BAYLISS & PALMER (2012). A não conservação de elementos repetitivos entre linhagens e espécies se mostra um importante mecanismo de fuga do sistema imune do hospedeiro, e se torna um processo bastante importante no sucesso do estabelecimento da infecção (LIN & KUSSELL, 2012).

Considerando os resultados obtidos com as análises genômicas comparativas foram realizados ensaios experimentais (qPCR) utilizando genes que apresentaram elementos tandem distintos nas regiões intergênicas em relação aos genes de micoplasmas analisados. Agregando todos os resultados da validação experimental, foi possível concluir, que em geral, quando os elementos repetitivos são conservados nos genes ortólogos dos organismos analisados (gene MHP7448_0623, por exemplo), não é observada expressão diferencial entre eles. Porém, quando não há conservação ou há a ausência dos mesmos elementos de repetição nos genes ortólogos, é observada variação nos níveis transcricionais. Interessantemente, uma SSR encontrada na região intergênica do gene MHP7448_0087 de *M. hyopneumoniae* linhagem 7448 e do seu respectivo ortólogo em *M. hyopneumoniae* linhagem 7422 (patogênica) não foi encontrada nas regiões intergênicas dos ortólogos de *M. hyopneumoniae* linhagem J e *M. flocculare* (não patogênicas). O mesmo foi observado para a repetição SSRM (encontrada nos genes *glyA* e *sipS*), conservada entre ortólogos das linhagens patogênicas e não conservada entre ortólogos das linhagens não patogênicas. Os níveis de transcritos dos ortólogos de *M. hyopneumoniae* linhagem J e *M. flocculare* em relação a *M. hyopneumoniae* linhagem 7448 desses genes (MHP7448_0087, *glyA* e *sipS*) apresentaram diferenças significativas, sugerindo que algumas das repetições encontradas podem estar envolvidas na patogenicidade de micoplasmas. Esses dados estão em concordância com experimentos já realizados em *Neisseria* spp. (MARTIN *et al.*, 2005) e outros organismos (BAYLISS & PALMER, 2012) onde cópias distintas de uma repetição em tandem (SSR e SSRM), bem como a sua ausência são capazes de interferir na expressão gênica e podem estar envolvidas na variação antigênica, entre linhagens e espécies.

Considerando a importância das adesinas para as células bacterianas e os diversos trabalhos que indicam a presença de elementos repetitivos influenciando a sua expressão (MOXON *et al.*, 1994; JORDAN *et al.*, 2003; VERSTREPEN *et al.*, 2004; JANULCZYK *et al.*, 2010), os genes codificantes para adesinas já descritos em *M. hyopneumoniae* linhagem 7448 (FERREIRA *et al.*, 2007), foram analisados quanto à presença de elementos palindromes nas suas respectivas regiões intergênicas. Inicialmente foi possível observar vários elementos repetitivos em uma mesma região intergênica, corroborando com dados de MOXON *et al.*, (1994), que mostra uma maior presença de repetições de DNA em genes que codificam proteínas de superfície bacterianas. A conservação dos elementos palindromes encontrados nesses genes foi investigada nos mesmos organismos utilizados para as análises comparativas das repetições em tandem. Com os resultados foi possível observar uma maior conservação em *M. hyopneumoniae* linhagem 7422 (87%), do que em *M. hyopneumoniae* linhagem J (69%) e *M. flocculare* (27%) se comparados com *M. hyopneumoniae* linhagem 7448. Portanto, podemos concluir que entre linhagens patogênicas existe uma maior conservação entre os elementos repetitivos analisados, se comparados aos micoplasmas não patogênicos. Variações no número de repetições envolvidas com a patogenicidade de bactérias já foram anteriormente descritas (MOXON *et al.*, 1994).

A avaliação dos índices de expressão de genes codificantes para adesinas que possuíam elementos repetitivos distintos nas suas regiões intergênicas foi realizada também através de ensaios de qPCR. Os resultados das análises experimentais sugeriram que a expressão dos genes codificantes para adesinas pode estar relacionada com a presença ou ausência de elementos palindromes. Diferenças no número de repetições palindromes e nos níveis de transcrição entre os genes ortólogos dos micoplasmas avaliados foram observados, principalmente nos genes MHP7448_0271 (codificante da adesina P102), MHP7448_0272 (adesina P97) e nos genes MHP7448_0005 e MHP7448_0006, codificantes para proteínas MgPa. A diferença de expressão de genes que codificam proteínas adesinas, dependendo da presença de repetições em tandem já foi descrita em diferentes organismos como *Haemophilus influenzae* (SHEETS & GEME, 2011) *Helicobacter pylori* (GOODWIN *et al.*, 2008), *Neisseria gonorrhoeae* (MURPHY *el al.*,1989) e em *Mycoplasma hominis* (ZHANG & WISE, 1997). No entanto, este é o primeiro trabalho relatando a relação de expressão diferencial de genes que codificam adesinas com a presença de repetições palindromes.

Com o objetivo de aprofundar os estudos relativos aos resultados obtidos com as análises comparativas e com os dados de análise transcricional, foram avaliados genes de *M. hyopneumoniae* que já mostraram expressão diferencial em diferentes condições de cultivo. Este trabalho foi realizado baseado na hipótese que as repetições presentes nas regiões regulatórias dos genes podem estar interferindo, de uma maneira direta ou indireta, na expressão de genes importantes de *M. hyopneumoniae* linhagem 7448. Portanto, foi realizado um estudo comparativo avaliando genes anteriormente descritos como diferencialmente expressos em *M. hyopneumoniae* linhagem 232 quando expostos a diferentes situações como choque térmico, estresse oxidativo, exposição ao ferro e noraepinefrina e em condições de infecção *in vitro* (MADSEN *et al.*, 2006; MADSEN *et al.*, 2006; SCHAFER *et al.*, 2007; MADSEN *et al.*, 2008; ONEAL *et al.*, 2008). A comparação entre genes ortólogos entre as duas linhagens de *M. hyopneumoniae* (232 e 7448) revelou altos índices de conservação, onde 80% das repetições palindromes se encontram presentes na região regulatória dos genes ortólogos de ambas as linhagens, reforçando a importância desses elementos.

Os estudos de descrição, caracterização e associação da presença de repetições de DNA do tipo tandem e palindrome em regiões intergênicas do genoma de *M. hyopneumoniae* linhagem 7448, demonstram a importância das repetições de DNA nesse organismo e, evidenciam a sua relação com a regulação da transcrição. Diferenças entre linhagens patogênicas e não patogênicas, bem como a validação experimental e a demonstração da presença de repetições em genes que são susceptíveis a regulação, ajudam a inferir esse possível papel regulador em *M. hyopneumoniae*. O presente estudo apresenta dados relevantes que podem servir como base para ensaios posteriores capazes de confirmar e aprimorar os resultados encontrados, contribuindo para a elucidação de mecanismos regulatórios ainda desconhecidos nesta espécie bacteriana.

5 PERSPECTIVAS

Como perspectivas importantes a serem ainda trabalhadas, podemos destacar:

- Ampliar os tipos de elementos repetitivos a serem preditos nos genomas, como por exemplo a busca de CRISPRS, padrões relacionados a H-DNA, G-DNA ou Z-DNA, repetições complementares, e repetições em tandem possuindo *gaps*.

- Expandir as regiões intergênicas a serem utilizadas como *input* e *output* dos *softwares,* tanto à montante como à jusante do sítio de início da tradução.

- Realizar ensaios de alteração de mobilidade eletroforética (EMSA) para avaliar se os elementos repetitivos, principalmente do tipo PAL e PALG poderiam ser sítios de ligação de proteínas regulatórias.

- Analisar comparativamente mais espécies relacionas a *M. hyopneumoniae*, em relação ao DNA repetitivo encontrado, incluindo análises de filogenia para o entendimento da origem dos elementos repetitivos.

- Relacionar os resultados descritos com dados já publicados de transcriptomas de micoplasmas do trato respiratório de suínos (SIQUEIRA *et al.*, 2014).

- Realizar um mapa completo com todos os elementos repetitivos encontrados no genoma de *M. hyopneumoniae* linhagem 7448, tanto em regiões intergênicas, quanto em regiões codificantes, tentando sempre associar com mudanças estruturais no cromossomo bacteriano que possam estar relacionados com processos regulatórios.

6 REFERÊNCIAS

BARATE, A. K. et al. Immunogenicity of IMS 1113 plus soluble subunit and chimeric proteins containing *Mycoplasma hyopneumoniae* P97 C-terminal repeat regions. FEMS Microbiol Lett, v. 352, n. 2, p. 213-20, Mar 2014.

BAYLISS, C. D.; PALMER, M. E. Evolution of simple sequence repeat-mediated phase variation in bacterial genomes. Ann N Y Acad Sci, v. 1267, p. 39-44, Sep 2012.

BRAZDA, V. et al. Cruciform structures are a common DNA feature important for regulating biological processes. BMC Mol Biol, v. 12, p. 33, 2011.

CALCUTT, M. J. et al. Complete Genome Sequence of *Mycoplasma flocculare* Strain Ms42T (ATCC 27399T). Genome Announc, v. 3, n. 2, 2015.

CASJENS, S. Evolution of the linear DNA replicons of the *Borrelia* spirochetes. Curr Opin Microbiol, v. 2, n. 5, p. 529-34, Oct 1999.

CHANG, L. J. et al. Mycoplasmas regulate the expression of heat-shock protein genes through CIRCE-HrcA interactions. Biochem Biophys Res Commun, v. 367, n. 1, p. 213-8, Feb 29 2008.

DEBEY, M. C.; ROSS, R. F. Ciliostasis and loss of cilia induced by *Mycoplasma hyopneumoniae* in porcine tracheal organ cultures. Infect Immun, v. 62, n. 12, p. 5312-8, Dec 1994

FERREIRA, H. B. et al. A preliminary survey of *M. hyopneumoniae* virulence factors based on comparative genomic analysis. Genet. Mol. Biol., v. 30, n. 1, p. 245-255, 00/2007 2007.

FRIIS, N. F.; FEENSTRA, A. A. *Mycoplasma hyorhinis* in the etiology of serositis among piglets. Acta Vet Scand, v. 35, n. 1, p. 93-8, 1994.

FRITSCH, T. E.; SIQUEIRA, F. M.; SCHRANK, I. S. Intrinsic terminators in *Mycoplasma hyopneumoniae* transcription. BMC Genomics, v. 16, p. 273, 2015.

GARDNER, S. W.; MINION, F. C. Detection and quantification of intergenic transcription in *Mycoplasma hyopneumoniae*. Microbiology, v. 156, n. Pt 8, p. 2305-15, Aug 2010.

GEMAYEL, R. et al. Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes (Basel), v. 3, n. 3, p. 461-80, 2012.

GOODWIN, A. C. et al. Expression of the *Helicobacter pylori* adhesin SabA is controlled via phase variation and the ArsRS signal transduction system. Microbiology, v. 154, n. Pt 8, p. 2231-40, Aug 2008.

GOSNEY, F. et al. Variable number tandem repeat (VNTR) typing of strains of *Mycoplasma mycoides* subspecies mycoides small colony isolated from the north-eastern regions of Italy between 1990 and 1993. Vet Microbiol, v. 147, n. 1-2, p. 220-2, Jan 10 2011

HALBEDEL, S. et al. Transcription in *Mycoplasma pneumoniae*: analysis of the promoters of the ackA and ldh genes. J Mol Biol, v. 371, n. 3, p. 596-607, Aug 17 2007.

HENDERSON, I. R.; OWEN, P.; NATARO, J. P. Molecular switches--the ON and OFF of bacterial phase variation. Mol Microbiol, v. 33, n. 5, p. 919-32, Sep 1999.

HUANG, Y.; MRAZEK, J. Assessing diversity of DNA structure-related sequence features in prokaryotic genomes. DNA Res, v. 21, n. 3, p. 285-97, Jun 2014. ISSN 1340-2838.

JACOB, F. Evolution and tinkering. Science, v. 196, n. 4295, p. 1161-6, Jun 10 1977.

JANULCZYK, R. et al. Simple sequence repeats and genome plasticity in *Streptococcus agalactiae*. J Bacteriol, v. 192, n. 15, p. 3990-4000, Aug 2010.

JORDAN, P.; SNYDER, L. A.; SAUNDERS, N. J. Diversity in coding tandem repeats in related *Neisseria* spp. BMC Microbiol, v. 3, p. 23, Nov 12 2003.

KAWASHIMA, K.; KATSUDA, K.; TSUNEMITSU, H. Epidemiological investigation of the prevalence and features of postweaning multisystemic wasting syndrome in Japan. J Vet Diagn Invest, v. 19, n. 1, p. 60-8, 2007.

KOBAYASHI, H. et al. *Mycoplasma hyorhinis* infection levels in lungs of piglets with porcine reproductive and respiratory syndrome (PRRS). J Vet Med Sci, v. 58, n. 2, p. 109-13, 1996.

KOBISCH, M.; FRIIS, N. F. Swine mycoplasmoses. Rev Sci Tech, v. 15, n. 4, p. 1569-605, Dec 1996.

LIN, W. H.; KUSSELL, E. Evolutionary pressures on simple sequence repeats in prokaryotic coding regions. Nucleic Acids Res, v. 40, n. 6, p. 2399-413, Mar 2012.

LIU, L. et al. GAA trinucleotide repeat region regulates M9/pMGA gene expression in *Mycoplasma gallisepticum*. Infect Immun, v. 68, n. 2, p. 871-6, Feb 2000.

LIU, W. et al. Complete genome sequence of *Mycoplasma hyopneumoniae* strain 168. J Bacteriol, v. 193, n. 4, p. 1016-7, Feb 2011.

______. Comparative genomic analyses of *Mycoplasma hyopneumoniae* pathogenic 168 strain and its high-passaged attenuated strain. BMC Genomics, v. 14, p. 80, 2013.

LLORENS-RICO, V.; LLUCH-SENAR, M.; SERRANO, L. Distinguishing between productive and abortive promoters using a random forest classifier in *Mycoplasma pneumoniae*. Nucleic Acids Res, v. 43, n. 7, p. 3442-53, Apr 20 2015.

LYSNYANSKY, I.; ROSENGARTEN, R.; YOGEV, D. Phenotypic switching of variable surface lipoproteins in *Mycoplasma bovis* involves high-frequency chromosomal rearrangements. J Bacteriol, v. 178, n. 18, p. 5395-401, Sep 1996.

MA, L. et al. Genetic variation in the complete MgPa operon and its repetitive chromosomal elements in clinical strains of *Mycoplasma genitalium*. PLoS One, v. 5, n. 12, p. e15660, 2010.

MADSEN, M. L. et al. Transcriptional profiling of *Mycoplasma hyopneumoniae* during heat shock using microarrays. Infect Immun, v. 74, n. 1, p. 160-6, Jan 2006.

______. Transcriptional profiling of *Mycoplasma hyopneumoniae* during iron depletion using microarrays. Microbiology, v. 152, n. Pt 4, p. 937-44, Apr 2006.

______. Transcriptome changes in *Mycoplasma hyopneumoniae* during infection. Infect Immun, v. 76, n. 2, p. 658-63, Feb 2008.

MAES, D. et al. Control of *Mycoplasma hyopneumoniae* infections in pigs. Vet Microbiol, v. 126, n. 4, p. 297-309, Jan 25 2008.

MANI, P. et al. Genome-wide analyses of recombination prone regions predict role of DNA structural motif in recombination. PLoS One, v. 4, n. 2, p. e4399, 2009.

MARCOTTE, E. M. et al. A census of protein repeats. J Mol Biol, v. 293, n. 1, p. 151-60, Oct 15 1999.

MARE, C. J.; SWITZER, W. P. New species: *Mycoplasma hyopneumoniae*; a causative agent of virus pig pneumonia. Vet Med Small Anim Clin, v. 60, p. 841-6, Aug 1965.

MARTIN, P. et al. Microsatellite instability regulates transcription factor binding and gene expression. In: (Ed.). Proc Natl Acad Sci U S A, v.102, 2005. p.3800-4.

MARTINEZ, H. M. An efficient method for finding repeats in molecular sequences. Nucleic Acids Res, v. 11, n. 13, p. 4629-34, Jul 11 1983.

METRUCCIO, M. M. et al. A novel phase variation mechanism in the meningococcus driven by a ligand-responsive repressor and differential spacing of distal promoter elements. PLoS Pathog, v. 5, n. 12, p. e1000710, Dec 2009.

MEYLING, A.; FRIIS, N. F. Serological identification of a new porcine mycoplasma species, M. flocculare. Acta Vet Scand, v. 13, n. 2, p. 287-9, 1972.

MIKHEIKIN, A. L.; LUSHNIKOV, A. Y.; LYUBCHENKO, Y. L. Effect of DNA supercoiling on the geometry of holliday junctions. Biochemistry, v. 45, n. 43, p. 12998-3006, Oct 31 2006.

MINION, F. C. et al. The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis. In: (Ed.). J Bacteriol. United States, v.186, 2004. p.7123-33.

MORITA, T. et al. Demonstration of *Mycoplasma hyorhinis* as a possible primary pathogen for porcine otitis media. Vet Pathol, v. 32, n. 2, p. 107-11, 1995.

MOXON, E. R. et al. Adaptive evolution of highly mutable loci in pathogenic bacteria. Curr Biol, v. 4, n. 1, p. 24-33, Jan 1 1994.

MRAZEK, J. Analysis of distribution indicates diverse functions of simple sequence repeats in Mycoplasma genomes. Mol Biol Evol, v. 23, n. 7, p. 1370-85, Jul 2006

MURPHY, G. L. et al. Phase variation of gonococcal protein II: regulation of gene expression by slipped-strand mispairing of a repetitive DNA sequence. Cell, v. 56, n. 4, p. 539-47, Feb 24 1989.

MUSATOVOVA, O.; KANNAN, T. R.; BASEMAN, J. B. *Mycoplasma pneumoniae* large DNA repetitive elements RepMP1 show type specific organization among strains. PLoS One, v. 7, n. 10, p. e47625, 2012.

NAAS, T. et al. Dynamics of IS-related genetic rearrangements in resting *Escherichia coli* K-12. Mol Biol Evol, v. 12, n. 2, p. 198-207, Mar 1995.

NATHUES, H. et al. Individual risk factors for *Mycoplasma hyopneumoniae* infections in suckling pigs at the age of weaning. Acta Vet Scand, v. 55, n. 1, p. 44, Jun 3 2013.

NIAUDET, B.; JANNIERE, L.; EHRLICH, S. D. Integration of linear, heterologous DNA molecules into the *Bacillus subtilis* chromosome: mechanism and use in induction of predictable rearrangements. J Bacteriol, v. 163, n. 1, p. 111-20, Jul 1985.

NOORMOHAMMADI, A. H. et al. *Mycoplasma synoviae* has two distinct phase-variable major membrane antigens, one of which is a putative hemagglutinin. Infect Immun, v. 65, n. 7, p. 2542-7, Jul 1997.

ONEAL, M. J. et al. Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to norepinephrine. Microbiology, v. 154, n. Pt 9, p. 2581-8, Sep 2008.

PALMER, G. H.; BRAYTON, K. A. Gene conversion is a convergent strategy for pathogen antigenic variation. Trends Parasitol, v. 23, n. 9, p. 408-13, Sep 2007.

PETER, B. J. et al. Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. Genome Biol, v. 5, n. 11, p. R87, 2004.

POLLACK, J. D.; TRYON, V. V.; BEAMAN, K. D. The metabolic pathways of *Acholeplasma* and *Mycoplasma*: an overview. Yale J Biol Med, v. 56, n. 5-6, p. 709-16, Sep-Dec 1983.

RAZIN, S. Peculiar properties of mycoplasmas: the smallest self-replicating prokaryotes. FEMS Microbiol Lett, v. 79, n. 1-3, p. 423-31, Dec 15 1992.

RAZIN, S.; HAYFLICK, L. Highlights of mycoplasma research--an historical perspective. Biologicals, v. 38, n. 2, p. 183-90, Mar 2010.

RAZIN, S.; YOGEV, D.; NAOT, Y. Molecular biology and pathogenicity of mycoplasmas. Microbiol Mol Biol Rev, v. 62, n. 4, p. 1094-156, Dec 1998.

SCHAFER, E. R. et al. Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to hydrogen peroxide. Microbiology, v. 153, n. Pt 11, p. 3785-90, Nov 2007.

SHEETS, A. J.; ST GEME, J. W., 3RD. Adhesive activity of the *haemophilus cryptic* geno species cha autotransporter is modulated by variation in tandem peptide repeats. J Bacteriol, v. 193, n. 2, p. 329-39, Jan 2011.

SIMIONATTO, S. et al. *Mycoplasma hyopneumoniae*: From disease to vaccine development. Vet Microbiol, v. 165, n. 3-4, p. 234-42, Aug 30 2013.

SIMMONS, W. L.; DENISON, A. M.; DYBVIG, K. Resistance of Mycoplasma pulmonis to complement lysis is dependent on the number of Vsa tandem repeats: shield hypothesis. Infect Immun, v. 72, n. 12, p. 6846-51, Dec 2004

SIQUEIRA, F. M. et al. Genome organization in *Mycoplasma hyopneumoniae*: identification of promoter-like sequences. Mol Biol Rep, v. 41, n. 8, p. 5395-402, Aug 2014.

SIQUEIRA, F. M. et al. Unravelling the transcriptome profile of the Swine respiratory tract mycoplasmas. PLoS One, v. 9, n. 10, p. e110327, 2014.

SIQUEIRA, F. M.; SCHRANK, A.; SCHRANK, I. S. *Mycoplasma hyopneumoniae* transcription unit organization: genome survey and prediction. DNA Res, v. 18, n. 6, p. 413-22, Dec 2011.

SIQUEIRA, F. M. et al. New insights on the biology of swine respiratory tract mycoplasmas from a comparative genome analysis. BMC Genomics, v. 14, p. 175, 2013.

STEMKE, G. W. et al. Phylogenetic relationships of three porcine mycoplasmas, *Mycoplasma hyopneumoniae*, *Mycoplasma flocculare*, and *Mycoplasma hyorhinis*, and complete 16S rRNA sequence of *M. flocculare*. Int J Syst Bacteriol, v. 42, n. 2, p. 220-5, Apr 1992.

THACKER, E. L. et al. *Mycoplasma hyopneumoniae* potentiation of porcine reproductive and respiratory syndrome virus-induced pneumonia. J Clin Microbiol, v. 37, n. 3, p. 620-7, Mar 1999.

THACKER E.L.; MINION F.C. Mycoplasmosis. In: Zimmerman JJ, Karriker LA, Ramirez A, Schwartz KJ, Stevenson GW, editors. Diseases of Swine. 10th ed. Ames, Iowa: Wiley-Blackwell; 2012. pp. 779–797.

THEISS, P.; WISE, K. S. Localized frameshift mutation generates selective, high-frequency phase variation of a surface lipoprotein encoded by a mycoplasma ABC transporter operon. J Bacteriol, v. 179, n. 12, p. 4013-22, Jun 1997.

TREANGEN, T. J. et al. Genesis, effects and fates of repeats in prokaryotic genomes. FEMS Microbiol Rev, v. 33, n. 3, p. 539-71, May 2009.

TREANGEN, T. J.; SALZBERG, S. L. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet, v. 13, n. 1, p. 36-46, Jan 2012.

VAN DER ENDE, A.; HOPMAN, C. T.; DANKERT, J. Multiple mechanisms of phase variation of PorA in *Neisseria meningitidis*. Infect Immun, v. 68, n. 12, p. 6685-90, Dec 2000.

VASCONCELOS, A. T. et al. Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. J Bacteriol, v.187, 2005. p.5568-77.

VERSTREPEN, K. J.; REYNOLDS, T. B.; FINK, G. R. Origins of variation in the fungal cell surface. Nat Rev Microbiol, v. 2, n. 7, p. 533-40, Jul 2004.

VICCA, J. et al. In vitro susceptibilities of *Mycoplasma hyopneumoniae* field isolates. Antimicrob Agents Chemother, v. 48, n. 11, p. 4470-2, Nov 2004

VILLARREAL, I. et al. Effect of vaccination of pigs against experimental infection with high and low virulence *Mycoplasma hyopneumoniae* strains. Vaccine, v. 29, n. 9, p. 1731-5, Feb 2011.

VINCES, M. D. et al. Unstable tandem repeats in promoters confer transcriptional evolvability. Science, v. 324, n. 5931, p. 1213-6, May 29 2009.

VOGEL, C. et al. Structure, function and evolution of multidomain proteins. Curr Opin Struct Biol, v. 14, n. 2, p. 208-16, Apr 2004.

WEBER SDE, S.; SANT'ANNA, F. H.; SCHRANK, I. S. Unveiling *Mycoplasma hyopneumoniae* promoters: sequence definition and genomic distribution. DNA Res, v. 19, n. 2, p. 103-15, Apr 2012.

WEINER, J., 3RD; HERRMANN, R.; BROWNING, G. F. Transcription in *Mycoplasma pneumoniae*. Nucleic Acids Res, v. 28, n. 22, p. 4488-96, Nov 15 2000.

WERBOWY, K.; CIESLINSKI, H.; KUR, J. Characterization of a cryptic plasmid pSFKW33 from *Shewanella* sp. 33B. Plasmid, v. 62, n. 1, p. 44-9, Jul 2009.

YOGEV, D. et al. Molecular basis of *Mycoplasma* surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5' regulatory sequences. Embo j, v. 10, n. 13, p. 4069-79, Dec 1991.

YOUNG, T. F. et al. A tissue culture system to study respiratory ciliary epithelial adherence of selected swine mycoplasmas. Vet Microbiol, v. 71, n. 3-4, p. 269-79, Feb 2000.

ZHANG, Q.; WISE, K. S. Molecular basis of size and antigenic variation of a *Mycoplasma hominis* adhesin encoded by divergent vaa genes. Infect Immun, v. 64, n. 7, p. 2737-44, Jul 1996.

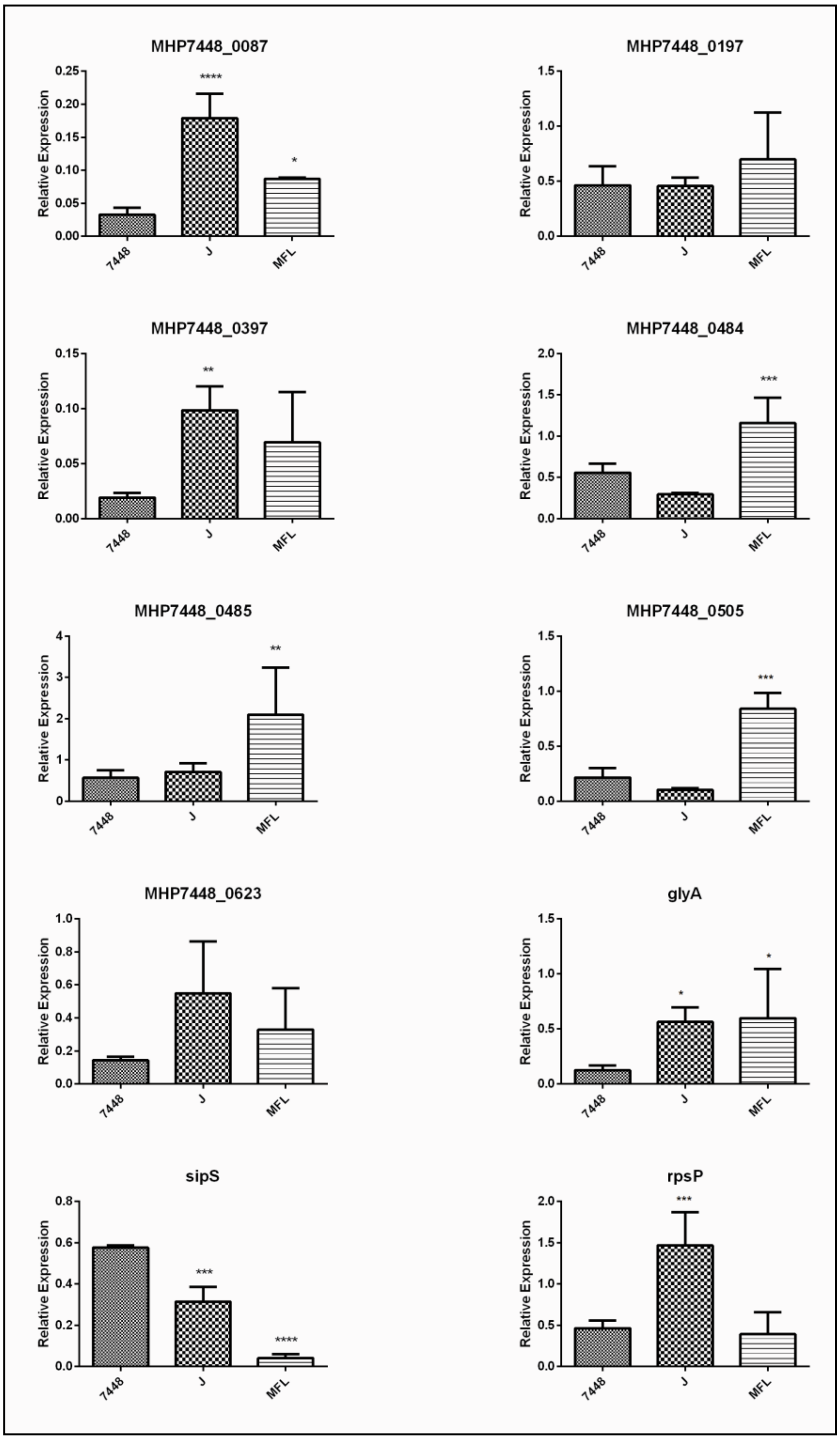
______. Localized reversible frameshift mutation in an adhesin gene confers a phase-variable adherence phenotype in mycoplasma. Mol Microbiol, v. 25, n. 5, p. 859-69, Sep 1997.

ZHANG, Q.; YOUNG, T. F.; ROSS, R. F. Identification and characterization of a *Mycoplasma hyopneumoniae* adhesin. Infect Immun, v. 63, n. 3, p. 1013-9, Mar 1995.

ZIELINSKI, G. C.; ROSS, R. F. Effect of growth in cell cultures and strain on virulence of *Mycoplasma hyopneumoniae* for swine. Am J Vet Res, v. 51, n. 3, p. 344-8, Mar 1990.

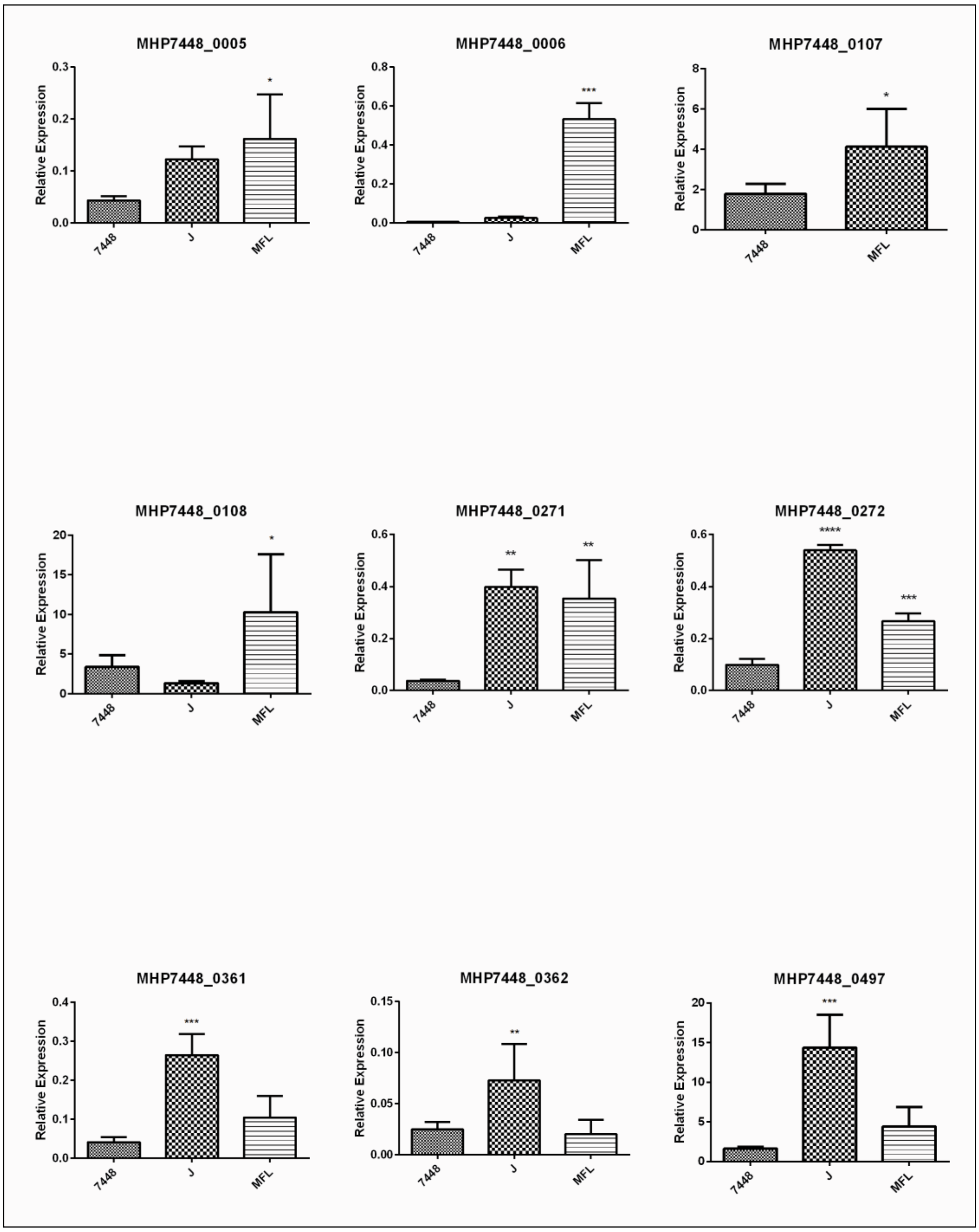
7 ANEXOS

Supplementary Figure S1: Experimental validation through qPCR assay of tandem elements comparison.



Asterisks indicate statistically significant differences in levels of expression, *0.01 < P < 0.05, **0.001 < P <0.01, ***P < 0.001. Abbreviations: *M. hyopneumoniae* strain 7448 (7448); *M. hyopneumoniae* strain J (J) and *M. flocculare* (MFL).

Supplementary Figure S2: Experimental validation through qPCR assay of palindrome elements comparison in adhesin coding genes.



Asterisks indicate statistically significant differences in levels of expression, *0.01 < P < 0.05, **0.001 < P <0.01, ***P < 0.001. Abbreviations: *M. hyopneumoniae* strain 7448 (7448); *M. hyopneumoniae* strain J (J) and *M. flocculare* (MFL).

Supplementary Table S1: Primers used in qPCR assays

Target (gene_name)	Primer_name	Sequence (5´- 3´)	Lenght (bp)	TM (°C)	Product (bp)	Analysis
MYF_00025	005 _MFL_F	GATCCAACTTTTTCTGCCACTTCA	24	56.0	156	A
	005 _MFL_R	CGCAAATGTTCTTGACGACGTTG	23	57.3		
MHP7448_0005	005 F	GAACAGAACAAGAAGTCGCTTTTT	24	54.3	160	A
	005 R	GCATCGCCAATGTTTGCCAAA	21	57.6		
MYF_00030	006 _MFL_F	CCTTCAAATTCTGTACGCCCCAAT	24	57.7	162	A
	006 _MFL_R	TAAATGACCGCCACCGCCTCA	21	61.4		
MHP7448_0006	006 F	GACCAGACGGGGATTGTATTGG	22	57.6	184	A
	006 R	GCATCAACAATTATCGCCAAAGAA	24	54.5		
MYF_00550	107 _MFL_F	GATGCAAAGCCAACAAATCCGAGC	24	59.5	172	A
	107 _MFL_R	GCGTCTCTTTTGGACTTGCTGGA	23	59.7		
MHP7448_0107	107 F	AGCTTTATTCACTACCTTCAGCAC	24	55.0	165	A
	107 R	GTTGTTGGGCTGGCTTCTGAC	21	59.0		
MHP7448_0108	108 F	CGGCGTCAGTTCTGCCAAA	19	58.6	173	A/T
	108 R	CTCCGATTTTGCTCCATCTCGA	22	56.9		
MHJ_0105	108_J F	AATCAGAAACAGCTAAACCCGTGG	24	57.6	155	A/T
	108_J R	CTCCGATTTTGCTCCATCACGA	22	57.4		
MYF_00575	108_MFL F	TGAAAGCGGAATTGGATTTTATGG	24	54.2	168	A/T
	108_MFL R	GAGACTGCTTGTTGTGACCCG	21	58.3		
MYF_00545	197_MFL F	GCAATCTTTCGGCGTAAAATACAG	24	54.8	175	T
	197_MFL R	TTAAGGCGGTGGCAATCG	18	55.8		
MHP7448_0197	197F	TCGGAGAAAATTACAAAAACGCTC	24	54.1	168	T
	197R	CTAAGAACGCTGCTAAGGGCAG	22	57.9		
MYF_01800	271 _MFL_F	TCACAGTAGTTTTGATGTTGCAGC	24	56.0	167	A
	271 _MFL_R	CTTGGGTCACTGTATTTCTTTGCG	24	56.4		
MHP7448_0271	271 F	CATTTTCACGGCTTGATTCAGGGC	24	59.1	174	A
	271 R	GTCCTGCAATTCCTTGGATTGTGG	24	58.5		
MHP7448_0272	272 F	TCTAACTTGATTTTACGAGCACGC	24	55.5	163	A/T
	272 R	GCTCAGTTAATTTTGCCTCAAGAT	24	54.0		
MYF_01805	272_MFL F	GAAAATTAGCGGGTTTGGGCA	21	56.4	182	A/T
	272_MFL R	TTTCTGCAACAATAAGTTGTCCA	23	53.2		
rpsP	286 F	CAACCCGTTCTACAAAATTGTCGT	24	56.0	163	T
	286 R	AACCGTTTGAGTTACTTGTGCG	22	56.0		
MF01456	286_MFL_F	CGACTTCAAAGAAAGGGTTCGA	22	55.2	191	T
	286_MFL_R	CTAACAGTGTTTGTCGATTGAGCG	24	56.3		
MYF_01300	361 _MFL_F	CCAAGCGGGGTAGGAAAAACATC	23	58.4	159	A
	361 _MFL_R	AGTTGGAATCTGGCTTAAAAACCC	24	55.8		
MHP7448_0361	361 F	CCAAGTGGTCAGGGCAAATCATC	23	58.4	159	A
	361 R	AGTAGCCGTCTGACTTAAAAATCC	24	54.8		
MYF_01305	362 _MFL_F	TGTTTCGATAATTTCGCTTGCTTC	24	54.4	180	A
	362 _MFL_R	ATTGCCGCTCATAAAATTGGGC	22	56.8		
MHP7448_0362	362 F	TACGATTAAATACACGCTTGCTGG	24	55.4	168	A
	362 R	CGCATCGGGTCGATTACAAAA	21	55.7		
MHP7448_0397	397 F	CCGCTGACTCTGACATCCTTTTGC	24	60.2	179	T
	397 R	GCACCAGAGGAACCAAACGGA	21	59.9		
MF00009	397_MFL_F	CAAAATAGCACTTCCAAAACGATT	24	52.6	174	T
	397_MFL_R	GGTATAATCCCATATTCCCAAGTT	24	52.3		
MHP7448_0484	484 F	GTCTCTGGCTCTGAATTGGTCGG	23	59.5	188	T
	484 R	CGGGTTACCATAACTTACAGTCGC	24	57.4		
MYF_00935	484_MFL F	TTCTGTGAATCTGAATTGGGCGA	23	56.7	160	T
	484_MFL R	CCATCATCCTCTGAACCTGAAAAA	24	55.1		
MHP7448_0485	485 F	ACGGTTCAATTTCCCTTAAAGAGC	24	55.7	158	T
	485 R	CTTTTTCCCCTTGAAGATGTGGG	23	56.2		
MF00865	485_MFL F	GGCAATCCGTTAAGCTCACG	20	56.5	165	T
	485_MFL R	GGCAATTTCAAAGACGAATTTTTG	24	52.3		

MYF_00885	497 _MFL_F	CGGTAGTGATGCCCAACAACAAA	23	58.0	163	A
	497 _MFL_R	CCCCTCAGTATTGTGCTGTTCAG	23	57.6		
MHP7448_0497	497 F	GATCCAGTGGTGGAAGTTCCTCAG	24	59.0	175	A
	497 R	GCTGGCTTTTCTGCTTCTTTAGGC	24	58.8		
MHP7448_0505	505 F	CCGAATTTTTTGCTACTTCGCTTG	24	55.4	166	T
	505 R	CGCTCACTTGGAACCGAAAGTC	22	58.4		
MF00046	505_MFL F	GCGAACAGACAATTATGGGTGTT	23	55.9	157	T
	505_MFL R	GAGCCAGGCACAAGAACTGAATAG	24	57.8		
MHP7448_0525	525 F	CCAAAAATGATTCCAGGGAA	20	56.3	190	GR
	525 R	GGCGTTTTAATTACACCGG	19	58.0		
MYF_00795	525_MFL_F	GCCGCTTTTTCAAAATCTTGGTTT	21	55.5	160	GR
	525_MFL_R	GATATCAATTCCACCAAAACGGTT	24	53.9		
MHP7448_0623	623 F	TCGAAGACACTTTGGCAAACA	21	55.1	165	T
	623 R	CCAATGCAACTTTTTGTTTTTCGC	24	54.9		
MYF_02820	623_MFL F	GTGAAAAACAGAAAATTGCACTTG	24	52.2	171	T
	623_MFL R	CTTCTGGAATATGATGGCTTACAA	24	52.9		
MHP7448_0081	81 F	CGAGACAAAAAGGTTCAAAAAG	22	57.1	171	GR
	81 R	GTAAAAAGGGATGTGCCTGC	20	60.4		
MF01193	81_MFL_F	GGGCTGATTTTAACGGGTGCTT	22	58.2	162	GR
	81_MFL_R	CGGAGCAACGCTTTGAAAAGTCTC	24	58.8		
MHP7448_0087	87 F	GTTGTCGGAGTGCCTAACACTGG	23	59.9	154	T
	87 R	GCAATAGGATACCTGGGGTGTC	22	57.2		
MF01266	87_MFL F	CGGGAATTTTGCCACCTAAA	20	53.5	183	T
	87_MFL R	CAGCCTCATTTTCAGCAGGTTT	22	56.1		
deoC	deoC F	GCTTTGATTGCCAATTTTTCG	21	56.7	163	GR
	deoC R	CAGCCCTTTTATCACCTCATG	21	60.6		
deoC	deoC_MFL_F	GCCAGTTTGGGTTAGGCTTG	20	56.7	160	GR
	deoC_MFL_R	CATTACCATGTCAATTTCGTCAGC	24	54.7		
glyA	glyA_MFL F	GTGGCTATTCTGCTTATTCAGGGA	24	56.7	168	T
	glyA_MFL R	GGGTTGTGGCGGTTATCACATG	22	58.7		
glyA	glyAF	TTGATTTTGCCCGCTTTCGAC	21	56.7	170	T
	glyAR	GCCACCTCTTGGTCCTCTCAAAG	23	59.6		
lon	lon F	TAAAACTTCAAGCATCGCCG	20	58.4	165	GR
	lon R	GAAATCCGAGGACATCGCC	19	62.3		
lon	lon_MFL_F	CGGGCGTGGGGATATTAAACTAAC	24	57.6	176	GR
	lon_MFL_R	AAGGTCCATCTTTTGGAATTGCC	23	56.0		
sipS	sipS F	CGCTCACTTGGAACCGAAAGTC	22	55.6	179	T
	sipS R	CCTTCGGGAATAACTCCGTTAA	22	54.4		
MYF_01338	sipS_MFL F	GTGGTTTTCAAATATAAGGAGCAA	24	51.8	167	T
	sipS_MFL R	AATTTACCTTCAGGAATTGTGCC	23	53.8		
upp	upp F	GGCTCTAAAACAATCGCAACTG	22	60.8	153	GR
	upp R	CGGTTTGGGTATGCTTGACTC	21	62.6		
upp	upp_MFL_F	CAACACAAGTTGTTTTTACCCCG	23	55.1	173	GR
	upp_MFL_R	GGCAACAGAATCGGCATCAGG	21	58.8		

Primers were used in the follow analysis: Adhesins coding genes (A), Tandem repeats (T) and establishment of gene reference (GR) used in $2^{-\Delta Ct}$ calculus. Primers used for *M. hyopneumoniae* strain 7448 were the same used for strain J, except for gene MHP7448_0108.

Supplementary Table S2: DNA repeats after filtered parameters.

REP ID*	ORF Name	Product	Sequence**	Start	End	Strand	ATG Distance***	GOZ****	POZ****
PAL_E_144	MHP7448_0301	hypothetical protein	TCTTGAGGATCCTAAAGA	364101	364118	+	118	TU 47 (I)	R
PAL_E_145	MHP7448_0302	permease	AATTAGTTTATATACTAACT	364529	364548	+	68	TU 47 (I)	2.R
PAL_E_148	MHP7448_0320	hypothetical protein	TATAAAAAAAATTTTCTTATA	395557	395576	+	58	TU 51 (I)	R
PAL_E_158	glpK	c glycerol kinase	TAAAAAAAAATATTTTTTTATTA	450674	450695	-	139	TU 62 (F)	1.R
PAL_E_16	MHP7448_0110	c ISMHp1 transposase	AATTAAAAATAATTAATGTTAATT	152882	152905	-	39	Class 1	1.R
PAL_E_163	MHP7448_0375	c PTS system enzyme IIB component	TTATAAAAAAAATATTTTTAAAA	474355	474376	-	143	TU 64 (I)	R
PAL_E_184	MHP7448_0489	c hypothetical protein	AAAATTATTAACAATTTT	640174	640191	-	259	TU 84 (F)	2.R
PAL_E_201	MHP7448_0557	c d-ribulose-5-phosphate 3 epimerase	TGTTTTTTATAGAAAACA	742475	742492	-	114	TU 97 (I)	R
PAL_E_208	pdhD-1	c dihydrolipoamide dehydrogenase	TTTTTTTATATAATAAAA	760176	760193	-	73	mC 36	1.R.1
PAL_E_209	nagA	N-acetylglucosamine-6-phosphate deacetylase	TTTTATTATATAAAAAAA	760176	760193	+	84	mC 37	1.R
PAL_E_210	dnaE	c DNA polymerase III alpha subunit	TAAAAAACGCTTTTTTTA	773067	773084	-	21	TU 104 (F)	R
PAL_E_22	MHP7448_0162	hypothetical protein	AAATTATGTTAAAATAGTTT	200692	200711	+	227	TU 26 (I)	R
PAL_E_234	MHP7448_0722	c conserved hypothetical protein	ATTAAAAAAATATTTTTTTATT	777713	777732	-	146	TU 104 (F)	2.R.2
PAL_E_3	MHP7448-tRNA-Leu1	tRNA-Leu	TTAAATTATTAAAAATTTTA	154620	154639	+	119	Class 3	R
PAL_E_35	MHP7448_0323	transposase	TTAATTTTTATTAATTACAATTAA	396982	397005	+	39	Class 1	1.R
PAL_E_37	MHP7448_0325	c hypothetical protein	AAATTAATATACTAACTT	400998	401015	-	51	TU 52 (F)	1.R
PAL_E_38	MHP7448_0328	hypothetical protein	TAAATTGGTATACTAACTTA	404239	404258	+	53	TU 54 (F)	1.R.2
PAL_E_5	MHP7448-tRNA-Cys1	tRNA-Cys	TTACATTTTTATAAAAATCAAA	433283	433304	-	2	Class 3	R.1
PAL_E_57	MHP7448_0495	c ISMHp1 transposase	AATTAAAAATAATTAATGTTAATT	647107	647130	-	39	Class 1	1.R
PAL_E_71	MHP7448_0636	c ISMHp1 transposase	AATTAAAAATAATTAATGTTAATT	856711	856734	-	39	Class 1	R
PAL_E_76	MHP7448_0673	c ISMHp1 transposase	AATTAAAAATAATTAATGTTAATT	913858	913881	-	39	Class 1	1.R
PAL_E_77	MHP7448_0681	hypothetical protein	AAATTTTTCCATATAAAAATTT	87214	87235	+	107	TU 07 (F)	1.R
PAL_E_79	MHP7448_0687	hypothetical protein	TTTATGAATATTTAAAAA	260145	260162	+	107	mC 15	R
PAL_E_80	MHP7448_0687	hypothetical protein	TTTATGAATATTTAAAAA	260242	260259	+	10	mC 15	R
PAL_E_88	MHP7448_0030	hypothetical protein	CAAAAAAATTAATTTATTTTTTTG	36497	36520	+	146	TU 03 (L)	1.R

PAL_E_99	MHP7448_0095	c conserved hypothetical protein	TATAAAAAAATTTTAAAATTTTTTTATA	124949	124976	+	55	mC 09	R.1
PAL_ES_241	dnaA	chromosomal replication initiator protein	CTTAAATTTAAGTTTTAG	18	35	+	171	TU 01 (F)	1.R
PAL_ES_260	MHP7448_0081	amino acid permease	AATTATAACATATTCTAATT	101397	101416	+	143	TU 09 (F)	3.R
PAL_ES_304	MHP7448_0248	c hypothetical protein	AAGCCCTAATTTGGTCTT	303139	303156	-	91	TU 37 (I)	R
PAL_ES_312	MHP7448_0290	hypothetical protein	AGATAAAAGCATTTATTT	348982	348999	+	66	TU 45 (I)	R
PAL_ES_317	MHP7448_0308	conserved hypothetical protein	TTCATTGTTATACCAATTAA	372807	372826	+	210	TU 49 (L)	3.R
PAL_ES_334	MHP7448_0342	c hypothetical protein	AATATTAAATATTTTGTATT	426828	426847	-	229	TU 57 (I)	1.R
PAL_ES_342	MHP7448_0356	c hypothetical protein	TTTAATTATTAATAAAAAAA	446227	446246	-	67	TU 61(F)	2.R
PAL_ES_347	MHP7448_0375	c PTS system enzyme IIB component	TAATCCTTGCAAGATTTA	474337	474354	-	125	TU 64 (I)	R
PAL_ES_354	MHP7448_0391	conserved hypothetical protein	AATTTTTATATAAAGAAT	492897	492914	+	121	TU 67 (F)	1.R
PAL_ES_367	MHP7448_0429	hypothetical protein	ATAATATTATTTTATTAT	543372	543389	+	11	TU 72(I)	R.1
PAL_ES_383	MHP7448_0486	c hypothetical protein	AATTTAATTAATAAATTT	634922	634939	-	79	TU 81 (I)	1.R
PAL_ES_418	pulA	pullulanase	GTAAATTTTAACATTTTC	777840	777857	+	15	mC 38	R.2
PAL_ES_420	xylF	ABC transporter xylose-binding lipoprotein	TTTTGTAATATAATAGAAAA	802106	802125	+	75	TU 106 (I)	R
PAL_ES_425	dam	c DNA adenine methylase	AAATGTAAATAATTTCATTT	833633	833652	-	125	mC 39	1.R.1
PAL_ES_427	MHP7448_0623	ABC transporter ATP-binding - Pr1	TAATATAAAAATTATATAATTTTTATATTA	834064	834093	+	68	TU 109 (F)	R.2
PAL_ES_442	MHP7448_0681	hypothetical protein	TTATTATAATTTTAACAA	87269	87286	+	56	TU 07 (F)	1.R
PAL_ES_448	MHP7448_0688	c ISMHp1 transposase	AAAAGCTGTAAAACTTTT	267005	267022	-	186	Class 1	1.R
PAL_ES_449	MHP7448_0693	hypothetical protein	ATTATTTATAAAAATTAT	400956	400973	+	73	mC 20	R
PAL_ES_450	MHP7448_0693	hypothetical protein	AAGTTAGTATATTAATTT	400998	401015	+	31	mC 20	R
PAL_ES_458	MHP7448_0722	c conserved hypothetical protein	GAAAATGTTAAAATTTAC	777840	777857	-	273	TU 104 (F)	2.R.1
PAL_S_464	MHP7448-tRNA-Tyr1	tRNA-Tyr	TTTTAAGTTAAAAGTTTAATTTAAAA	487852	487877	-	112	Class 3	1.R.1
PAL_S_466	MHP7448-tRNA-Pro1	tRNA-Pro	ATTTTTTTAAAAGAAATATTTTAAAAAAAT	111440	111469	+	44	Class 3	1.R
PAL_S_474	MHP7448_0019	c ABC transporter ATP-binding protein	TAATTTAATTAAAATTATTTAAATTA	21202	21227	-	65	TU 02 (I)	R
PAL_S_477	MHP7448_0054	conserved hypothetical protein	TTTTTTTATGGAAAATTTCATAAAAAAA	65647	65674	+	73	TU 06 (I)	R
PAL_S_484	MHP7448_0080	c conserved hypothetical protein	TAGTTAATTAATTCACTG	101434	101451	-	138	TU 08 (F)	1.R
PAL_S_495	clpB	c ATP-dependent protease binding protein	TTTTTTCATTTAAGAAAAAA	136899	136918	-	84	mC 10	1.R

PAL_S_514	MHP7448_0155	c conserved hypothetical protein	AAAAATACTAGGATTTTG	194974	194991	-	268	TU 25 (I)	R
PAL_S_516	rpsM	c 30S ribosomal protein S13	AAATATTTAACTAAATATTC	207398	207417	-	3	TU 27 (I)	R
PAL_S_519	rpsJ	c 30S ribosomal protein S10	AATTTTTTTTAAAAATTT	220796	220813	-	195	TU 27 (F)	1.R.1
PAL_S_530	smf	c DNA processing protein SMF	CTTAATTTATAAATGAAA	304144	304161	-	9	TU 37 (F)	R
PAL_S_534	eno	enolase	TTTCATTTATAAATTAAG	304144	304161	+	266	TU 38 (F)	1.R
PAL_S_536	eno	enolase	TTTTTTTGAAATCAAAGAAA	304220	304239	+	188	TU 38 (F)	1.R
PAL_S_540	MHP7448_0272	c P97-like protein	AAATTCAAGGTTGATTTT	331980	331997	-	240	TU 42 (F)	2.R
PAL_S_557	MHP7448_0318	hypothetical protein	TAATGTAAAATTAAATTA	392971	392988	+	220	TU 51(I)	1.R
PAL_S_566	MHP7448_0333	c lipoprotein	CTACTAATTTATTAGTAA	412933	412950	-	221	TU 55 (F)	2.R.1
PAL_S_568	MHP7448_0334	hypothetical protein	TTACTAATAAATTAGTAG	412933	412950	+	174	TU 56 (F)	R.2
PAL_S_569	MHP7448_0335	conserved hypothetical protein	ATATTTTCAAAAAAATAT	414201	414218	-	39	TU 56 (L)	R.1
PAL_S_576	glpK	c glycerol kinase	AAATTAACAAAATTGCTAATTT	450600	450621	-	65	TU 62 (F)	1.R
PAL_S_577	MHP7448_0363	hypothetical protein	TTAGTTTCTCGAAATTAA	454601	454618	+	82	TU 63 (I)	R
PAL_S_614	MHP7448_0489	c hypothetical protein	TAAAATTATAAAATTTTT	640074	640091	-	159	TU 84 (F)	2.R
PAL_S_617	pgk	phosphoglycerate kinase	AAAAATTTTATAATTTTA	640074	640091	+	173	TU 85 (F)	1.R.1
PAL_S_648	MHP7448_0557	c d-ribulose-5-phosphate 3 epimerase	CATCTTCAGCTGTAGATT	742565	742582	-	204	TU 97 (I)	R
PAL_S_654	pdhD-1	c dihydrolipoamide dehydrogenase	TAGCTATAACTTTGTAGCTA	760242	760261	-	139	mC 36	2.R
PAL_S_655	nagA	N-acetylglucosamine-6-phosphate deacetylase	TAGCTACAAAGTTATAGCTA	760242	760261	+	16	mC 37	R.1
PAL_S_656	rpiB	ribose 5-phosphate isomerase B	TAATTCTAATTAGTATTT	765882	765899	+	86	TU 103 (I)	1.R
PAL_S_658	glpD	c glycerol-3-phosphate dehydrogenase	TTATAATTTATATTATAG	781165	781182	-	27	TU 105 (L)	R.1
PAL_S_708	MHP7448_0706	hypothetical protein	ATCAAATTCGAATTTAAA	570526	570543	+	283	TU 75 (I)	R
PALG_E_10	MHP7448-tRNA-Trp1	tRNA-Trp	AATATTTTTTT-N10-AAAAATATAAT	124948	124979	+	70	Class 3	2.R.1
PALG_E_1000	MHP7448_0399	c hypothetical protein	ATTTTTATTT-N15-AAATAACTAT	501533	501567	-	267	TU 68 (F)	1.R
PALG_E_1001	asnS	asparaginyl-tRNA synthetase	ATTTAAGATT-N10-AATTTTACAT	503555	503584	+	179	TU 69 (I)	1.R
PALG_E_1002	asnS	asparaginyl-tRNA synthetase	ACTTTTTGA-N7-TAAAAAAGT	503594	503618	+	145	TU 69 (I)	1.R
PALG_E_1003	ruvA	holliday junction DNA helicase RuvA	TTTTAGTTG-N8-CAACCAAAA	509524	509549	+	11	TU 69 (I)	R
PALG_E_1004	MHP7448_0408	c hypothetical protein	TAATTTTATT-N10-AATATAATTA	514329	514358	-	85	TU 70 (I)	1.R

PALG_E_1005	MHP7448_0410	pseudogene	ATAATTTTT-N14-AAAAATTAT	517051	517082	-	44	Class 2	R
PALG_E_1007	MHP7448_0410	pseudogene	AAAAAAAAGTT-N9-AAGATTTTTTT	517210	517240	-	203	Class 2	R
PALG_E_1008	MHP7448_0412	c conserved hypothetical protein	GTTTTCATC-N5-GTTGAAAAC	519084	519106	-	27	TU 70 (I)	R
PALG_E_1010	MHP7448_0414	c putative ICEF-II	AACAAAAGTAT-N8-ATACTTTTTTT	523195	523224	-	12	TU 70 (I)	R
PALG_E_1011	MHP7448_0414	c putative ICEF-II	ATTCAATAGAAA-N14-TTCCTATTGAAT	523356	523393	-	173	TU 70 (I)	R
PALG_E_1012	MHP7448_0414	c putative ICEF-II	AGAAATTGAA-N5-TGCAAATTCT	523459	523483	-	276	TU 70 (I)	R
PALG_E_1013	MHP7448_0422	c hypothetical protein	AAAATGAAAA-N9-TTTTCAATTT	534803	534831	-	221	TU 71 (I)	R
PALG_E_1014	MHP7448_0431	c conserved hypothetical protein	TAGAAAAAA-N15-TTTTATCTA	546882	546914	-	200	TU 73 (L)	R
PALG_E_1015	MHP7448_0432	c hypothetical protein	TAGAAAAAAT-N13-ATTTTTTGCA	548506	548538	-	39	TU 73 (I)	R
PALG_E_1017	sgaH	c 3-hexulose-6-phosphate synthase	TTGAATTTTT-N6-AAAAATTCAA	554408	554433	-	57	TU 73 (I)	1.R
PALG_E_1018	sgaH	c 3-hexulose-6-phosphate synthase	CCCCGTTTTTT-N11-AAAAAACATGG	554616	554648	-	265	TU 73 (I)	1.R
PALG_E_1019	MHP7448_0439	c conserved hypothetical protein	TTATAATTTA-N8-TAAATTTTAA	556882	556909	-	8	TU 73 (I)	R
PALG_E_1020	MHP7448_0440	c hypothetical protein	TTTTTAATTTC-N6-GAAGTTAAAAA	557896	557923	-	70	TU 73 (F)	R.1
PALG_E_1021	MHP7448_0441	hypothetical protein	TTTTTAACTTC-N6-GAAATTAAAAA	557896	557923	+	77	mC 27	R.1
PALG_E_1022	MHP7448_0442	pseudogene	TACAATTAATT-N9-AAGTAATTTTA	558900	558930	+	17	Class 2	1.R
PALG_E_1023	MHP7448_0443	c hypothetical protein	TATCTATGGT-N8-AGCATACATA	562673	562700	-	29	TU 74 (L)	R.1
PALG_E_1024	MHP7448_0443	c hypothetical protein	CAAAAAACTAG-N6-CTAGTTTTTTG	562755	562782	-	111	TU 74 (L)	2.R
PALG_E_1025	MHP7448_0447	hypothetical protein	CAAAAGTTGC-N7-GCAATTTTTG	585304	585330	+	227	TU 76 (I)	R
PALG_E_1026	MHP7448_0448	hypothetical protein	GAATTTATTC-N9-GACTAAATTC	586194	586222	+	127	TU 76 (L)	R
PALG_E_1027	MHP7448_0452	c ABC transporter ATP-binding protein	CAATTTTTTCT-N10-AGAAAAAAATTG	590834	590865	-	44	TU 77 (I)	1.R
PALG_E_1028	MHP7448_0452	c ABC transporter ATP-binding protein	TTTTTTCAA-N8-TAAAAAAAA	590867	590892	-	77	TU 77 (I)	1.R
PALG_E_1029	MHP7448_0453	c ABC transporter permease protein	AGCAAAATTT-N13-AAATTTTCCT	598894	598926	-	37	TU 77 (F)	R
PALG_E_1030	MHP7448_0453	c ABC transporter permease protein	TAATTTTGT-N14-ACAAAATTA	598979	599010	-	122	TU 77 (F)	1.R
PALG_E_1032	acpD	acyl carrier protein phosphodiesterase	AGGAAAATTT-N13-AAATTTTGCT	598894	598926	+	121	TU 78 (F)	1.R
PALG_E_1033	rplK	c 50S ribosomal protein L11	AATAAAAAA-N15-TTTTTAATT	603194	603226	-	107	TU 79 (F)	1.R.1
PALG_E_1034	MHP7448_0463	hypothetical protein	ATAATTTAAT-N14-ATAAAATAAT	605321	605354	+	279	TU 80 (L)	R
PALG_E_1035	MHP7448_0463	hypothetical protein	TTTAAGTTA-N8-TAACTTAAA	605571	605596	+	37	TU 80 (L)	R

PALG_E_1036	MHP7448_0470	c conserved hypothetical protein	TTATCTAATATT-N7-AATATTAGATAA	616574	616604	-	10	TU 81 (F)	1.R
PALG_E_1037	MHP7448_0471	conserved hypothetical protein	TTATCTAATATT-N7-AATATTAGATAA	616574	616604	+	140	TU 82 (F)	1.R
PALG_E_1038	MHP7448_0471	conserved hypothetical protein	CACTTTTAAA-N15-TCTAAAACTG	616652	616686	+	58	TU 82 (F)	R.2
PALG_E_1039	ptsI	phosphoenolpyruvate-protein phosphotransferase	AAAAATTTTAT-N9-ATATAATTTTT	617663	617693	+	40	TU 82 (I)	R
PALG_E_1040	nadE	NH 3 -dependent NAD synthetase	AAACCGAAAAATT-N15-AATTTTTCGGTTT	619524	619564	+	79	TU 82 (I)	2.R
PALG_E_1041	MHP7448_0477	c conserved hypothetical protein	ATTTTTGTTT-N5-AAACTAAAAT	624055	624079	-	266	TU 83 (I)	1.R
PALG_E_1042	MHP7448_0484	c hypothetical protein	CATTTTTTT-N11-AAAAAAAAG	633022	633050	-	44	TU 83 (F)	R.1
PALG_E_1043	MHP7448_0484	c hypothetical protein	AAAAAAAATA-N11-TGATTTTTTT	633072	633102	-	94	TU 83 (F)	R.1
PALG_E_1044	MHP7448_0484	c hypothetical protein	AAAAAACTG-N8-CTGTTTTTT	633234	633259	-	256	TU 83 (F)	1.R
PALG_E_1045	MHP7448_0485	conserved hypothetical protein	CTTTTTTTT-N11-AAAAAAATG	633022	633050	+	248	mC 29	2.R
PALG_E_1046	MHP7448_0485	conserved hypothetical protein	AAAAAAATCA-N11-TATTTTTTTT	633072	633102	+	196	mC 29	2.R
PALG_E_1047	MHP7448_0485	conserved hypothetical protein	AAAAAACAG-N8-CAGTTTTTT	633234	633259	+	39	mC 29	R.2
PALG_E_1049	MHP7448_0489	c hypothetical protein	TTTAAAATTT-N5-AAATCTTGAA	640128	640152	-	213	TU 84 (F)	2.R
PALG_E_1050	pgk	phosphoglycerate kinase	ATAAAACAG-N13-CTTATTTAT	640040	640070	+	194	TU 85 (F)	2.R.1
PALG_E_1051	pgk	phosphoglycerate kinase	AATAGTTTAT-N12-ATAAAATTTT	640221	640252	+	12	TU 85 (F)	R.2
PALG_E_1052	MHP7448_0491	hypothetical protein	TGAATTAGTT-N11-AACTAACTTA	641619	641649	+	5	TU 85 (I)	R.1
PALG_E_1053	fruA	PTS system, fructose-specific IIABC component	ATTTTTTTAA-N14-TTAAAAAAAT	641942	641975	+	45	TU 85 (I)	R
PALG_E_1054	MHP7448_0495	c ISMHp1 transposase	TTACAATTTTA-N14-TATAATTATAA	647071	647106	-	3	Class 1	R
PALG_E_1055	MHP7448_0497	c p76 membrane protein precursor	TAAATTTTAAA-N6-TTTATAATTTA	657663	657690	-	78	TU 86 (I)	1.R
PALG_E_1056	MHP7448_0497	c p76 membrane protein precursor	TAAATTTTTTAG-N8-CTAAAAAATTTA	657737	657768	-	152	TU 86 (I)	1.R
PALG_E_1057	MHP7448_0500	c conserved hypothetical protein	TTTTAAAAA-N11-TTTTAAAAA	661834	661862	-	10	TU 86 (I)	R
PALG_E_1060	MHP7448_0505	c lipoprotein	TTTAAAAAAACCAA-N13-TTGGTTTTTTTAAA	670642	670682	-	228	TU 86 (F)	2.R
PALG_E_1061	MHP7448_0505	c lipoprotein	TTTTGTTTTA-N7-TAAAACAAAA	670683	670709	-	269	TU 86 (F)	2.R
PALG_E_1062	pdhC	dihydrolipoamide acetyltransferase	CAAAAAAATT-N14-ATTTTTTTTG	671196	671229	+	222	TU 87 (F)	2.R
PALG_E_1063	pdhC	dihydrolipoamide acetyltransferase	AAATTCTAA-N5-TTAGATTTT	671239	671261	+	190	TU 87 (F)	2.R.1
PALG_E_1064	MHP7448_0512	c hypothetical protein	AAAAAAATA-N11-TAATTTTTT	679900	679928	-	16	TU 88 (F)	R.1
PALG_E_1065	MHP7448_0512	c hypothetical protein	AAATTGTTA-N10-TAAAATTTT	680032	680059	-	148	TU 88 (F)	R.1

PALG_E_1066	MHP7448_0512	c hypothetical protein	TATTTGTATTTTT-N14-AAAAATTCAAATA	680130	680169	-	246	TU 88 (F)	R.2
PALG_E_1067	pepF	c oligoendopeptidase F	GTTTTATTTT-N8-AAAATTATAC	691456	691483	-	46	TU 90 (F)	R.1
PALG_E_1068	MHP7448_0522	conserved hypothetical protein	GTATAATTTT-N8-AAAATAAAAC	691456	691483	+	154	mC 30	1.R
PALG_E_1070	tufA	c elongation factor EF-Tu	TATTTTTTTTT-N13-AAAAAAGGATA	698979	699013	-	110	TU 91 (L)	4.R
PALG_E_1071	upp	c uracil phosphoribosyltransferase	TAAAAATTGTTTT-N5-AAAACAATTTTTA	703277	703307	-	14	TU 91 (I)	R.1
PALG_E_1072	deoC	c deoxyribose-phosphate aldolase	TTTAATTTTT-N15-AAAAATTCCA	704162	704196	-	141	TU 91 (F)	1.R.2
PALG_E_1073	deoC	c deoxyribose-phosphate aldolase	TTTTAATTTATA-N13-TTTAAATTAAAA	704261	704297	-	240	TU 91 (F)	2.R.1
PALG_E_1074	gyrA	DNA gyrase subunit A	TGGAATTTTT-N15-AAAAATTAAA	704162	704196	+	264	mC 31	2.R
PALG_E_1075	gyrA	DNA gyrase subunit A	TTTTAATTTAAA-N13-TATAAATTAAAA	704261	704297	+	163	mC 31	1.R
PALG_E_1076	MHP7448_0530	c hypothetical protein	TTTTAAAAA-N11-TTCTTAAAA	708059	708087	-	22	TU 92 (I)	1.R
PALG_E_1077	pyrH	c uridylate kinase	TTTTTTTCA-N14-TAAAAGAAA	716184	716215	-	14	TU 92 (F)	R
PALG_E_1078	pyrH	c uridylate kinase	ATTTTTTTTG-N11-CAAAAAAAAT	716260	716290	-	90	TU 92 (F)	1.R
PALG_E_1079	MHP7448_0536	hypothetical protein	ATTTTTTTTG-N11-CAAAAAAAAT	716260	716290	+	38	TU 93 (F)	R
PALG_E_1080	potA	c spermidine/putrescine ABC transporter ATP-binding	ATTTTTTTAA-N14-TTAAAAATAT	724106	724139	-	271	TU 94 (F)	1.R
PALG_E_1082	MHP7448_0543	hypothetical protein	TTTCCTATTCA-N14-TGATCAGGAAA	724726	724761	+	72	TU 95 (I)	R.1
PALG_E_1083	ktrB	potassium uptake protein	AATAAATTT-N15-AAATTTTTT	727339	727371	+	212	mC 33	1.R
PALG_E_1084	ktrB	potassium uptake protein	TTTTTTAGCT-N7-ATATAAAAAA	727518	727544	+	39	mC 33	R.1
PALG_E_1086	MHP7448_0547	c conserved hypothetical protein	TACTTTTAT-N11-ATAAAATTA	733590	733618	-	157	TU 96 (L)	R
PALG_E_1087	MHP7448_0557	c d-ribulose-5-phosphate 3 epimerase	AATTTTTTTC-N9-GAAAAAAATT	742435	742463	-	74	TU 97 (I)	R
PALG_E_1088	MHP7448_0557	c d-ribulose-5-phosphate 3 epimerase	TTTTGTCAC-N13-GTGCCAAAA	742522	742552	-	161	TU 97 (I)	R
PALG_E_1089	MHP7448_0559	c conserved hypothetical protein	TTTGACACAC-N10-GTGTCTGAAA	745711	745740	-	239	TU 97 (I)	1.R
PALG_E_1090	MHP7448_0559	c conserved hypothetical protein	GACAATTTTAA-N5-TTAAATTTGTC	745767	745793	-	295	TU 97 (I)	1.R
PALG_E_1091	MHP7448_0562	pseudogene	AAAATTCTTA-N8-TTATAATTTT	749506	749533	-	13	Class 2	R
PALG_E_1092	MHP7448_0565	pseudogene	TTATAATTAT-N9-ATAATTTTAA	753423	753451	-	49	Class 2	1.R
PALG_E_1093	MHP7448_0565	pseudogene	GACAATTTTAA-N5-TTAAATTTGTC	753591	753617	-	217	Class 2	1.R
PALG_E_1094	MHP7448_0567	c hypothetical protein	TTTTAATTTT-N8-AAACTTAAAA	756461	756488	-	135	TU 100 (F)	R
PALG_E_1095	MHP7448_0568	pseudogene	TTATAATTAT-N9-ATAATTTTAA	757341	757369	-	31	Class 2	R

PALG_E_1096	MHP7448_0569	hypothetical protein	TTAAAATTAT-N9-ATAATTATAA	757341	757369	+	76	TU 101 (I)	R
PALG_E_1097	pulA	pullulanase	AAAAGGTCGTC-N5-GACAACCTTTT	777626	777652	+	220	mC 38	2.R
PALG_E_1098	pulA	pullulanase	AAAAAATATT-N15-AATATTTATT	777716	777750	+	122	mC 38	2.R
PALG_E_1099	pulA	pullulanase	TAATAATAT-N5-ATATAATTA	777762	777784	+	88	mC 38	R.1
PALG_E_11	MHP7448-tRNA-Leu1	tRNA-Leu	ATTTTCAGTTTA-N12-TAAACAGAAAAT	154509	154544	-	8	Class 3	R.2
PALG_E_1103	MHP7448_0598	hypothetical protein	TTATTAAAATT-N7-AATTTTTAGAA	794099	794127	+	54	TU 106 (I)	R
PALG_E_1104	MHP7448_0599	hypothetical protein	AAATTTTGT-N8-ATAAAATTT	795972	795997	+	64	TU 106 (I)	R
PALG_E_1106	xylF	ABC transporter xylose-binding lipoprotein	TAAAAAATATT-N10-AAAATTTTTTA	802072	802103	+	97	TU 106 (I)	R
PALG_E_1107	xylF	ABC transporter xylose-binding lipoprotein	AATTAATTT-N5-AAAATAATT	802137	802159	+	41	TU 106 (I)	R
PALG_E_1108	xylF	ABC transporter xylose-binding lipoprotein	TCGTAATTAA-N5-TTAATTAGGA	802168	802192	+	8	TU 106 (I)	R
PALG_E_1109	ddem	c cytosine specific DNA methyltransferase	TAATATATAAT-N9-ATTAAATACTA	816874	816904	-	220	TU 107 (I)	1.R
PALG_E_1110	rpoC	c DNA-directed RNA polymerase subunit beta	AAAATAAAA-N6-TCTTATTTT	821629	821652	-	18	TU 107 (I)	R
PALG_E_1111	rpoB	c DNA-directed RNA polymerase beta subunit	TACGTTTCTAA-N6-TTAAAAACTTA	825355	825382	-	25	TU 107 (I)	R
PALG_E_1112	rpoB	c DNA-directed RNA polymerase beta subunit	TATTTTTTTAT-N6-ATAAAAAAATA	825443	825470	-	113	TU 107 (I)	1.R
PALG_E_1114	rplJ	c 50S ribosomal protein L10	AAAATTTCT-N15-ATAAATTTT	826601	826633	-	136	TU 107 (F)	1.R.3
PALG_E_1115	rplJ	c 50S ribosomal protein L10	CATAAAAATTTT-N10-AAATTTTTTATG	826727	826760	-	262	TU 107 (F)	2.R.1
PALG_E_1117	MHP7448_0620	conserved hypothetical protein	AAAAATTTTT-N15-AAAAAATCTT	826684	826718	+	121	TU 108 (F)	1.R.2
PALG_E_1118	MHP7448_0620	conserved hypothetical protein	CATAAAAAATTT-N10-AAAATTTTTATG	826727	826760	+	79	TU 108 (F)	1.R.2
PALG_E_1119	dam	c DNA adenine methylase	TAAATAATTT-N15-ATATAATTTA	833512	833546	-	4	mC 39	1.R.2
PALG_E_1120	MHP7448_0623	ABC transporter ATP-binding - Pr1	TATATTTATT-N14-AATAAATATA	833946	833979	+	182	TU 109 (F)	1.R.2
PALG_E_1121	MHP7448_0625	pseudogene	TTAGATTTTAC-N15-GTAAATTTTAA	837472	837508	+	66	Class 2	R
PALG_E_1122	MHP7448_0626	c conserved hypothetical protein	TTAATAATT-N13-AATTATAAA	840521	840551	-	20	mC 40	R.1
PALG_E_1123	MHP7448_0626	c conserved hypothetical protein	ATATTAAAAAAATAAC-N6-GTTATTTTTTTAATAT	840761	840798	-	260	mC 40	2.R
PALG_E_1124	MHP7448_0627	ABC transporter ATP-binding - Pr1-like protein	TTTATAATT-N13-AATTATTAA	840521	840551	+	282	TU 110 (F)	1.R
PALG_E_1125	MHP7448_0627	ABC transporter ATP-binding - Pr1-like protein	ATATTAAAAAAATAAC-N6-GTTATTTTTTTAATAT	840761	840798	+	35	TU 110 (F)	R.2
PALG_E_1126	MHP7448_0630	c 5'-nucleotidase precursor	AAGGAAGAAAA-N11-TTTTCTTCTTT	848227	848259	-	52	TU 111 (F)	2.R
PALG_E_1127	MHP7448_0630	c 5'-nucleotidase precursor	AAAAAATACT-N9-AGTGTTTTTT	848351	848379	-	176	TU 111 (F)	2.R

PALG_E_1128	MHP7448_0630	c 5'-nucleotidase precursor	TTATAAATT-N5-AATTTTTAA	848397	848419	-	222	TU 111 (F)	2.R
PALG_E_1129	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase	AAAAATTTT-N8-ATAATTTTT	848399	848424	+	102	TU 112 (F)	2.R.1
PALG_E_1130	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase	TTTTTAATTT-N13-AAATTAAAAA	848453	848485	+	41	TU 112 (F)	1.R.2
PALG_E_1131	gcp	c O-sialoglycoprotein endopeptidase	TTTTATATTAT-N9-ATAATATAGAA	854422	854452	-	35	mC 41	R.2
PALG_E_1132	MHP7448_0636	c ISMHp1 transposase	TTACAATTTTA-N14-TATAATTATAA	856675	856710	-	3	Class 1	R
PALG_E_1133	MHP7448_0636	c ISMHp1 transposase	ATTTTGATAA-N13-TTTTCAAAAT	856829	856861	-	157	Class 1	1.R
PALG_E_1134	MHP7448_0640	c tRNA/rRNA methyltransferase	TTTCCATCTT-N8-AAGATTGAAA	859169	859196	-	2	TU 113 (I)	1.R
PALG_E_1135	leuS	c leucyl-tRNA synthetase	GCATTTTTTG-N11-CAAAAAATGC	868580	868610	-	104	TU 113 (F)	1.R.1
PALG_E_1136	leuS	c leucyl-tRNA synthetase	TTAAAATTTT-N5-AAAATTTTAA	868633	868657	-	157	TU 113 (F)	1.R.1
PALG_E_1137	uvrB	excinuclease ABC subunit B	GCATTTTTTG-N11-CAAAAAATGC	868580	868610	+	66	TU 114 (F)	1.R.1
PALG_E_1138	uvrB	excinuclease ABC subunit B	TTAAAATTTT-N5-AAAATTTTAA	868633	868657	+	19	TU 114 (F)	R.1
PALG_E_1139	MHP7448_0656	prolipoprotein p65	ATATAAAAATC-N5-GATTTTTATAT	876464	876490	+	82	TU 116 (I)	1.R
PALG_E_1141	rpmG	50S ribosomal protein L33	AACATTTTTC-N15-GAAAAATGTT	881603	881637	+	78	TU 116 (I)	1.R
PALG_E_1142	rpmG	50S ribosomal protein L33	TTTCTAAACT-N13-AATTTAGAAA	881669	881701	+	14	TU 116 (I)	R.1
PALG_E_1143	MHP7448_0660	hypothetical protein	AATTTTTTA-N10-TAAAAAAAT	882925	882952	+	8	TU 116 (I)	R
PALG_E_1144	MHP7448_0661	conserved hypothetical protein	TAATAATTTTACA-N7-TGTAAAATTATTA	884674	884706	+	3	TU 116 (L)	R
PALG_E_1145	MHP7448_0663	c adhesin like-protein P146	ATTATTGTA-N13-TATAATAAT	894385	894415	-	12	TU 117 (F)	R.1
PALG_E_1146	MHP7448_0663	c adhesin like-protein P146	TTTTTATTTT-N15-AAAATAAAAA	894515	894549	-	142	TU 117 (F)	1.R
PALG_E_1147	MHP7448_0663	c adhesin like-protein P146	AATTTTTTT-N10-AAAAAACTT	894553	894580	-	180	TU 117 (F)	3.R
PALG_E_1148	MHP7448_0663	c adhesin like-protein P146	ATTTTTTAA-N9-TTTAAAAAT	894656	894682	-	283	TU 117 (F)	3.R
PALG_E_1149	MHP7448_0664	ABC transporter ATP-binding - Pr1	TTATTAAAAAA-N13-TTTTGTTATAA	894775	894809	+	122	TU 118 (F)	R
PALG_E_1151	MHP7448_0670	conserved hypothetical protein	AATTTATAAT-N10-ACTAAAAATT	906638	906667	+	20	TU 120 (I)	R.1
PALG_E_1152	MHP7448_0673	c ISMHp1 transposase	TTACAATTTTA-N14-TATAATTATAA	913822	913857	-	3	Class 1	R
PALG_E_1153	MHP7448_0673	c ISMHp1 transposase	TTTACTTTTT-N5-AAAAAGCTAA	913977	914001	-	158	Class 1	1.R
PALG_E_1154	MHP7448_0675	c hypothetical protein	TAAAATTTTACT-N8-AGTAAAATTTTA	917708	917739	-	40	TU 121 (I)	R
PALG_E_1155	MHP7448_0709	hypothetical protein	ATAAATATTA-N15-TATAATTTAT	573044	573066	+	10	TU 75 (L)	R.1
PALG_E_1156	MHP7448_0712	hypothetical protein	TTGGATCAA-N6-TTGTTGCAA	584326	584353	+	96	TU 76 (I)	R

PALG_E_12	MHP7448-tRNA-Leu1	tRNA-Leu	AAAAATTAGTAAA-N11-TTTACTAATTTTT	154642	154678	-	141	Class 3	R
PALG_E_13	MHP7448-tRNA-His1	tRNA-His	TAAAAAACCCTT-N7-AAGTGTTTTTTA	160731	160761	-	67	Class 3	1.R
PALG_E_14	MHP7448-tRNA-His1	tRNA-His	ATCAAAAATT-N7-AATATTTAAT	160794	160820	-	130	Class 3	1.R
PALG_E_15	MHP7448-tRNA-His1	tRNA-His	TGATTTAGT-N9-ACTAAACAA	160862	160888	-	198	Class 3	1.R
PALG_E_16	MHP7448-tRNA-His1	tRNA-His	AAAAATGTGAAT-N14-ATTGAAATTTTT	160900	160937	-	236	Class 3	1.R
PALG_E_17	MHP7448-tRNA-Trp2	tRNA-Trp	ATTATTTTT-N13-ATAATTAAT	185847	185877	-	9	Class 3	R
PALG_E_18	MHP7448-tRNA-Trp2	tRNA-Trp	AAAAAATAAA-N6-TTTAATTTTT	185922	185947	-	84	Class 3	R
PALG_E_19	MHP7448-tRNA-Asn1	tRNA-Asn	AATAAAAATTT-N15-ATACTTTTATT	335421	335457	+	5	Class 3	R.1
PALG_E_20	MHP7448-tRNA-Cys1	tRNA-Cys	AAAAAATAA-N12-TTAGTTTTT	433319	433348	-	38	Class 3	1.R
PALG_E_21	MHP7448-tRNA-Ser2	tRNA-Ser	CTCAAAATTT-N6-AAAGTTTCAG	603655	603680	-	164	Class 3	1.R
PALG_E_22	MHP7448-tRNA-Ser3	tRNA-Ser	TATAATATA-N10-TATATAATA	854441	854468	+	1	Class 3	R.1
PALG_E_23	MHP7448-tRNA-Ile1	tRNA-Ile	ATAATTTTAATG-N6-CATAAAAACTAT	871683	871712	-	18	Class 3	R.1
PALG_E_237	dnaA	chromosomal replication initiator protein	TTATTTTAA-N6-TTAAAAAAA	39	62	+	144	TU 01 (F)	1.R
PALG_E_238	dnaA	chromosomal replication initiator protein	ATCAACTTTATT-N10-AATAAAGTTGAT	86	119	+	87	TU 01 (F)	1.R
PALG_E_239	dnaA	chromosomal replication initiator protein	TGTTTGCAA-N14-TTTTAAACA	124	155	+	51	TU 01 (F)	R.1
PALG_E_24	MHP7448-tRNA-Arg1	tRNA-Arg	TAAATTATAA-N13-TAATTATTTA	111559	111591	-	5	Class 3	1.R.2
PALG_E_240	dnaA	chromosomal replication initiator protein	AAATTATACT-N10-AGGAAAATTT	172	201	+	5	TU 01 (F)	R.1
PALG_E_241	ftsY	cell division protein ftsY	TTAAAATTATTTATT-N14-AAAAAGTAATTTTAA	7939	7982	+	41	TU 01 (I)	1.R
PALG_E_242	MHP7448_0017	hypothetical protein	TTCTAATCT-N13-ATAATAGAA	17187	17217	+	142	mC 02	R
PALG_E_243	MHP7448_0020	c ABC transporter ATP-binding protein	AATTACTAA-N12-TGAATAATT	23587	23616	-	9	TU 02 (I)	R
PALG_E_244	MHP7448_0023	c ABC transporter ATP-binding protein	ATTTTTATGA-N15-TACTAAAAAT	29307	29341	-	65	TU 02 (I)	R
PALG_E_245	sipS	signal peptidase I	AACACTTTTTATT-N11-AATTAAAATTGTT	32767	32803	+	39	TU 03 (F)	2.R.2
PALG_E_246	gatC	Asp-tRNAAsn/Glu-tRNAGln amidotransferase C subunit	GTTATTTAT-N10-ATAAAGAAC	33390	33417	+	14	TU 03 (I)	R.1
PALG_E_248	MHP7448_0030	hypothetical protein	TAAACCTTT-N6-AAAATTTTA	36531	36554	+	112	TU 03 (L)	1.R
PALG_E_249	parE	c topoisomerase IV subunit B	TTTAGAAAA-N11-TTTTCAGAA	44996	45024	-	16	TU 04 (F)	R.2
PALG_E_25	MHP7448-tRNA-Leu1	tRNA-Leu	AAAAAAATAAT-N7-ATAATTTTTTT	154588	154616	-	87	Class 3	R
PALG_E_251	gap	glyceraldehyde 3-phosphate dehydrogenase	AATTTTACA-N10-TGTAAAATT	45050	45077	+	112	mC 03	1.R

PALG_E_252	recA	c recombination protein	TTTTCCTAAA-N11-TTTATTAAAA	53428	53458	-	68	TU 05 (F)	R.2
PALG_E_253	licA	PTS system, lichenan-specific IIA component	TTTTAATAAA-N11-TTTAGGAAAA	53428	53458	+	53	TU 06 (F)	R
PALG_E_255	rpsB	30S ribosomal protein S2	TGTAAATTT-N5-AGAATTACA	66249	66271	+	177	TU 06 (I)	R
PALG_E_257	MHP7448_0064	conserved hypothetical protein	AAAACATTTAT-N7-ATACATATTTT	77908	77936	+	29	TU 06 (I)	R
PALG_E_258	dnaK	chaperone protein dnaK - heat shock protein 70	ATAAATTTT-N11-AAAAATTAT	87681	87709	+	129	TU 07 (I)	1.R
PALG_E_259	rpsG	30s ribosomal protein S7	TAAAAAAAT-N15-ATATTTTTA	99321	99353	-	15	TU 08 (I)	R
PALG_E_26	MHP7448-tRNA-His1	tRNA-His	AAATATGGTA-N13-TTCCATATTT	160665	160697	-	1	Class 3	R
PALG_E_260	rpsL	c 30S ribosomal protein S12	AATTTAAAA-N14-TTTTATATT	99867	99898	-	89	TU 08 (I)	R
PALG_E_261	MHP7448_0080	c conserved hypothetical protein	TTTAACAAA-N12-TATGTTATA	101400	101429	-	104	TU 08 (F)	R
PALG_E_262	MHP7448_0080	c conserved hypothetical protein	TAAAAAATA-N5-TAAATTTTA	101455	101477	-	159	TU 08 (F)	1.R
PALG_E_263	MHP7448_0080	c conserved hypothetical protein	AAGATTTTAA-N12-TTAAAATTAT	101509	101540	-	213	TU 08 (F)	1.R
PALG_E_265	MHP7448_0081	amino acid permease	TTAACAATT-N14-AGTTCTTAA	101309	101340	+	219	TU 09 (F)	3.R
PALG_E_266	MHP7448_0081	amino acid permease	TAAAATTTA-N5-TATTTTTTA	101455	101477	+	82	TU 09 (F)	2.R
PALG_E_267	nox	NADH oxidase	TTTTATTTT-N13-AAAAGTAAA	102968	102998	+	122	TU 09 (L)	2.R
PALG_E_268	nox	NADH oxidase	AATCTAAAAAAAT-N8-ATTTTTTTAGATT	103021	103054	+	66	TU 09 (L)	1.R
PALG_E_269	nox	NADH oxidase	AAAAATAAA-N6-TTTATTTTT	103055	103078	+	42	TU 09 (L)	1.R.1
PALG_E_27	MHP7448-tRNA-Trp2	tRNA-Trp	AATTTAAAT-N9-AGTTAAATT	185882	185908	-	44	Class 3	1.R.1
PALG_E_270	nox	NADH oxidase	AAATTTAAAC-N11-GTTATAATTT	103090	103120	+	0	TU 09 (L)	R.1
PALG_E_271	secA	c preprotein translocase SecA subunit	AAATTGAAA-N14-TTTAACTTT	110317	110348	-	78	TU 10 (F)	R.1
PALG_E_272	secA	c preprotein translocase SecA subunit	AAACTTTTC-N10-GAATAATTT	110372	110399	-	133	TU 10 (F)	1.R
PALG_E_273	MHP7448_0092	c conserved hypothetical protein	TTTATTTATAA-N12-TTATAAAAAAA	122064	122097	-	30	TU 13 (I)	R
PALG_E_274	MHP7448_0092	c conserved hypothetical protein	TTAAAAAAC-N11-GTTTTTTGA	122121	122149	-	87	TU 13 (I)	R
PALG_E_275	MHP7448_0092	c conserved hypothetical protein	TAATCTTAAA-N13-TTTATGATTA	122155	122187	-	121	TU 13 (I)	R
PALG_E_276	MHP7448_0093	c hypothetical protein	TTAAATTTT-N8-ATTATTTAA	122619	122644	-	0	TU 13 (F)	R
PALG_E_277	MHP7448_0094	conserved hypothetical protein	AATAATTAT-N15-AAATTTATT	122622	122654	+	63	mC 08	1.R.1
PALG_E_278	MHP7448_0094	conserved hypothetical protein	TTTTATTTTTA-N5-TAAAAATAAAA	122664	122690	+	27	mC 08	1.R.1
PALG_E_28	MHP7448-tRNA-Tyr1	tRNA-Tyr	AATTTGCTAA-N5-TTAGCTAATT	487937	487961	-	197	Class 3	1.R

PALG_E_280	tpx	thiol peroxidase	AAAAAGTAAAATTA-N5-TAATTTTAATTTTT	125169	125201	+	62	TU 14 (F)	R
PALG_E_281	tpx	thiol peroxidase	CTTTTTAAAT-N13-AATTAAAAAG	125205	125237	+	26	TU 14 (F)	R.3
PALG_E_282	lgt	prolipoprotein diacylglyceryl transferase	TTATTCTAAA-N5-TTCATAATAA	125989	126013	+	270	TU 14 (I)	1.R
PALG_E_283	lgt	prolipoprotein diacylglyceryl transferase	CAACTATTTT-N10-AAAAAAATTG	126173	126202	+	81	TU 14 (I)	1.R
PALG_E_284	lgt	prolipoprotein diacylglyceryl transferase	ATTAAAAAAT-N13-ATTATTCAAT	126226	126258	+	25	TU 14 (I)	R
PALG_E_285	tpiA	triosephosphate isomerase	TAATTTTTT-N14-AATAAATTA	136827	136858	+	138	TU 15 (F)	1.R
PALG_E_286	tpiA	triosephosphate isomerase	TTAAATGAAA-N6-TTTAATTTAA	136906	136931	+	65	TU 15 (F)	1.R
PALG_E_287	tpiA	triosephosphate isomerase	TAATTTAAA-N9-TTTAAATTA	136936	136962	+	34	TU 15 (F)	1.R
PALG_E_288	rpsO	30S ribosomal protein S15	ATTAATACT-N9-ATTTTTAAT	137794	137820	+	23	TU 15 (L)	R
PALG_E_29	MHP7448-tRNA-Tyr1	tRNA-Tyr	ATAATTTTT-N14-AAAAATTAT	487967	487998	-	227	Class 3	1.R
PALG_E_290	MHP7448_0106	c hypothetical protein	TAATTTTTC-N9-GATAAACTA	142493	142519	-	34	TU 16 (L)	R
PALG_E_291	MHP7448_0108	c protein P97 - copy 2	TTTTTTAGTTTTT-N5-AAAAAATAAGAAA	148809	148839	-	102	TU 16 (I)	R
PALG_E_293	gyrB	c DNA gyrase subunit B	AAAAAACCA-N10-TGGAATTTT	150922	150949	-	35	TU 16 (I)	1.R
PALG_E_294	gyrB	c DNA gyrase subunit B	TTTTCTTCGT-N8-AAGAAAAAAA	151039	151066	-	152	TU 16 (I)	1.R
PALG_E_295	MHP7448_0110	c ISMHp1 transposase	ATTCACTTT-N6-AATATGAAT	152949	152972	-	106	Class 1	1.R
PALG_E_296	MHP7448_0110	c ISMHp1 transposase	TTTTATTAAA-N10-TTTAAAAAAA	153095	153124	-	252	Class 1	1.R
PALG_E_298	MHP7448_0112	hypothetical protein	AAAATAAAA-N14-TTTAATTTT	154540	154571	+	147	TU 16 (F)	R
PALG_E_299	MHP7448_0112	hypothetical protein	AAAAATTAGTAAA-N11-TTTACTAATTTTT	154642	154678	+	40	TU 16 (F)	R
PALG_E_30	MHP7448-tRNA-Thr2	tRNA-Thr	CATTTTTTTA-N12-TAAAAAAATG	588596	588627	-	21	Class 3	1.R
PALG_E_301	pdhA	pyruvate dehydrogenase E1-alpha subunit	ATATAAATAATA-N12-TATAATTTATAT	158208	158243	+	91	TU 17 (I)	R
PALG_E_303	MHP7448_0117	conserved hypothetical protein	ATGTATGTT-N9-AAAATACTT	161002	161028	+	266	TU 18 (F)	1.R
PALG_E_304	MHP7448_0117	conserved hypothetical protein	AAATGTTAT-N14-ATTAAATTT	161241	161272	+	22	TU 18 (F)	R
PALG_E_305	MHP7448_0120	conserved hypothetical protein	TTTGACAAA-N15-TTTGTTAAA	163499	163531	+	130	TU 18 (I)	1.R
PALG_E_306	MHP7448_0120	conserved hypothetical protein	TATAATTATA-N10-TATAAATAGA	163619	163648	+	13	TU 18 (I)	R
PALG_E_309	infC	c translation initiation factor IF-3	GACAAACAAA-N8-TTTTTTTGTC	166956	166983	-	97	TU 19 (I)	R
PALG_E_31	MHP7448-tRNA-Thr2	tRNA-Thr	CTTTTTCTAAA-N11-TTTAAAAAAAG	588628	588660	-	53	Class 3	1.R
PALG_E_310	infC	c translation initiation factor IF-3	ATATTTTTTT-N14-AATATAATAT	167051	167084	-	192	TU 19 (I)	R

PALG_E_313	rplU	50S ribosomal protein L21	TTATTGCTTTTT-N6-AAAATGCTATAA	171615	171644	+	89	TU 21 (F)	R
PALG_E_314	lip2	c lipase-esterase	TAATATTAT-N13-AAAAAATTA	173399	173429	-	20	mC 12	R
PALG_E_318	scr	pseudogene	TAATTTTTT-N13-ATAATATTA	173399	173429	+	84	Class 2	1.R
PALG_E_319	scr	pseudogene	AAATTTTTTC-N12-GAAAAAACTT	173449	173480	+	33	Class 2	1.R.1
PALG_E_32	MHP7448-tRNA-Ser2	tRNA-Ser	TAATTTTTTATC-N6-GATATAATATTA	603506	603535	-	15	Class 3	R
PALG_E_323	MHP7448_0138	conserved hypothetical protein	TTTTTTCAAGG-N11-CCTGGAAAAAA	177630	177662	+	71	TU 23 (F)	2.R
PALG_E_325	MHP7448_0145	chromate transport protein	ATTAATTAT-N13-AAAAATAAT	185847	185877	+	123	TU 24 (F)	R
PALG_E_326	tig	trigger factor	AATGATAAAA-N9-TTTTATTATT	189257	189285	+	118	TU 24 (I)	R
PALG_E_327	MHP7448_0155	c conserved hypothetical protein	AAATTAAAT-N6-ATTTATTTT	194740	194763	-	34	TU 25 (I)	R
PALG_E_328	MHP7448_0155	c conserved hypothetical protein	TTAATTAAAAA-N11-TTTTTAATTAA	194780	194812	-	74	TU 25 (I)	R
PALG_E_329	MHP7448_0162	hypothetical protein	TTTTCAATTT-N9-AAATTCAAAA	200729	200757	+	181	TU 26 (I)	R
PALG_E_33	MHP7448-tRNA-Met2	tRNA-Met	TTTAATTTTT-N14-AAAAATTAAA	111108	111141	-	8	Class 3	R
PALG_E_330	rpsM	c 30S ribosomal protein S13	AAATCAAAATTTTTT-N14-AAAAAATTTTGATTT	207460	207503	-	65	TU 27(I)	R
PALG_E_331	rpmJ	c 50S ribosomal protein L36	TTTATATAA-N9-TTAAATAAA	207671	207697	-	18	TU 27(I)	R
PALG_E_332	rplB	c 50S ribosomal protein L2	ACTAAAAAAGAC-N9-GTCTAGTTTAGT	218319	218351	-	48	TU 27(I)	R
PALG_E_333	rpsJ	c 30S ribosomal protein S10	ATAATTTAA-N8-TTAAAGGAT	220657	220682	-	56	TU 27 (F)	R.1
PALG_E_334	rpsJ	c 30S ribosomal protein S10	GAAAATTTTT-N11-AATAATATTC	220764	220794	-	163	TU 27 (F)	1.R.1
PALG_E_336	MHP7448_0197	conserved hypothetical protein	AGTAAGCAAAA-N9-TTTTGCAAACT	221000	221030	+	98	TU 28 (F)	1.R
PALG_E_337	MHP7448_0197	conserved hypothetical protein	ATAATAATAT-N13-AATTTATTAT	221084	221116	+	12	TU 28 (F)	R.1
PALG_E_338	MHP7448_0198	protein P97 - copy 1	AATTTTAATA-N10-TATTTAAATT	221645	221674	+	294	TU 28 (I)	1.R
PALG_E_339	MHP7448_0198	protein P97 - copy 1	ACAAGTTAAA-N6-TTTCACTTAT	221833	221858	+	110	TU 28 (I)	1.R
PALG_E_340	MHP7448_0202	conserved hypothetical protein	TTTGCAATTT-N7-AAACTGAAAA	231729	231755	+	75	TU 29 (I)	R
PALG_E_342	MHP7448_0210	c ABC transporter ATP-binding protein	ATTTAATAA-N12-TAAATAAAT	241672	241701	-	87	TU 30 (F)	1.R
PALG_E_343	MHP7448_0211	hypothetical protein	GAAAAATAA-N15-TTCTTTTTC	241638	241670	+	88	TU 31 (F)	R.2
PALG_E_344	MHP7448_0211	hypothetical protein	ATTTATTTA-N12-TTATTAAAT	241672	241701	+	57	TU 31 (F)	R.2
PALG_E_345	MHP7448_0211	hypothetical protein	TTTAAAAAA-N11-TTAATTAAA	241723	241751	+	7	TU 31 (F)	R.2
PALG_E_346	oppB-1	oligopeptide ABC transporter system permease	ATTAATAAAA-N10-TTTGAATAAT	243114	243143	+	126	TU 31 (I)	R

PALG_E_347	oppB-1	oligopeptide ABC transporter system permease	ATCTAATAATT-N14-AAATATTAGAT	243172	243207	+	62	TU 31 (I)	R
PALG_E_349	MHP7448_0219	c hypothetical protein	AAGAATTTAT-N15-ATAATTTTTT	259800	259834	-	139	mC 14	R.1
PALG_E_350	MHP7448_0219	c hypothetical protein	AATTATTAAA-N15-TTAAATAACT	259896	259930	-	235	mC 14	R.1
PALG_E_352	nrdF	c ribonucleoside-diphosphate reductase beta chain	TACAAAAAAAT-N7-ATTTTTTGGTA	264502	264530	-	148	TU 32 (I)	R.1
PALG_E_353	nrdF	c ribonucleoside-diphosphate reductase beta chain	AAAAAGTCTAT-N6-ATAGGGTTTTT	264572	264599	-	218	TU 32 (I)	1.R.1
PALG_E_354	glyA	c glycine hydroxymethyltransferase	TTTTAAACA-N11-TGTTTTAAA	268510	268538	-	68	TU 32 (F)	2.R
PALG_E_355	glyA	c glycine hydroxymethyltransferase	TTTGCTTTCAA-N5-TTTAAACCAAA	268708	268734	-	266	TU 32 (F)	2.R
PALG_E_356	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	TTTGGTTTAAA-N5-TTGAAAGCAAA	268708	268734	+	204	TU 33 (F)	1.R
PALG_E_357	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	ATAAAAAATAA-N6-TTATTTTTTAT	268808	268835	+	103	TU 33 (F)	R
PALG_E_358	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	AAATTATAA-N8-TTAAAAATT	268840	268865	+	73	TU 33 (F)	R
PALG_E_359	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	TAATTTTTA-N13-TAAAACTTA	268868	268898	+	40	TU 33 (F)	R.1
PALG_E_360	mglA	ribose ABC transport ATP-binding protein	AAATATTGAAT-N13-ATCCATTATTT	276102	276136	+	54	TU 33 (I)	R
PALG_E_362	mglA	ribose ABC transport ATP-binding protein	TTACAATTTTA-N14-TATAATTATAA	276152	276187	+	3	TU 33 (I)	R
PALG_E_363	MHP7448_0236	transposase	ATTCACTTT-N6-AATATGAAT	281861	281884	+	106	Class 1	1.R
PALG_E_364	MHP7448_0240	c hypothetical protein	TAAATTAAA-N8-TTAAAATTA	287253	287278	-	62	TU 35 (F)	1.R.2
PALG_E_365	MHP7448_0240	c hypothetical protein	TTTTACTATTT-N12-AAATACTAAAA	287293	287326	-	102	TU 35 (F)	1.R.2
PALG_E_366	secD	protein-export membrane protein	TTAAAATATA-N6-TATTTTTTAA	287429	287454	+	263	TU 36 (F)	1.R
PALG_E_367	secD	protein-export membrane protein	TAATTAAAATT-N6-AATTTATATTA	287538	287565	+	152	TU 36 (F)	1.R
PALG_E_368	secD	protein-export membrane protein	TCATATTAT-N6-ATGATATAA	287670	287693	+	24	TU 36 (F)	R
PALG_E_369	hisS	histidyl-tRNA synthetase	TTTTCTATT-N8-AATTAAAAA	290328	290353	+	51	TU 36 (I)	R
PALG_E_370	hisS	histidyl-tRNA synthetase	TTTTTTCTAAG-N14-CTTAAAAAAAA	290361	290396	+	8	TU 36 (I)	R
PALG_E_371	MHP7448_0244	conserved hypothetical protein	AAAAAATAA-N7-TTATTTTTT	293394	293418	+	43	TU 36 (I)	1.R
PALG_E_372	rpsT	30S ribosomal protein S20	TATTGTATA-N6-TATATAATA	293965	293988	+	69	TU 36 (L)	R
PALG_E_374	smf	c DNA processing protein SMF	ATAAAAAAAACTTTT-N6-AAAAGTTTTTTTTAT	304181	304216	-	46	TU 37 (F)	R
PALG_E_375	smf	c DNA processing protein SMF	TTATTTTTTTTA-N11-TATGAAAAATAA	304298	304332	-	163	TU 37 (F)	R
PALG_E_377	eno	enolase	ATAAAAAAAACTTTT-N6-AAAAGTTTTTTTTAT	304181	304216	+	211	TU 38 (F)	1.R
PALG_E_378	eno	enolase	ACTTTAATAA-N14-TTTTTAAATT	304314	304347	+	80	TU 38 (F)	1.R

PALG_E_379	dnaX	DNA polymerase III gamma and tau subunit	AATTAATTT-N11-AAATTTATT	309930	309958	+	22	TU 40 (F)	1.R
PALG_E_380	mgtA	c cation-transporting P-type ATPase	TATAACCTAA-N7-TTATTTTATA	324311	324337	-	36	mC 17	R.1
PALG_E_381	MHP7448_0272	c P97-like protein	AATTTTAAA-N9-TTGAAAATT	331844	331870	-	104	TU 42 (F)	1.R
PALG_E_382	MHP7448_0272	c P97-like protein	TAAAAAGTTG-N11-CATTTTTTTA	331883	331913	-	143	TU 42 (F)	2.R
PALG_E_383	MHP7448_0272	c P97-like protein	TTATAAAAA-N8-TTTTTATTA	331923	331948	-	183	TU 42 (F)	2.R
PALG_E_385	MHP7448_0272	c P97-like protein	TAATAAAAA-N8-TTTTTATAA	332031	332056	-	291	TU 42 (F)	2.R
PALG_E_386	rpmF	50S ribosomal protein L32	TTTATGAAAA-N9-TTATAATAAA	335109	335137	+	8	TU 43 (I)	R
PALG_E_387	MHP7448_0279	c transcriptional regulator	TTTTAAAAATTTTTAA-N9-TTAAAAATTTTTAAAA	339289	339329	-	21	mC 18	1.R.2
PALG_E_388	MHP7448_0279	c transcriptional regulator	AAAAGTTAA-N7-TTAATTTTT	339335	339359	-	67	mC 18	1.R.2
PALG_E_389	MHP7448_0279	c transcriptional regulator	TAACTAAATTT-N5-AAATTAAGTTA	339396	339422	-	128	mC 18	1.R.1
PALG_E_390	MHP7448_0279	c transcriptional regulator	AAAAGTTTC-N15-GAAACCTTT	339431	339463	-	163	mC 18	2.R.1
PALG_E_391	pyrG	CTP synthase	ATAAATTTTT-N10-AAAAAACTAT	339471	339500	+	285	mC 19	3.R
PALG_E_392	pyrG	CTP synthase	AAATCAAAAA-N10-TTTTTGATAT	339625	339654	+	131	mC 19	2.R.1
PALG_E_393	pyrG	CTP synthase	TATTTTTAA-N13-TTTATAATA	339672	339702	+	83	mC 19	1.R.1
PALG_E_394	MHP7448_0282	c hypothetical protein	ATATTTTTA-N8-TAAGAAAAT	342622	342647	-	3	TU 44 (I)	R
PALG_E_395	rplS	c 50S ribosomal protein L19	TTATAATTT-N8-AAATTAAAA	344944	344969	-	14	TU 44 (I)	R
PALG_E_396	rpsP	c 30S ribosomal protein S16	AAATCTTTT-N7-AAACGACTT	346105	346129	-	151	TU 44 (I)	1.R
PALG_E_398	MHP7448_0287	c hypothetical protein	TAAAATTTT-N5-AAAATTTTA	346675	346697	-	16	TU 44 (F)	R.1
PALG_E_399	MHP7448_0287	c hypothetical protein	ATTTACTTTT-N10-AAAAGTTTAT	346860	346889	-	201	TU 44 (F)	R.1
PALG_E_400	MHP7448_0289	hypothetical protein	TTCTTCAAAA-N12-TTTTGATAAA	347591	347622	+	106	TU 45 (F)	R
PALG_E_401	MHP7448_0289	hypothetical protein	AATTTTTTTT-N6-AAGAAAAATT	347630	347655	+	73	TU 45 (F)	R
PALG_E_402	MHP7448_0289	hypothetical protein	AAATTTAACTT-N12-AATTTAAATTT	347662	347695	+	33	TU 45 (F)	R
PALG_E_403	rpsF	c 30S ribosomal protein S6	AAAAAATTGTAAA-N8-TTTAAAACTTTTT	357503	357536	-	52	TU 46 (I)	R.1
PALG_E_404	rpsF	c 30S ribosomal protein S6	AAAAAAATAAAA-N10-TTTGATTTCTTT	357632	357665	-	181	TU 46 (I)	1.R
PALG_E_405	rpsF	c 30S ribosomal protein S6	TGATAATTAT-N13-AAAATTGTCA	357685	357717	-	234	TU 46 (I)	1.R
PALG_E_406	MHP7448_0300	hypothetical protein	AATTTGAAC-N9-GTTCATATT	361890	361916	+	216	TU 47 (I)	2.R
PALG_E_407	MHP7448_0300	hypothetical protein	ATAATTTAAT-N14-ATAAAATAAT	362027	362060	+	72	TU 47 (I)	R.2

PALG_E_408	MHP7448_0301	hypothetical protein	AAAATTATC-N14-GATGGTTTT	363968	363999	+	237	TU 47 (I)	R
PALG_E_409	MHP7448_0301	hypothetical protein	AAGACAAAA-N14-TTATGTCTT	364121	364152	+	84	TU 47 (I)	R
PALG_E_410	MHP7448_0301	hypothetical protein	ATTTAAATC-N14-GCTCTAAAT	364157	364188	+	48	TU 47 (I)	R
PALG_E_411	MHP7448_0302	permease	AAATTTATTATAT-N8-ATATAGTAAATTT	364562	364595	+	21	TU 47 (I)	R
PALG_E_412	MHP7448_0305	c ABC transporter ATP-binding protein	ATATAAAAAAA-N9-TTTTTTAATTT	369338	369368	-	15	TU 48 (L)	R
PALG_E_413	MHP7448_0308	conserved hypothetical protein	AAAATGCTAT-N9-ATAGGATTGT	372904	372932	+	104	TU 49 (L)	3.R
PALG_E_414	MHP7448_0308	conserved hypothetical protein	TATTTAGTTAT-N15-ATAATTATATA	372963	372999	+	37	TU 49 (L)	R
PALG_E_415	MHP7448_0312	c glycine cleavage system H protein	TTAGATATT-N5-AAAATTTAA	382670	382692	-	10	TU 50 (F)	R.1
PALG_E_416	MHP7448_0313	hypothetical protein	TAAATAATT-N12-AATTATGGA	382930	382959	+	107	TU 51 (F)	1.R
PALG_E_417	MHP7448_0313	hypothetical protein	ATATTTTGC-N7-GTATAAATAT	382992	383016	+	50	TU 51 (F)	R.1
PALG_E_418	MHP7448_0315	ABC transporter ATP-binding protein	AAAAATTAT-N7-ATATTTTTT	386230	386254	+	60	TU 51 (I)	R
PALG_E_421	MHP7448_0318	hypothetical protein	TAAAAATTT-N8-AAATTATAA	393010	393035	+	173	TU 51 (I)	1.R
PALG_E_422	MHP7448_0318	hypothetical protein	AACCAGTTA-N10-TCAATGGTT	393049	393076	+	132	TU 51 (I)	1.R
PALG_E_423	MHP7448_0322	c putative transposase	TAGATTTTT-N15-AAAAATTAA	396568	396600	-	140	Class 1	R.1
PALG_E_424	MHP7448_0323	transposase	TTTTTCTAGA-N12-TCTATAAAAA	396762	396793	+	251	Class 1	1.R
PALG_E_425	MHP7448_0323	transposase	TGCTTAAAA-N10-TTTTTACCA	396797	396824	+	220	Class 1	1.R
PALG_E_426	MHP7448_0323	transposase	ATTCACTTT-N6-AATATGAAT	396915	396938	+	106	Class 1	1.R
PALG_E_427	MHP7448_0325	c hypothetical protein	AAAATTAAAA-N12-TTATAATTTT	400966	400997	-	19	TU 52 (F)	R
PALG_E_429	MHP7448_0326	c hypothetical protein	TTATCTTCA-N10-TCAAGATCA	402326	402353	-	42	TU 53 (L)	R
PALG_E_430	MHP7448_0327	c conserved hypothetical protein	AAATAAAAT-N10-AGTTTATTT	403804	403831	-	0	TU 53 (F)	R.2
PALG_E_431	MHP7448_0327	c conserved hypothetical protein	TAAAAGCCC-N9-GAGGTTTTA	403915	403941	-	111	TU 53 (F)	2.R
PALG_E_432	MHP7448_0328	hypothetical protein	ATAGATAAAA-N14-TTTTTTCTAT	404192	404225	+	86	TU 54 (F)	1.R.2
PALG_E_433	MHP7448_0328	hypothetical protein	TTATATTAT-N8-ATAATAGTA	404269	404294	+	17	TU 54 (F)	R.2
PALG_E_434	MHP7448_0331	c hypothetical protein	AATTCCGGCA-N15-TGCAGAAATT	409584	409618	-	86	TU 55 (L)	R
PALG_E_435	MHP7448_0332	c subtilisin-like serine protease	TTTATTTAAAA-N14-TTTTCAATAAA	411828	411863	-	91	TU 55 (I)	2.R
PALG_E_436	MHP7448_0332	c subtilisin-like serine protease	TCATAATTT-N7-AAAGTATAA	411873	411897	-	136	TU 55 (I)	2.R
PALG_E_437	MHP7448_0332	c subtilisin-like serine protease	AATTTTAAA-N13-TTGAAAATT	411925	411955	-	188	TU 55 (I)	2.R

PALG_E_438	MHP7448_0333	c lipoprotein	TTTGTTTTT-N9-AAAAATAAA	412735	412761	-	23	TU 55 (F)	R.3
PALG_E_439	MHP7448_0333	c lipoprotein	AATTCAAAATTA-N14-TAATTTTAAATT	412799	412836	-	87	TU 55 (F)	R.2
PALG_E_440	MHP7448_0333	c lipoprotein	AAAAATTTAGA-N14-TCTATAATTTT	412977	413012	-	265	TU 55 (F)	1.R
PALG_E_441	MHP7448_0334	hypothetical protein	AATTTATTA-N11-TAAAAAATT	412833	412861	+	263	TU 56 (F)	1.R
PALG_E_442	MHP7448_0334	hypothetical protein	TGTCTAAATT-N9-AAATTAGCCA	413000	413028	+	96	TU 56 (F)	R.2
PALG_E_443	MHP7448_0334	hypothetical protein	TTTTTAGCTATTTT-N6-AAAATAACTAAAAA	413031	413064	+	60	TU 56 (F)	R.2
PALG_E_444	MHP7448_0334	hypothetical protein	ATTAAAAAA-N6-TTTTATAAT	413069	413092	+	32	TU 56 (F)	R.2
PALG_E_445	MHP7448_0336	c hypothetical protein	AATTTAGAAT-N8-AGTCTAAATT	415992	416019	-	77	mC 21	1.R
PALG_E_446	MHP7448_0336	c hypothetical protein	TTCAAATTTT-N13-AACATTTGCA	416026	416058	-	111	mC 21	1.R
PALG_E_447	MHP7448_0336	c hypothetical protein	TCATAAAAAAT-N12-ATATTTTATGA	416164	416197	-	249	mC 21	1.R
PALG_E_448	MHP7448_0336	c hypothetical protein	AAACAAATTA-N7-TATTTTTTTT	416205	416231	-	290	mC 21	1.R
PALG_E_449	MHP7448_0337	hypothetical protein	ATAATTAATTAT-N12-ATATTTAATTAT	416147	416182	+	237	mC 22	R.1
PALG_E_450	MHP7448_0342	c hypothetical protein	AAAAATATAATTTTT-N13-AAAAATTATATTTTT	426677	426719	-	78	TU 57 (I)	R.1
PALG_E_451	MHP7448_0342	c hypothetical protein	ATTTATAGT-N12-AATATGAAT	426720	426749	-	121	TU 57 (I)	R.1
PALG_E_452	MHP7448_0342	c hypothetical protein	AAATATTTAA-N11-TTTTATATTT	426755	426785	-	156	TU 57 (I)	R.1
PALG_E_453	MHP7448_0343	pseudogene	TATTTAAATAA-N11-TTATTTAAATA	427199	427231	+	236	Class 2	1.R
PALG_E_454	MHP7448_0343	pseudogene	AAGTTAAAA-N9-TTTTAGTTT	427347	427373	+	94	Class 2	R
PALG_E_455	MHP7448_0344	c hypothetical protein	ATTTATTATAT-N9-ATATACTAAAT	429243	429273	-	28	TU 57 (F)	R.1
PALG_E_456	MHP7448_0344	c hypothetical protein	AAAAAATTT-N8-AAATTTATT	429281	429306	-	66	TU 57 (F)	1.R.1
PALG_E_457	MHP7448_0344	c hypothetical protein	AATTATAAA-N10-TTTTTCATT	429324	429351	-	109	TU 57 (F)	2.R.1
PALG_E_458	MHP7448_0347	hypothetical protein	TTTAATATAAA-N9-TTTATATGAAA	432353	432383	+	51	TU 58 (L)	R
PALG_E_459	MHP7448_0348	c hypothetical protein	ATATTTAAT-N15-ATTATATAT	434215	434247	-	29	TU 59 (L)	R
PALG_E_460	MHP7448_0348	c hypothetical protein	AATCAAACTTA-N14-TTAGTTTTATT	434485	434520	-	299	TU 59 (L)	R
PALG_E_461	MHP7448_0351	c conserved hypothetical protein	ATAATTTTA-N11-TATAATTAT	438554	438582	-	29	TU 59 (F)	R.1
PALG_E_462	MHP7448_0351	c conserved hypothetical protein	TTTTAAGAA-N14-TTATTCAAA	438590	438621	-	65	TU 59 (F)	2.R.1
PALG_E_463	MHP7448_0351	c conserved hypothetical protein	AAAATTATAT-N6-ATATAGTTTT	438630	438655	-	105	TU 59 (F)	2.R
PALG_E_465	MHP7448_0352	hypothetical protein	TTTTTTAGA-N14-TATAATAAA	438536	438567	+	116	TU 60 (F)	3.R

PALG_E_466	MHP7448_0352	hypothetical protein	AAATTATTA-N11-TAATTATTT	438576	438604	+	79	TU 60 (F)	2.R
PALG_E_467	MHP7448_0356	c hypothetical protein	AAAATTAAAAT-N11-ATTTAAATTTT	446183	446215	-	23	TU 61 (F)	R.1
PALG_E_469	MHP7448_0357	amino acid permease	AAAATTATT-N15-AATATTTTT	446199	446231	+	41	mC 25	1.R.1
PALG_E_470	glpK	c glycerol kinase	TTTGGATTA-N8-TAATTCAAA	450563	450588	-	28	TU 62 (F)	R
PALG_E_471	glpK	c glycerol kinase	TAAATTTTAT-N6-ATAAGATTTA	450640	450665	-	105	TU 62 (F)	1.R
PALG_E_472	MHP7448_0360	P37-like ABC transporter substrate-binding lipoprotein	TAAATCTTAT-N6-ATAAAATTTA	450640	450665	+	207	TU 63 (F)	1.R
PALG_E_473	MHP7448_0360	P37-like ABC transporter substrate-binding lipoprotein	AATTTATAAT-N15-ATAATAATTT	450693	450727	+	145	TU 63 (F)	R.1
PALG_E_474	MHP7448_0360	P37-like ABC transporter substrate-binding lipoprotein	GTCTTTTTT-N12-ATAAAACAC	450807	450836	+	36	TU 63 (F)	R.2
PALG_E_476	MHP7448_0365	hypothetical protein	AATTTATTAA-N8-TTATTAATTT	456509	456536	+	214	TU 63 (I)	R
PALG_E_477	MHP7448_0366	lipoprotein	ATTTTAAATT-N5-AATATAAATT	457570	457594	+	43	TU 63 (I)	R
PALG_E_478	MHP7448_0373	c conserved hypothetical protein	AAAATGTTATA-N10-TATAAAACTTT	473302	473333	-	68	TU 64 (I)	R
PALG_E_479	MHP7448_0375	c PTS system enzyme IIB component	TTTTTAATT-N15-AATTAAAAA	474231	474263	-	19	TU 64 (I)	R
PALG_E_480	MHP7448_0377	c conserved hypothetical protein	TTAACGAAA-N8-TTTATTTAA	477564	477589	-	3	TU 64 (I)	R
PALG_E_481	MHP7448_0377	c conserved hypothetical protein	AAAATGTTTAA-N13-TAAAAAATTTT	477607	477641	-	46	TU 64 (I)	R
PALG_E_482	MHP7448_0378	c lipoprotein	AAGATTATT-N8-AAAACTCTT	479762	479787	-	201	TU 64 (F)	1.R
PALG_E_483	MHP7448_0379	ABC transport ATP-binding protein	TACCTTCAA-N9-TTGAATATA	479873	479899	+	29	TU 65 (F)	R.2
PALG_E_484	MHP7448_0386	pseudogene	ACTTTTTAA-N14-TTAAAATTT	487849	487880	+	153	Class 2	1.R
PALG_E_485	MHP7448_0386	pseudogene	AATTAGCTAA-N5-TTAGCAAATT	487937	487961	+	72	Class 2	1.R
PALG_E_486	MHP7448_0390	c hypothetical protein	TTAACTTTGC-N11-GCTAACTTAA	492336	492366	-	71	TU 66 (F)	1.R.1
PALG_E_487	MHP7448_0390	c hypothetical protein	TAAATTCTA-N11-TATAATTTA	492487	492515	-	222	TU 66 (F)	1.R
PALG_E_488	MHP7448_0390	c hypothetical protein	AATTTTATA-N14-TATTAAATT	492531	492562	-	266	TU 66 (F)	2.R
PALG_E_489	MHP7448_0391	conserved hypothetical protein	TTATTTTTTTC-N15-GTAAAAAGTAA	492917	492953	+	82	TU 67 (F)	1.R
PALG_E_490	MHP7448_0394	conserved hypothetical protein	TTTTAAATA-N14-TTTCTAAAA	495411	495442	+	168	TU 67 (I)	1.R
PALG_E_492	MHP7448_0397	c conserved hypothetical protein	TTTTGATTG-N10-CAATAAAAA	499913	499940	-	122	TU 68 (L)	1.R
PALG_E_494	MHP7448_0399	c hypothetical protein	TTATTAATTA-N14-TCATTAATAA	501332	501365	-	66	TU 68 (F)	1.R
PALG_E_495	MHP7448_0399	c hypothetical protein	ATAATGAAA-N10-TTTCCTTTT	501460	501487	-	194	TU 68 (F)	1.R
PALG_E_496	MHP7448_0400	hypothetical protein	AAAAGGAAA-N10-TTTCATTAT	501460	501487	+	230	TU 69 (F)	R.1

PALG_E_497	MHP7448_0400	hypothetical protein	ATAGTTATTT-N15-AAATAAAAAT	501533	501567	+	150	TU 69 (F)	R.1
PALG_E_498	asnS	asparaginyl-tRNA synthetase	AAATTTTTT-N10-AACAAATTT	503449	503476	+	287	TU 69 (I)	1.R
PALG_E_499	asnS	asparaginyl-tRNA synthetase	AATTTTAAAA-N10-TTTTATAATT	503495	503524	+	239	TU 69 (I)	1.R
PALG_E_500	asnS	asparaginyl-tRNA synthetase	TAAAAAATT-N15-ATTTTTCTA	503636	503668	+	95	TU 69 (I)	1.R
PALG_E_501	asnS	asparaginyl-tRNA synthetase	TAATTTTAA-N10-TTTTAATTA	503676	503703	+	60	TU 69 (I)	1.R
PALG_E_503	MHP7448_0408	c hypothetical protein	TTGAAAAAT-N6-ATTTTTTAA	514274	514297	-	30	TU 70 (I)	R
PALG_E_504	MHP7448_0410	pseudogene	ACTTTTTAA-N14-TTAAAATTT	517169	517200	-	162	Class 2	R
PALG_E_505	MHP7448_0411	c hypothetical protein	TTACTAAATTT-N13-AAATAAAGTAA	517716	517750	-	7	TU 70 (I)	R.2
PALG_E_506	MHP7448_0411	c hypothetical protein	TTAAAAATT-N7-AATTTATAA	517796	517820	-	87	TU 70 (I)	R.1
PALG_E_507	MHP7448_0411	c hypothetical protein	ATTTTTAAA-N13-TTCAGAAAT	517847	517877	-	138	TU 70 (I)	1.R.1
PALG_E_508	MHP7448_0411	c hypothetical protein	TAAAATGTTT-N15-ACAAATTTTA	517904	517938	-	195	TU 70 (I)	1.R
PALG_E_509	MHP7448_0411	c hypothetical protein	TCCAAAAAA-N14-TTATTTGTA	517956	517987	-	247	TU 70 (I)	2.R
PALG_E_510	MHP7448_0412	c conserved hypothetical protein	TTTACAAGAT-N7-AACTTGCAAA	519142	519168	-	85	TU 70 (I)	R
PALG_E_511	MHP7448_0413	c conserved hypothetical protein	TTTTACTTA-N15-TAATTAAAA	519640	519672	-	23	TU 70 (I)	1.R
PALG_E_512	MHP7448_0429	hypothetical protein	TTTATTTTTGT-N6-ACATTAATAAA	543241	543268	+	132	TU 72 (I)	1.R
PALG_E_513	MHP7448_0429	hypothetical protein	AAGTAAAAA-N8-TTTTTATTT	543273	543298	+	102	TU 72 (I)	1.R
PALG_E_514	MHP7448_0431	c conserved hypothetical protein	ATTTAACCT-N8-AGGAAAAAT	546699	546724	-	17	TU 73 (L)	R
PALG_E_515	MHP7448_0431	c conserved hypothetical protein	TTTTTTTAACTT-N14-AAGGTAAAAAAA	546784	546821	-	102	TU 73 (L)	R
PALG_E_517	sgaH	c 3-hexulose-6-phosphate synthase	TTTTTTGGTAATTTT-N6-AAAATTCGCAAAAAA	554449	554484	-	98	TU 73 (I)	1.R
PALG_E_519	MHP7448_0440	c hypothetical protein	TCGAAAATT-N11-AATTATCAA	557928	557956	-	102	TU 73 (F)	R.1
PALG_E_520	MHP7448_0440	c hypothetical protein	ATCAACTTT-N9-AAATTTTAT	557965	557991	-	139	TU 73 (F)	R.1
PALG_E_521	MHP7448_0441	hypothetical protein	TTGATAATT-N11-AATTTTCGA	557928	557956	+	44	mC 27	R.1
PALG_E_522	MHP7448_0441	hypothetical protein	ATAAAATTT-N9-AAAGTTGAT	557965	557991	+	9	mC 27	R.1
PALG_E_523	MHP7448_0445	c hypothetical protein	TTTTAAAATT-N10-AATTTGAAAA	581369	581398	-	24	mC 28	R
PALG_E_524	MHP7448_0445	c hypothetical protein	TAAAACTACC-N5-GGTTTTTTTA	581437	581461	-	92	mC 28	R
PALG_E_525	MHP7448_0447	hypothetical protein	TGCCCCTGC-N10-GCTGAGGCA	585257	585284	+	273	TU 76 (I)	R
PALG_E_526	MHP7448_0448	hypothetical protein	TTGTTGATAC-N10-GTATCAACAA	586045	586074	+	275	TU 76 (L)	R

PALG_E_527	MHP7448_0448	hypothetical protein	TGATAATTT-N6-AAAGAATCA	586296	586319	+	30	TU 76 (L)	R
PALG_E_528	MHP7448_0449	c GTP-binding protein	AAGTAAGTA-N10-TATTTAATT	587485	587512	-	1	TU 77 (L)	R
PALG_E_529	MHP7448_0449	c GTP-binding protein	ATTTATTTAGTT-N7-AACTAAATTAAT	587542	587572	-	58	TU 77 (L)	1.R
PALG_E_531	metK	c S-adenosylmethionine synthetase	TTTTTTCTAAAAAA-N9-TTTTTTAGAAAAAA	589841	589877	-	35	TU 77 (I)	1.R
PALG_E_534	acpD	acyl carrier protein phosphodiesterase	TAATTTTGT-N14-ACAAAATTA	598979	599010	+	37	TU 78 (F)	R
PALG_E_535	acpD	acyl carrier protein phosphodiesterase	TCCAAAAAA-N13-TTAATTGGA	599011	599041	+	6	TU 78 (F)	R.1
PALG_E_537	MHP7448_0461	hypothetical protein	AATTAAAAA-N13-TTTAGAATT	603531	603561	+	154	TU 80 (F)	1.R
PALG_E_538	MHP7448_0461	hypothetical protein	TTTAAATTTT-N13-AAATTGTAAA	603668	603700	+	15	TU 80 (F)	R.1
PALG_E_539	MHP7448_0463	hypothetical protein	AAAAAACATC-N11-GATTTTTTAT	605363	605393	+	240	TU 80 (L)	R
PALG_E_540	pepA	c leucyl aminopeptidase	ATTAGAAAAA-N14-TATTTCTAAT	609429	609462	-	76	TU 81 (L)	1.R
PALG_E_541	MHP7448_0466	c hypothetical protein	TTTTATCTG-N6-CAAAAAAAA	610914	610937	-	9	TU 81 (I)	R.1
PALG_E_542	MHP7448_0466	c hypothetical protein	ATAAAAAAA-N14-TTTTTTTAT	611015	611046	-	110	TU 81 (I)	1.R
PALG_E_543	MHP7448_0470	c conserved hypothetical protein	CAGTTTTAGA-N15-TTTAAAAGTG	616652	616686	-	88	TU 81 (F)	R.1
PALG_E_547	nadE	NH 3 -dependent NAD synthetase	TTAAATTTAAA-N7-TTTATATTAAA	619479	619507	+	136	TU 82 (I)	2.R
PALG_E_550	MHP7448_0477	c conserved hypothetical protein	TTATATTTTAA-N5-TTATAGTATAA	623798	623824	-	9	TU 83 (I)	R
PALG_E_551	MHP7448_0477	c conserved hypothetical protein	TAATAATAA-N10-TTATTTTTA	623835	623862	-	46	TU 83 (I)	1.R
PALG_E_552	MHP7448_0477	c conserved hypothetical protein	AATTAAAAA-N9-TTATTAATT	623941	623967	-	152	TU 83 (I)	1.R
PALG_E_553	MHP7448_0477	c conserved hypothetical protein	ATTTTTAAT-N13-ATTAAAAAT	624003	624033	-	214	TU 83 (I)	1.R
PALG_E_554	MHP7448_0484	c hypothetical protein	ATGAATTTT-N5-AAAATTTCT	633199	633221	-	221	TU 83 (F)	1.R
PALG_E_555	MHP7448_0485	conserved hypothetical protein	AGAAATTTT-N5-AAAATTCAT	633199	633221	+	77	mC 29	R.2
PALG_E_556	MHP7448_0488	c conserved hypothetical protein	ATTGAAATAT-N7-ATAATTAAAT	637099	637125	-	67	TU 84 (I)	1.R
PALG_E_557	MHP7448_0488	c conserved hypothetical protein	AAAAATTTTGT-N7-ACAAAATTTTT	637227	637255	-	195	TU 84 (I)	1.R
PALG_E_558	MHP7448_0489	c hypothetical protein	TTTAAATTAT-N9-ATAATCAAAA	639931	639959	-	16	TU 84 (F)	R.1
PALG_E_559	MHP7448_0489	c hypothetical protein	AAATTTTAT-N9-ATATAATCT	639976	640002	-	61	TU 84 (F)	1.R.1
PALG_E_560	MHP7448_0489	c hypothetical protein	TATTTTTATT-N15-AATTTAAATA	640025	640059	-	110	TU 84 (F)	1.R
PALG_E_561	pgk	phosphoglycerate kinase	ATAATTTAAATT-N9-AATTGAGATTAT	639950	639982	+	282	TU 85 (F)	2.R
PALG_E_562	pgk	phosphoglycerate kinase	AAATAAAATTTT-N14-AAAAATCTATTT	639992	640029	+	235	TU 85 (F)	2.R.1

PALG_E_563	MHP7448_0491	hypothetical protein	GAGCTTTAAA-N13-TTTAAATCGC	641482	641514	+	140	TU 85 (I)	1.R
PALG_E_564	MHP7448_0491	hypothetical protein	TAACTAATATTTTT-N13-AAAAATATTAGTTA	641525	641565	+	89	TU 85 (I)	1.R
PALG_E_565	fruA	PTS system, fructose-specific IIABC component	AAAAATTTA-N8-TATATTTTT	641910	641935	+	85	TU 85 (I)	R
PALG_E_566	MHP7448_0495	c ISMHp1 transposase	ATTCACTTT-N6-AATATGAAT	647174	647197	-	106	Class 1	1.R
PALG_E_568	MHP7448_0495	c ISMHp1 transposase	AAATAATGAT-N6-ATAATTTTTT	647279	647304	-	211	Class 1	1.R
PALG_E_570	MHP7448_0496	c putative p216 surface protein	AATTTATAC-N13-GGAAAAATT	653219	653249	-	23	TU 86 (L)	R.1
PALG_E_571	MHP7448_0496	c putative p216 surface protein	AAAAAATTT-N15-AAAATTTTT	653263	653295	-	67	TU 86 (L)	1.R
PALG_E_572	MHP7448_0497	c p76 membrane protein precursor	AATTTTAAA-N15-TTTAACATT	657585	657617	-	0	TU 86 (I)	R
PALG_E_573	MHP7448_0497	c p76 membrane protein precursor	AAAAATAAGGAAAA-N8-TTTTTCATATTTTT	657807	657842	-	222	TU 86 (I)	1.R
PALG_E_574	MHP7448_0500	c conserved hypothetical protein	TTTTTTAATAATT-N13-ATTTATGAAAAAA	661864	661902	-	40	TU 86 (I)	R
PALG_E_575	MHP7448_0500	c conserved hypothetical protein	TAAATCTTAATT-N6-AATTAAAAATTA	661903	661932	-	79	TU 86 (I)	R
PALG_E_576	MHP7448_0505	c lipoprotein	AAAATAGTA-N7-TATTTTTTT	670444	670468	-	30	TU 86 (F)	1.R.1
PALG_E_577	MHP7448_0505	c lipoprotein	TAGTAAATT-N10-AATTTAAAA	670478	670505	-	64	TU 86 (F)	1.R.1
PALG_E_578	MHP7448_0505	c lipoprotein	TTTTACTAA-N11-TAAGAAAAA	670570	670598	-	156	TU 86 (F)	1.R.1
PALG_E_579	MHP7448_0505	c lipoprotein	ATAATATTT-N7-AAATACAAT	670608	670632	-	194	TU 86 (F)	1.R
PALG_E_580	pdhC	dihydrolipoamide acetyltransferase	TTTTTTACTTTTA-N5-TAAAGGCAAAAAA	671164	671194	+	257	TU 87 (F)	3.R
PALG_E_581	pdhC	dihydrolipoamide acetyltransferase	TTTTCCATAA-N7-TTATGCTAAA	671305	671331	+	120	TU 87 (F)	2.R.1
PALG_E_582	pdhC	dihydrolipoamide acetyltransferase	ATAATTATA-N15-TGAAATTAT	671367	671399	+	52	TU 87 (F)	1.R.1
PALG_E_583	pta	c phosphate acetyltransferase	ATTTTTTAA-N7-TTAGAAAAT	676713	676737	-	73	TU 88 (I)	1.R
PALG_E_584	MHP7448_0511	c conserved hypothetical protein	GTTTTATAT-N10-AAATAGAAC	678684	678711	-	107	TU 88 (I)	1.R
PALG_E_585	MHP7448_0512	c hypothetical protein	TATAAATTA-N5-TAATTTTTA	680061	680083	-	177	TU 88 (F)	R.1
PALG_E_586	MHP7448_0513	46K surface antigen precursor	AAAATTTTA-N10-TAACAATTT	680032	680059	+	212	TU 89 (F)	2.R.1
PALG_E_587	MHP7448_0513	46K surface antigen precursor	AAAAATTAC-N13-GTATTTTTT	680062	680092	+	179	TU 89 (F)	2.R.2
PALG_E_588	MHP7448_0513	46K surface antigen precursor	TATTTGAATTTTT-N14-AAAAATACAAATA	680130	680169	+	102	TU 89 (F)	2.R.2
PALG_E_589	MHP7448_0513	46K surface antigen precursor	AAAAAAATT-N14-AATTTATTT	680204	680235	+	36	TU 89 (F)	R.2
PALG_E_591	MHP7448_0520	c hypothetical protein	AAAAAATTT-N13-AAATATTTT	688666	688696	-	197	TU 90 (I)	2.R
PALG_E_592	MHP7448_0520	c hypothetical protein	TTGTTAATT-N9-AACTAGCAA	688699	688725	-	230	TU 90 (I)	2.R

PALG_E_593	pepF	c oligoendopeptidase F	TTATTTTCT-N7-AGAAGGTAA	691415	691439	-	5	TU 90 (F)	R.1
PALG_E_594	pepF	c oligoendopeptidase F	AATTTAAAAT-N11-ATTTTAAAAT	691502	691532	-	92	TU 90 (F)	R.1
PALG_E_595	MHP7448_0522	conserved hypothetical protein	AAATAGAAAA-N6-TTTAATATTT	691427	691452	+	185	mC 30	1.R
PALG_E_597	tufA	c elongation factor EF-Tu	ATTTTTTTA-N13-TAAAATAAT	699074	699104	-	205	TU 91 (L)	4.R
PALG_E_598	tufA	c elongation factor EF-Tu	TTTAAAAACT-N15-AATTTCTAAA	699107	699141	-	238	TU 91 (L)	4.R
PALG_E_599	deoC	c deoxyribose-phosphate aldolase	AAAATTTTA-N9-TAGTATTTT	704054	704080	-	33	TU 91 (F)	R.3
PALG_E_600	deoC	c deoxyribose-phosphate aldolase	TAAGTAAAA-N12-TTTTACAAA	704085	704114	-	64	TU 91 (F)	R.3
PALG_E_601	deoC	c deoxyribose-phosphate aldolase	AGTTTTTAT-N10-ATCAAAACT	704230	704257	-	209	TU 91 (F)	1.R.1
PALG_E_602	deoC	c deoxyribose-phosphate aldolase	AAATTTGAC-N13-GTAAAATTT	704305	704335	-	284	TU 91 (F)	R
PALG_E_603	gyrA	DNA gyrase subunit A	AGTTTTGAT-N10-ATAAAAACT	704230	704257	+	203	mC 31	2.R
PALG_E_604	gyrA	DNA gyrase subunit A	AAATTTTAC-N13-GTCAAATTT	704305	704335	+	125	mC 31	1.R.1
PALG_E_605	gyrA	DNA gyrase subunit A	GACTTTTTA-N13-TAAAAATAC	704337	704367	+	93	mC 31	1.R.1
PALG_E_606	gyrA	DNA gyrase subunit A	TTTTCTGTTT-N14-AAACATAAAA	704374	704407	+	53	mC 31	1.R.1
PALG_E_607	pmsR	c peptide methionine sulfoxide reductase	TTTTTTTGA-N9-TTTAAAAAA	707546	707572	-	45	TU 92 (L)	1.R
PALG_E_609	MHP7448_0536	hypothetical protein	TTATACTTT-N13-AATGTATTA	716193	716223	+	105	TU 93 (F)	R
PALG_E_610	potA	c spermidine/putrescine ABC transporter ATP-binding	TATTTTTATTTT-N15-AAAATTTAAATA	723835	723873	-	0	TU 94 (F)	R.1
PALG_E_611	potA	c spermidine/putrescine ABC transporter ATP-binding	TAAGAATAAAA-N5-TTTTATTCAAA	723882	723908	-	47	TU 94 (F)	R.1
PALG_E_612	MHP7448_0543	hypothetical protein	AAATATTGT-N8-ACGATATTT	724775	724800	+	33	TU 95 (I)	R.1
PALG_E_613	ktrA	c potassium uptake protein	TTTTAAATTTAT-N9-ATAATTTTAAAA	727319	727351	-	32	mC 32	1.R.1
PALG_E_614	ktrA	c potassium uptake protein	AAAAAATACT-N9-AGTGTTTTTT	727386	727414	-	99	mC 32	1.R.1
PALG_E_615	ktrA	c potassium uptake protein	ATCTAAAAGAA-N6-TTTTTTTATAT	727420	727447	-	133	mC 32	1.R.1
PALG_E_616	ktrA	c potassium uptake protein	ATAATTATA-N8-TATAAAAAT	727459	727484	-	172	mC 32	1.R
PALG_E_617	ktrA	c potassium uptake protein	TTTTATATAT-N12-AAATCTAAAA	727511	727542	-	224	mC 32	2.R
PALG_E_619	ktrB	potassium uptake protein	TGTTTTTATA-N15-TATAAAAAAA	727396	727430	+	153	mC 33	1.R
PALG_E_620	ktrB	potassium uptake protein	AAAACTAGT-N11-ACTAACTTT	727468	727496	+	87	mC 33	1.R
PALG_E_622	MHP7448_0547	c conserved hypothetical protein	AAAATTCAAT-N8-ATTTTATTTT	733677	733704	-	244	TU 96 (L)	R
PALG_E_632	MHP7448_0556	c hypothetical protein	AGATTATTT-N15-AAAATATCT	741610	741642	-	191	TU 97 (I)	R

PALG_E_633	MHP7448_0559	c conserved hypothetical protein	TAAAATTAA-N7-TTAATATTA	745503	745527	-	31	TU 97 (I)	R.1
PALG_E_634	MHP7448_0559	c conserved hypothetical protein	AATTACATAA-N5-TTATAGAATT	745550	745574	-	78	TU 97 (I)	R.1
PALG_E_635	MHP7448_0559	c conserved hypothetical protein	AATATGAAAA-N9-TTTTTCTATT	745593	745621	-	121	TU 97 (I)	R
PALG_E_636	MHP7448_0559	c conserved hypothetical protein	AAAAAATCAATAAA-N5-TTTATTAATTTTTT	745628	745660	-	156	TU 97 (I)	1.R
PALG_E_637	MHP7448_0559	c conserved hypothetical protein	AACTTATCC-N10-GGATATTTT	745673	745700	-	201	TU 97 (I)	1.R
PALG_E_638	MHP7448_0562	pseudogene	TCTTTTAAAA-N8-TTTTAAATGA	749544	749571	-	51	Class 2	1.R
PALG_E_640	MHP7448_0564	c hypothetical protein	GACTAAAAAC-N15-GTTATTAGTC	752506	752540	-	116	TU 99 (F)	R
PALG_E_641	MHP7448_0565	pseudogene	TCTTTTAAAA-N8-TTTTAAATGA	753480	753507	-	106	Class 2	1.R
PALG_E_642	MHP7448_0565	pseudogene	AAATTAGCA-N10-TATTAATTT	753510	753537	-	136	Class 2	1.R
PALG_E_643	MHP7448_0568	pseudogene	TCTTTTAAAA-N8-TTTTAAATGA	757398	757425	-	88	Class 2	1.R
PALG_E_644	MHP7448_0569	hypothetical protein	TCATTTAAAA-N8-TTTTAAAAGA	757398	757425	+	20	TU 101 (I)	1.R
PALG_E_645	MHP7448_0570	conserved hypothetical protein	TTAAATTTCTAT-N11-ATAAAAAATTTAA	757731	757765	+	180	TU 101 (L)	1.R
PALG_E_646	MHP7448_0570	conserved hypothetical protein	AATTATGAA-N11-TTTAGAATT	757773	757801	+	144	TU 101 (L)	1.R
PALG_E_647	MHP7448_0570	conserved hypothetical protein	ATTTTAATTA-N7-TAATTATAAT	757823	757849	+	96	TU 101 (L)	1.R
PALG_E_649	pulA	pullulanase	GTTTTTAATTA-N8-TAAATAAAAAC	777595	777624	+	248	mC 38	2.R
PALG_E_650	pulA	pullulanase	TTATCGTTA-N14-TAATGATAA	777790	777821	+	51	mC 38	R.2
PALG_E_651	trpS	c tryptophanyl-tRNA synthetase	AAAACTTTT-N15-AAAATTTTT	784342	784374	-	5	TU 105 (F)	R
PALG_E_654	MHP7448_0596	hypothetical protein	AATTTATAGT-N14-ACCAAAAATT	790567	790600	+	102	TU 106 (I)	R
PALG_E_655	MHP7448_0598	hypothetical protein	TTTTTAGTTTT-N10-AAACCTAAAAA	794021	794052	+	129	TU 106 (I)	R
PALG_E_656	MHP7448_0599	hypothetical protein	TAATTATTAT-N13-ATAAGATTTA	795865	795897	+	164	TU 106 (I)	1.R
PALG_E_657	MHP7448_0599	hypothetical protein	ACTAGTTTTT-N6-AAAAATTAGT	796008	796033	+	28	TU 106 (I)	R.1
PALG_E_658	ppa	inorganic pyrophosphatase	TTATTTTTAA-N11-TAAAATATAA	801359	801389	+	0	TU 106 (I)	R
PALG_E_659	xylF	ABC transporter xylose-binding lipoprotein	TTTTTGTAAATAA-N13-TAATATACAAAAA	802027	802065	+	135	TU 106 (I)	R
PALG_E_660	MHP7448_0605	sugar ABC transporter ATP-binding protein	AAGAAAAAAT-N14-ATTATATCTT	803559	803592	+	125	TU 106 (I)	1.R
PALG_E_661	MHP7448_0605	sugar ABC transporter ATP-binding protein	TAATTTAGTTTAATTT-N6-AAATTATATTAAATTA	803615	803652	+	65	TU 106 (I)	1.R
PALG_E_662	MHP7448_0610	conserved hypothetical protein	AATAAATAAA-N7-TTTCTTAATT	808984	809010	+	191	TU 106 (L)	1.R
PALG_E_664	MHP7448_0610	conserved hypothetical protein	AAAAATTATA-N6-TATAATTTTT	809093	809118	+	83	TU 106 (L)	1.R

PALG_E_665	MHP7448_0610	conserved hypothetical protein	AAATTTAAGGAAA-N13-TATGCTTAAATTT	809153	809191	+	10	TU 106 (L)	R.1
PALG_E_666	MHP7448_0612	c hypothetical protein	TAAAATACTAA-N11-TTAGGATTTTA	814731	814763	-	34	TU 107 (I)	R.1
PALG_E_667	MHP7448_0612	c hypothetical protein	ATTTTCGGT-N7-ACCAAAAAT	814831	814855	-	134	TU 107 (I)	2.R
PALG_E_668	ddem	c cytosine specific DNA methyltransferase	GAATTTGATT-N6-AATGAAAATC	816736	816761	-	82	TU 107 (I)	1.R
PALG_E_669	ddem	c cytosine specific DNA methyltransferase	AATTGTAGAC-N13-GTCAAAAATT	816829	816861	-	175	TU 107 (I)	1.R
PALG_E_670	ddem	c cytosine specific DNA methyltransferase	TAAATTTTT-N14-AAAAATTTA	816913	816944	-	259	TU 107 (I)	1.R
PALG_E_673	rplJ	c 50S ribosomal protein L10	TAGCCAGAA-N7-TTTTTGCTA	826680	826704	-	215	TU 107 (F)	2.R.2
PALG_E_674	MHP7448_0620	conserved hypothetical protein	AAAATTTAT-N15-AGAAATTTT	826601	826633	+	206	TU 108 (F)	2.R
PALG_E_675	dam	c DNA adenine methylase	AAAAAATTT-N11-AAATTATTT	833600	833628	-	92	mC 39	2.R.2
PALG_E_676	dam	c DNA adenine methylase	TTTTAAAAA-N12-TTCTTAAAA	833660	833689	-	152	mC 39	2.R.2
PALG_E_677	dam	c DNA adenine methylase	GAATTTTCAT-N9-ATAAAAACTC	833720	833748	-	212	mC 39	2.R.2
PALG_E_678	MHP7448_0623	ABC transporter ATP-binding - Pr1	TTTTAAAAA-N8-TTTTCAAAA	834015	834040	+	121	TU 109 (F)	1.R.2
PALG_E_679	MHP7448_0623	ABC transporter ATP-binding - Pr1	TTTTTATGAA-N11-TTAATATAAA	834123	834153	+	8	TU 109 (F)	R.3
PALG_E_680	MHP7448_0624	ABC transporter ATP-binding protein - Pr2	TCAAATTTA-N5-TGAAATTGA	835805	835827	+	6	TU 109 (L)	R.1
PALG_E_681	MHP7448_0625	pseudogene	AAAAATGAAAA-N5-TTTTTATTTTT	837428	837454	+	120	Class 2	R
PALG_E_682	MHP7448_0626	c conserved hypothetical protein	TTTTATTAA-N5-TTACTAAAA	840571	840593	-	70	mC 40	1.R.1
PALG_E_683	MHP7448_0627	ABC transporter ATP-binding - Pr1-like protein	TTTTAGTAA-N5-TTAATAAAA	840571	840593	+	240	TU 110 (F)	1.R.1
PALG_E_684	MHP7448_0630	c 5'-nucleotidase precursor	AAAATTATTTT-N14-AAACAAATTTT	848277	848312	-	102	TU 111 (F)	2.R
PALG_E_685	MHP7448_0630	c 5'-nucleotidase precursor	TTAATAAAA-N10-TTTAATAAA	848318	848345	-	143	TU 111 (F)	2.R
PALG_E_686	MHP7448_0630	c 5'-nucleotidase precursor	TTTTTAATTT-N13-AAATTAAAAA	848453	848485	-	278	TU 111 (F)	2.R
PALG_E_687	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase	TTTTTCTTC-N14-GAATAAAAA	848248	848279	+	247	TU 112 (F)	3.R.1
PALG_E_688	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase	CCTATATTAT-N5-ATAATTTTGG	848290	848314	+	212	TU 112 (F)	3.R.1
PALG_E_689	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase	TTTATTAAA-N10-TTTTATTAA	848318	848345	+	181	TU 112 (F)	3.R.1
PALG_E_690	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase	AAAAAACACT-N9-AGTATTTTTT	848351	848379	+	147	TU 112 (F)	3.R.1
PALG_E_691	MHP7448_0634	conserved hypothetical protein	AAACTAGGT-N14-AACTACTTT	850600	850631	+	127	TU 112 (L)	R
PALG_E_692	MHP7448_0634	conserved hypothetical protein	TTATTTCAA-N10-TTCAATTAA	850703	850730	+	28	TU 112 (L)	R
PALG_E_694	nusG	c transcription antitermination protein	TCTTTTTAT-N7-ATGAAAAGA	857609	857633	-	157	TU 113 (L)	R

PALG_E_695	rpmG-1	c 50S ribosomal protein L33	AATTATAAGTTA-N11-TAGTTTATAATT	857852	857886	-	3	TU 113 (I)	R.1
PALG_E_696	leuS	c leucyl-tRNA synthetase	TTAATTTTA-N11-TATAATTTA	868486	868514	-	10	TU 113 (F)	R.1
PALG_E_697	leuS	c leucyl-tRNA synthetase	CTTTTTTAAATA-N14-TTTTTAGAAAAG	868528	868565	-	52	TU 113 (F)	1.R.1
PALG_E_699	uvrB	excinuclease ABC subunit B	TAAATTATA-N11-TAAAATTAA	868486	868514	+	162	TU 114 (F)	R.1
PALG_E_700	uvrB	excinuclease ABC subunit B	CTTTTCTAAAAA-N14-TATTTAAAAAAG	868528	868565	+	111	TU 114 (F)	1.R.1
PALG_E_701	MHP7448_0653	c conserved hypothetical protein	AATTTTTAT-N9-ATCAAAATT	874725	874751	-	4	TU 115 (F)	R.2
PALG_E_702	MHP7448_0653	c conserved hypothetical protein	TATTATAAT-N8-ATTATAATA	874752	874777	-	31	TU 115 (F)	R.1
PALG_E_703	MHP7448_0653	c conserved hypothetical protein	TTTTATTTCT-N15-ATAATTAAAA	874782	874816	-	61	TU 115 (F)	1.R
PALG_E_704	prsA	ribose-phosphate pyrophosphokinase	ATAATTAAT-N8-ATTTATTAT	874732	874757	+	58	TU 116 (F)	2.R.1
PALG_E_705	prsA	ribose-phosphate pyrophosphokinase	TATTATAAT-N10-ATTATAATA	874768	874795	+	20	TU 116 (F)	R.2
PALG_E_707	MHP7448_0657	putative ABC transporter ATP-binding protein P115-like protein	TTAAAAAAAAC-N13-GACTTTTTTAA	878568	878602	+	52	TU 116 (I)	1.R
PALG_E_708	MHP7448_0663	c adhesin like-protein P146	AATTATTTTT-N6-AAAAATAATT	894453	894478	-	80	TU 117 (F)	2.R.1
PALG_E_709	MHP7448_0663	c adhesin like-protein P146	TATTGTTTTT-N12-AAAAATATTA	894583	894614	-	210	TU 117 (F)	3.R
PALG_E_710	MHP7448_0664	ABC transporter ATP-binding - Pr1	TAATAAACTT-N13-AAGTTTTTTA	894666	894698	+	233	TU 118 (F)	1.R
PALG_E_711	MHP7448_0664	ABC transporter ATP-binding - Pr1	TTATTATTAA-N7-TTAATGTTAA	894733	894759	+	172	TU 118 (F)	1.R
PALG_E_712	MHP7448_0664	ABC transporter ATP-binding - Pr1	ATTTATAAT-N5-ATTAGAAAT	894857	894879	+	52	TU 118 (F)	R.1
PALG_E_713	MHP7448_0664	ABC transporter ATP-binding - Pr1	AAAATTTTTG-N13-CAAAAATTTT	894884	894916	+	15	TU 118 (F)	R.1
PALG_E_714	MHP7448_0667	hypothetical protein	TAAAGAACA-N10-TTTTCCTTA	903676	903703	+	204	TU 120 (I)	R.1
PALG_E_715	MHP7448_0667	hypothetical protein	TAATGTTTTT-N15-AAAAACATAA	903776	903810	+	97	TU 120 (I)	R.1
PALG_E_724	MHP7448_0682	hypothetical protein	TAATTTCATTT-N9-AATTAAAATTA	110570	110602	+	160	TU 10 (I)	R
PALG_E_725	MHP7448_0682	hypothetical protein	AAAATTATAA-N12-TTTTAATTTT	110603	110630	+	132	TU 10 (I)	R
PALG_E_731	MHP7448_0684	hypothetical protein	AAAAATTAAA-N15-TTTATTTTTT	163110	163134	+	25	TU 18 (I)	R
PALG_E_734	MHP7448_0684	hypothetical protein	AATAAAAAAA-N15-TTTTTTTATT	163011	163041	+	118	TU 18 (I)	R
PALG_E_735	MHP7448_0685	hypothetical protein	ATAAAAATTT-N10-AATATTTTAT	190801	190835	+	41	TU 24 (L)	R
PALG_E_742	MHP7448_0688	c ISMHp1 transposase	TTTCAGACAC-N14-GTGTCTCAAA	267093	267122	-	274	Class 1	1.R
PALG_E_749	MHP7448_0694	c hypothetical protein	CAAAATATT-N14-AAAATTTAG	421138	421174	-	10	TU 57 (F)	R
PALG_E_750	MHP7448_0694	c hypothetical protein	ACATTATTT-N13-AAATAATTT	421175	421206	-	47	TU 57 (F)	1.R

PALG_E_751	MHP7448_0694	c hypothetical protein	TGGATTTCT-N5-AGAAATCTA	421243	421272	-	115	TU 57 (F)	1.R
PALG_E_752	MHP7448_0694	c hypothetical protein	TTTTTTCTTT-N12-AATGAAAAAA	421339	421371	-	211	TU 57 (F)	1.R
PALG_E_754	MHP7448_0695	c hypothetical protein	GAAAATTATA-N8-TATAATCATC	429370	429392	+	95	mC 23	R.3
PALG_E_759	MHP7448_0703	hypothetical protein	AATTCAAAA-N5-TTTGGAATT	569011	569045	+	112	TU 75 (I)	R
PALG_E_760	MHP7448_0703	hypothetical protein	AATTATCAAA-N6-TTTAATAGTT	568851	568879	+	278	TU 75 (I)	R
PALG_E_761	MHP7448_0704	hypothetical protein	AAATTATTA-N10-TAATTGTTT	569662	569684	+	11	TU 75 (I)	R
PALG_E_762	MHP7448_0704	hypothetical protein	AATTTCTAATT-N15-AAGTAGCAATT	569462	569484	+	211	TU 75 (I)	R
PALG_E_763	MHP7448_0704	hypothetical protein	TTGATCTTTA-N15-TCAAGATAAA	569526	569552	+	143	TU 75 (I)	R
PALG_E_764	MHP7448_0706	hypothetical protein	GCTCAAGTT-N9-AATTTGATC	570765	570790	+	36	TU 75 (I)	R
PALG_E_765	MHP7448_0706	hypothetical protein	AACAATTGG-N11-CCCCTTGTT	570635	570662	+	164	TU 75 (I)	R
PALG_E_766	MHP7448_0707	hypothetical protein	AGGGAATTT-N13-ATATTCCTT	571321	571355	+	53	TU 75 (I)	R
PALG_E_768	MHP7448_0707	hypothetical protein	GCAACTTTA-N5-TAAAGTTGC	571282	571318	+	90	TU 75 (I)	R
PALG_E_769	MHP7448_0708	hypothetical protein	TAATTAGTT-N10-AATTAATTA	571856	571886	+	118	TU 75 (I)	R
PALG_E_770	MHP7448_0708	hypothetical protein	ATCTCAGGCAA-N11-TTACCTGAGAT	571927	571955	+	49	TU 75 (I)	R
PALG_E_771	MHP7448_0709	hypothetical protein	GTTGTCCAA-N5-TTGGTCAGC	573001	573027	+	49	TU 75 (L)	R.1
PALG_E_772	MHP7448_0709	hypothetical protein	GTACTAAAT-N5-ATTTTGTTC	572834	572856	+	220	TU 75 (L)	1.R
PALG_E_773	MHP7448_0709	hypothetical protein	AATTCCTGAA-N11-TTCAGCAAGT	572857	572891	+	185	TU 75 (L)	R
PALG_E_774	MHP7448_0709	hypothetical protein	GAATCTCAAG-N8-CTTGATATTC	572939	572966	+	110	TU 75 (L)	R.1
PALG_E_775	MHP7448_0709	hypothetical protein	AGAATTAAT-N15-ATATATTCT	572968	573000	+	76	TU 75 (L)	R.1
PALG_E_776	MHP7448_0712	hypothetical protein	ATCAAAATT-N9-AATCTTAAT	584388	584420	+	29	TU 76 (I)	R
PALG_E_777	MHP7448_0712	hypothetical protein	GCAGCATCA-N8-TGATGCTTC	584179	584202	+	247	TU 76 (I)	R
PALG_E_778	MHP7448_0712	hypothetical protein	ATATTTTTAA-N14-TTAAAAAAAT	584225	584251	+	198	TU 76 (I)	R
PALG_E_779	MHP7448_0712	hypothetical protein	AAAACTAATTT-N9-AATTTTGTTTT	584265	584287	+	162	TU 76 (I)	R
PALG_E_780	MHP7448_0712	hypothetical protein	TTCTTGATTT-N12-AAATTAAAAA	584291	584321	+	128	TU 76 (I)	R
PALG_E_781	MHP7448_0714	c hypothetical protein	TTAACAAAA-N13-TTTTGAAAA	724106	724139	+	132	TU 95 (F)	1.R.1
PALG_E_782	MHP7448_0714	c hypothetical protein	TTTGACTTT-N10-AAAATCAAA	724171	724201	+	70	TU 95 (F)	R.1
PALG_E_783	MHP7448_0714	c hypothetical protein	TGCTGCAGTAGC-N10-GCTACTGTAGCA	724206	724237	+	34	TU 95 (F)	R.2

PALG_E_790	MHP7448_0718	hypothetical protein	TTTTAAGTTT-N8-AAAATTAAAA	749506	749533	+	58	TU 98 (L)	R
PALG_E_791	MHP7448_0718	hypothetical protein	TTTCTAATTT-N13-ACAATAGAAA	749544	749571	+	20	TU 98 (L)	R
PALG_E_793	MHP7448_0719	hypothetical protein	AAAATTCAG-N10-CTGATATTT	752525	752552	+	198	mC 35	1.R.1
PALG_E_794	MHP7448_0719	hypothetical protein	TTTTAAGTTT-N8-AAAATTAAAA	752581	752613	+	137	mC 35	R.1
PALG_E_797	MHP7448_0721	c conserved hypothetical protein	AAAAGGTTGTC-N5-GACGACCTTTT	765208	765233	-	33	TU 102 (F)	R
PALG_E_800	MHP7448_0722	c conserved hypothetical protein	TTATCATTA-N14-TAACGATAA	777579	777610	-	12	TU 104 (I)	R.3
PALG_E_801	MHP7448_0722	c conserved hypothetical protein	ATATTAATAT-N11-ATGTTAAAAT	777626	777652	-	59	TU 104 (I)	1.R.3
PALG_E_805	dnaN	c 50S ribosomal protein L2	ATTTTTATT-N10-AATAAAAAT	1680	1707	+	35	TU 01 (I)	1.R
PALG_E_808	MHP7448_0009	conserved hypothetical protein	TTTTTTCTAATTTTTT-N11-AAAAAAATTAGAAAAAA	9078	9120	+	74	TU 01 (I)	1.R
PALG_E_810	fba	fructose-bisphosphate aldolase	TTTTTTTTTTTTTTT-N15-AAAAAAAAAAAAAAA	14897	14941	+	23	TU 01 (I)	R
PALG_E_811	MHP7448_0016	c hypothetical protein	TTTTTTTTTT-N9-AAAAAATAAA	17030	17058	-	46	mC 01	R
PALG_E_812	MHP7448_0017	hypothetical protein	TATAGTCAAG-N6-CTTGATTTTA	17313	17338	+	21	mC 02	R
PALG_E_814	MHP7448_0018	c conserved hypothetical protein	AATAATTTAA-N15-TAAAATTATT	18796	18830	-	90	TU 02 (L)	R
PALG_E_815	MHP7448_0019	c ABC transporter ATP-binding protein	TTAAAACAACTA-N14-TAATTGTATTAA	21152	21189	-	15	TU 02 (I)	R
PALG_E_816	MHP7448_0019	c ABC transporter ATP-binding protein	TTAAGGTTTTT-N14-AAAAACCCAAA	21264	21299	-	127	TU 02 (I)	R
PALG_E_817	MHP7448_0022	c hypothetical protein	TAGATAAAA-N15-TTTTTTCTA	26902	26934	-	115	TU 02 (I)	2.R
PALG_E_818	MHP7448_0025	c hypothetical protein	TTAAAATTA-N12-TTATTTTAA	32560	32589	-	90	TU 02 (F)	R.3
PALG_E_819	MHP7448_0025	c hypothetical protein	AACAATTTTAATT-N11-AATAAAAAGTGTT	32767	32803	-	297	TU 02 (F)	3.R
PALG_E_820	sipS	signal peptidase I	TTAAAATAA-N12-TAATTTTAA	32560	32589	+	253	TU 03 (F)	2.R.1
PALG_E_821	MHP7448_0030	hypothetical protein	TATTAAAATT-N13-AATTTTAGTA	36578	36610	+	56	TU 03 (L)	1.R
PALG_E_822	parE	c topoisomerase IV subunit B	AATTTTACA-N10-TGTAAAATT	45050	45077	-	70	TU 04 (F)	1.R
PALG_E_823	parE	c topoisomerase IV subunit B	AAAAAATTAA-N6-TTAATTATTT	45102	45127	-	122	TU 04 (F)	2.R
PALG_E_824	gap	glyceraldehyde 3-phosphate dehydrogenase	GTAAAAATTTT-N8-AAATTTTATAC	45008	45037	+	152	mC 03	2.R
PALG_E_825	gap	glyceraldehyde 3-phosphate dehydrogenase	AAATAATTAA-N6-TTAATTTTTT	45102	45127	+	62	mC 03	R.2
PALG_E_826	recA	c recombination protein	TGAAAAAAT-N5-ATTTTTTCA	53480	53502	-	120	TU 05 (F)	R.2
PALG_E_827	licA	PTS system, lichenan-specific IIA component	TGAAAAAAT-N5-ATTTTTTCA	53480	53502	+	9	TU 06 (F)	R
PALG_E_828	atpF	ATP synthase subunit B	AAATTTTTAT-N14-AAAAAAAGTT	60393	60426	+	3	TU 06 (I)	R

PALG_E_829	MHP7448_0054	conserved hypothetical protein	AATTACACC-N7-GGTGTGATT	65490	65514	+	233	TU 06 (I)	R
PALG_E_830	MHP7448_0054	conserved hypothetical protein	TTTTTTTGC-N8-GAAAAAAAA	65707	65732	+	15	TU 06 (I)	R
PALG_E_831	rpsB	30S ribosomal protein S2	AAAAATTAAA-N5-TTTAATTTTT	66291	66315	+	133	TU 06 (I)	R
PALG_E_832	rpsB	30S ribosomal protein S2	ATAATTTAATT-N15-AATAAAAATAT	66327	66363	+	85	TU 06 (I)	R
PALG_E_833	uvrC	c excinuclease ABC subunit C	TTTTTGTTAA-N11-TTATCAAAAA	87259	87289	-	66	mC 06	1.R
PALG_E_835	dnaK	chaperone protein dnaK - heat shock protein 70	TTCTTTTGTCT-N9-AGACAAAATAA	87768	87798	+	40	TU 07 (I)	R.1
PALG_E_837	cmk	cytidylate kinase	TGTTTATAA-N13-TTATTAACA	90993	91023	+	17	TU 07 (I)	R
PALG_E_838	MHP7448_0074	hypothetical protein	AAGCAAAAA-N6-TTTTTACTT	95434	95457	+	16	TU 07 (L)	R.1
PALG_E_839	rpsL	c 30S ribosomal protein S12	TTGTTTTCTA-N13-TATAAAATAA	99786	99818	-	8	TU 08 (I)	R
PALG_E_840	MHP7448_0080	c conserved hypothetical protein	TTTTTTACT-N13-ACTAAAAAA	101326	101356	-	30	TU 08 (F)	R.1
PALG_E_841	MHP7448_0081	amino acid permease	AAAAAATTTA-N12-TAAAATTTTT	101352	101383	+	176	TU 09 (F)	3.R
PALG_E_842	MHP7448_0081	amino acid permease	ATAATTTTAA-N12-TTAAAATCTT	101509	101540	+	19	TU 09 (F)	R.2
PALG_E_843	deoD	c purine-nucleoside phosphorylase	TATTATAAT-N11-ATTAACATA	106682	106710	-	10	TU 10 (I)	R
PALG_E_845	secA	c preprotein translocase SecA subunit	ATAATTTTTTAT-N5-ATAAAAAATTAT	110460	110488	-	221	TU 10 (F)	1.R
PALG_E_846	MHP7448_0089	c hypothetical protein	TTTGAGACAC-N12-GTGTCTGAAA	114873	114904	-	66	TU 11 (F)	1.R
PALG_E_847	MHP7448_0090	conserved hypothetical protein	AATTAAGTTA-N5-TAAACTAATT	115353	115377	+	254	TU 12 (L)	1.R
PALG_E_848	MHP7448_0090	conserved hypothetical protein	AAATTTATTA-N10-TAATAATTCT	115499	115528	+	103	TU 12 (L)	R.1
PALG_E_849	MHP7448_0093	c hypothetical protein	TTTTATTTTTA-N5-TAAAAATAAAA	122664	122690	-	45	TU 13 (F)	1.R
PALG_E_850	MHP7448_0095	c conserved hypothetical protein	AAAATCTTGA-N14-TCAAGATTTT	124913	124946	-	19	mC 09	1.R.1
PALG_E_851	trxB	thioredoxin reductase	GAAAAATTT-N11-AATTTTTTC	127307	127335	+	71	TU 14 (I)	3.R
PALG_E_852	trxB	thioredoxin reductase	ATAAAAATATTATAAT-N13-ATTATAATATTTTTAT	127347	127391	+	15	TU 14 (I)	R
PALG_E_853	clpB	c ATP-dependent protease binding protein	TAATTTATT-N14-AAAAAATTA	136827	136858	-	12	mC 10	R.1
PALG_E_854	clpB	c ATP-dependent protease binding protein	TAATTTAAA-N9-TTTAAATTA	136936	136962	-	121	mC 10	1.R
PALG_E_856	MHP7448_0106	c hypothetical protein	ATTATTAAA-N11-TTTTATAAT	142464	142492	-	5	TU 16 (L)	R
PALG_E_858	MHP7448_0106	c hypothetical protein	ATTTTATTAT-N9-AGAATAAAAT	142581	142609	-	122	TU 16 (L)	R
PALG_E_859	MHP7448_0108	c protein P97 - copy 2	ATAATAAAA-N9-TTTTATTAT	148777	148803	-	70	TU 16 (I)	R
PALG_E_860	MHP7448_0108	c protein P97 - copy 2	AAAAAAGTGCC-N8-GGCATTTTTTT	148906	148935	-	199	TU 16 (I)	R

PALG_E_861	gyrB	c DNA gyrase subunit B	TTAAAAAAGCG-N12-CACTTTTTTAA	151119	151152	-	232	TU 16 (I)	1.R
PALG_E_863	pfkA	c 6-phosphofructokinase	TCCTATTTTA-N10-TAAAAAAGGA	154292	154321	-	17	TU 16 (L)	1.R
PALG_E_864	MHP7448_0112	hypothetical protein	AAAAAAATTAT-N7-ATTATTTTTTT	154588	154616	+	102	TU 16 (F)	2.R.1
PALG_E_865	MHP7448_0113	amino acid permease	TAATTTAAA-N7-TTTAAAATA	155960	155984	+	21	TU 17 (I)	1.R
PALG_E_866	MHP7448_0117	conserved hypothetical protein	TATTTAACAAA-N15-TTTGTTAAATA	161129	161165	+	129	TU 18 (F)	1.R
PALG_E_867	MHP7448_0117	conserved hypothetical protein	TTAATTTTTA-N15-TAAAAATAAA	161186	161220	+	74	TU 18 (F)	1.R
PALG_E_868	MHP7448_0120	conserved hypothetical protein	TTAATTTTTA-N15-TAAAAATAAA	163554	163588	+	73	TU 18 (I)	1.R
PALG_E_869	ung	c uracil-dna glycosylase	TTATAGCATTTT-N6-AAAAAGCAATAA	171615	171644	-	31	TU 20 (F)	R
PALG_E_870	ung	c uracil-dna glycosylase	ATTTAAACC-N15-GGTTTGAAT	171656	171688	-	72	TU 20 (F)	R
PALG_E_872	rplU	50S ribosomal protein L21	ATTCAAACC-N15-GGTTTAAAT	171656	171688	+	45	TU 21 (F)	R.1
PALG_E_873	lip2	c lipase-esterase	TTCTAATTTTTTTTG-N14-CTAAAAAAATTAGAA	173430	173473	-	51	mC 12	1.R
PALG_E_876	MHP7448_0136	hexosephosphate transport protein	AAAAATCATTT-N12-AAATGATTTTT	174825	174858	+	87	TU 22 (I)	1.R
PALG_E_877	ldh	L-lactate dehydrogenase	TTAAATCTAT-N6-AAAAATTTAA	176479	176504	+	22	TU 22 (L)	R
PALG_E_880	MHP7448_0145	chromate transport protein	AATTTAACT-N9-ATTTAAATT	185882	185908	+	92	TU 24 (F)	R
PALG_E_881	MHP7448_0145	chromate transport protein	AAAAATTAAA-N6-TTTATTTTTT	185922	185947	+	53	TU 24 (F)	R
PALG_E_882	tig	trigger factor	TATTATAAT-N8-ATTATACTA	189216	189241	+	162	TU 24 (I)	R
PALG_E_883	tig	trigger factor	TAATTTGAATTAA-N6-TTAATTTAGATTA	189291	189322	+	81	TU 24 (I)	R
PALG_E_884	tig	trigger factor	TTTTTATTTTT-N14-AAAATTAACAA	189335	189370	+	33	TU 24 (I)	R
PALG_E_885	MHP7448_0155	c conserved hypothetical protein	TAATTTTTA-N5-TACAAATAA	194921	194943	-	215	TU 25 (I)	R
PALG_E_888	MHP7448_0162	hypothetical protein	TTATTTAATTT-N5-AAATTAACTAA	200786	200812	+	126	TU 26 (I)	R
PALG_E_889	rpsM	c 30S ribosomal protein S13	TCATTTTTT-N11-AAAAAAGGA	207419	207447	-	24	TU 27(I)	R
PALG_E_892	rpsJ	c 30S ribosomal protein S10	CAGTACGCGA-N13-TCGTGTATTG	220607	220639	-	6	TU 27 (F)	R.2
PALG_E_893	rpsJ	c 30S ribosomal protein S10	AAAAAATTA-N8-TTATTTTTT	220687	220712	-	86	TU 27 (F)	R.1
PALG_E_894	MHP7448_0197	conserved hypothetical protein	AAAATCCAAAACTTG-N11-CAAGTTTTGGATTTT	221036	221076	+	52	TU 28 (F)	1.R
PALG_E_895	MHP7448_0198	protein P97 - copy 1	TTTTTTTATA-N11-TATAAAAGAA	221682	221712	+	256	TU 28 (I)	1.R
PALG_E_896	MHP7448_0198	protein P97 - copy 1	ACTTTTTTGTT-N10-AAAAAAAAAGT	221787	221818	+	150	TU 28 (I)	1.R
PALG_E_897	MHP7448_0210	c ABC transporter ATP-binding protein	AAAATTTAAA-N9-TTTCAATTTT	241632	241660	-	47	TU 30 (F)	R

PALG_E_898	MHP7448_0210	c ABC transporter ATP-binding protein	AATTAAAAA-N10-TTTTAAACT	241721	241748	-	136	TU 30 (F)	1.R
PALG_E_9	MHP7448-tRNA-Trp1	tRNA-Trp	AAAATCTTGA-N14-TCAAGATTTT	124913	124946	+	103	Class 3	2.R.1
PALG_E_900	oppB-1	oligopeptide ABC transporter system permease	AATATGATTTT-N7-AAAATCAATTT	243042	243070	+	199	TU 31 (I)	R
PALG_E_901	MHP7448_0218	ISMHp1 transposase	TACAATTAATT-N9-AAGTAATTTTA	257075	257105	+	17	Class 1	1.R
PALG_E_902	nrdF	c ribonucleoside-diphosphate reductase beta chain	GTATAAAAAATA-N8-TATTTTTTTAAC	264637	264668	-	283	TU 32 (I)	1.R
PALG_E_903	glyA	c glycine hydroxymethyltransferase	TTTTTAAAA-N14-TTTTTAAAA	268603	268634	-	161	TU 32 (F)	2.R
PALG_E_904	glyA	c glycine hydroxymethyltransferase	CAAAAATTTT-N12-ACAATTTTTG	268646	268677	-	204	TU 32 (F)	2.R
PALG_E_905	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	CAAAAATTGT-N12-AAAATTTTTG	268646	268677	+	261	TU 33 (F)	1.R
PALG_E_906	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	AAAAACATA-N7-TATATTTTT	268901	268925	+	13	TU 33 (F)	R.1
PALG_E_907	MHP7448_0240	c hypothetical protein	AACTAATAT-N9-ATATTATTT	287210	287236	-	19	TU 35 (F)	R.2
PALG_E_908	MHP7448_0240	c hypothetical protein	AATATGTTTTT-N13-AAAAATATATT	287369	287403	-	178	TU 35 (F)	2.R.1
PALG_E_909	MHP7448_0240	c hypothetical protein	AAAAAATAAT-N12-AATATTTTTT	287421	287452	-	230	TU 35 (F)	2.R
PALG_E_910	MHP7448_0240	c hypothetical protein	TTAAAAAATT-N8-ATTTCTTTAA	287479	287506	-	288	TU 35 (F)	3.R
PALG_E_911	secD	protein-export membrane protein	TTAAAGAAAT-N8-AATTTTTTAA	287479	287506	+	211	TU 36 (F)	1.R
PALG_E_912	secD	protein-export membrane protein	AAAATTTATAAG-N5-CTTGAAAATTTT	287567	287595	+	122	TU 36 (F)	1.R
PALG_E_913	rpsT	30S ribosomal protein S20	TTTTATAAA-N12-TTTAAAAAA	293927	293956	+	101	TU 36 (L)	1.R
PALG_E_914	smf	c DNA processing protein SMF	TAATTTTAA-N7-TTAAACTTA	304256	304280	-	121	TU 37 (F)	R
PALG_E_915	eno	enolase	TAAGTTTAA-N7-TTAAAATTA	304256	304280	+	147	TU 38 (F)	1.R
PALG_E_916	MHP7448_0255	c hypothetical protein	AATAAATTT-N11-AAATTAATT	309930	309958	-	69	TU 39 (F)	3.R
PALG_E_917	dnaX	DNA polymerase III gamma and tau subunit	GTTTTTATTT-N13-AATTATAAAC	309889	309921	+	59	TU 40 (F)	1.R
PALG_E_918	MHP7448_0257	conserved hypothetical protein	TTTATATAATTT-N15-AAATTATCGAAA	312169	312207	+	9	TU 40 (I)	R
PALG_E_919	MHP7448_0267	conserved hypothetical protein	TAAAAAAATA-N13-TGTTTTTTTA	320642	320674	+	34	TU 40 (L)	R
PALG_E_920	mgtA	c cation-transporting P-type ATPase	TAATTTATA-N6-TATAAAAAA	324285	324308	-	10	mC 17	R.1
PALG_E_921	smpB	SsrA-binding protein	TATAAAATT-N5-AATTAGATA	324290	324312	+	51	TU 41 (F)	1.R
PALG_E_922	smpB	SsrA-binding protein	TAAATTACTATTA-N11-TAATATTATTTTA	324318	324354	+	9	TU 41 (F)	R
PALG_E_923	MHP7448_0270	nuclease, lipoprotein	ATATACAACTT-N6-AAGATGTCTAT	324953	324980	+	108	TU 41 (L)	R
PALG_E_924	MHP7448_0271	c P102-like protein	TTAATTTTTA-N6-TAAAAATACA	328608	328633	-	17	TU 42 (L)	R

PALG_E_925	MHP7448_0272	c P97-like protein	AATTAAAAT-N8-ATTTTAATT	331770	331795	-	30	TU 42 (F)	R.1
PALG_E_926	MHP7448_0272	c P97-like protein	TAAAATCACAA-N5-TTCGGATTTTA	331806	331832	-	66	TU 42 (F)	1.R.1
PALG_E_927	MHP7448_0279	c transcriptional regulator	TTTTTTTGA-N7-TCAAAAAAA	339513	339537	-	245	mC 18	2.R.1
PALG_E_928	pyrG	CTP synthase	TTTTTTTGA-N7-TCAAAAAAA	339513	339537	+	248	mC 19	3.R
PALG_E_929	pyrG	CTP synthase	TTAAATTTA-N11-TTAATTTAA	339749	339777	+	8	mC 19	R.2
PALG_E_930	rpsP	c 30S ribosomal protein S16	AACTTAAAA-N8-TTTTAAGTT	346043	346068	-	89	TU 44 (I)	1.R
PALG_E_931	rpsP	c 30S ribosomal protein S16	TTTTTTATG-N6-CAAAAAAAA	346070	346093	-	116	TU 44 (I)	1.R
PALG_E_932	MHP7448_0287	c hypothetical protein	AATTTTTTCT-N8-ACAAAAAATT	346727	346754	-	68	TU 44 (F)	R.1
PALG_E_933	MHP7448_0287	c hypothetical protein	AAAAAAAAT-N10-ATTTTTTTT	346956	346983	-	297	TU 44 (F)	1.R
PALG_E_934	MHP7448_0289	hypothetical protein	AAAAATTCA-N9-TCAATTTTT	347466	347492	+	236	TU 45 (F)	R
PALG_E_935	MHP7448_0290	hypothetical protein	CAAAAGAAA-N11-TTTATTTTG	349016	349044	+	21	TU 45 (I)	R
PALG_E_936	MHP7448_0291	putative type II DNA modification enzyme	AATTTTATTT-N9-AAAAAAACTT	350744	350772	+	20	TU 45 (F)	R
PALG_E_937	MHP7448_0298	c hypothetical protein	TATAATAATT-N13-AATTCTTATA	360284	360316	-	58	TU 46 (F)	R.4
PALG_E_938	MHP7448_0299	hypothetical protein	AATTATTATA-N5-TATAAGAATT	360307	360331	+	13	TU 47 (F)	R.1
PALG_E_939	MHP7448_0300	hypothetical protein	TTTTTATTTTTTT-N15-AAAAAAATAAAAA	361823	361863	+	269	TU 47 (I)	2.R
PALG_E_940	MHP7448_0300	hypothetical protein	AACATTTTCT-N9-AGACAATATT	361932	361960	+	172	TU 47 (I)	2.R
PALG_E_941	MHP7448_0301	hypothetical protein	AAATTTATT-N12-AACAAATTT	364056	364085	+	151	TU 47 (I)	R
PALG_E_942	MHP7448_0308	conserved hypothetical protein	AATTTTTTCTA-N11-TTAAAAAAATT	372712	372744	+	292	TU 49 (L)	3.R
PALG_E_943	MHP7448_0308	conserved hypothetical protein	TCATATAAAA-N8-TTTTATATGA	372761	372788	+	248	TU 49 (L)	3.R
PALG_E_944	MHP7448_0308	conserved hypothetical protein	TTTTGATAA-N12-TTTTCAAAA	372841	372870	+	166	TU 49 (L)	3.R
PALG_E_945	MHP7448_0315	ABC transporter ATP-binding protein	TTATAATCTT-N7-AAGAATACAA	386273	386299	+	15	TU 51 (I)	R
PALG_E_946	MHP7448_0321	conserved hypothetical protein	AAATACTTATTTA-N13-TATATAATTATTT	395915	395953	+	0	TU 51 (L)	R
PALG_E_947	MHP7448_0322	c putative transposase	TAGATAAAA-N15-TTTTTTCTA	396512	396544	-	84	Class 1	R.1
PALG_E_948	MHP7448_0322	c putative transposase	CTTGTTTCT-N8-AAAAACAAG	396697	396722	-	269	Class 1	1.R
PALG_E_949	MHP7448_0323	transposase	AATAAAAAT-N10-ATTTTTTTT	396834	396861	+	183	Class 1	1.R
PALG_E_950	MHP7448_0323	transposase	TTACAATTTTA-N14-TATAATTATAA	397006	397041	+	3	Class 1	R
PALG_E_951	MHP7448_0327	c conserved hypothetical protein	TTTTGATAA-N12-TTTTCAAAA	404009	404038	-	205	TU 53 (F)	2.R

PALG_E_952	MHP7448_0327	c conserved hypothetical protein	TCATATAAAA-N8-TTTTATATGA	404091	404118	-	287	TU 53 (F)	2.R
PALG_E_953	MHP7448_0328	hypothetical protein	TTTTGAAAA-N12-TTATCAAAA	404009	404038	+	273	TU 54 (F)	2.R
PALG_E_954	MHP7448_0328	hypothetical protein	TCATATAAAA-N8-TTTTATATGA	404091	404118	+	193	TU 54 (F)	2.R.1
PALG_E_955	MHP7448_0328	hypothetical protein	ATTTTTTAAAA-N8-TTAGAAAAAAT	404137	404166	+	145	TU 54 (F)	2.R.1
PALG_E_956	MHP7448_0331	c hypothetical protein	TTATTGTTAA-N8-TTAACAAAGA	409547	409574	-	49	TU 55 (L)	R
PALG_E_957	MHP7448_0332	c subtilisin-like serine protease	AAATTAAAA-N12-TTATAATTT	411766	411795	-	29	TU 55 (I)	R.1
PALG_E_958	MHP7448_0333	c lipoprotein	AAATCCAAAA-N10-TTTTAGAGTT	412848	412877	-	136	TU 55 (F)	1.R.2
PALG_E_959	MHP7448_0333	c lipoprotein	GATTTTATA-N14-TATAAATTC	412881	412912	-	169	TU 55 (F)	1.R.1
PALG_E_960	MHP7448_0334	hypothetical protein	GAATTTATA-N14-TATAAAATC	412881	412912	+	212	TU 56 (F)	R.1
PALG_E_961	MHP7448_0336	c hypothetical protein	TCAAAAAATT-N10-AATTTTTTGA	416066	416095	-	151	mC 21	1.R
PALG_E_962	MHP7448_0336	c hypothetical protein	ATAATTAAT-N6-ATTACTTAT	416135	416158	-	220	mC 21	1.R
PALG_E_963	MHP7448_0337	hypothetical protein	ATTTTTTATG-N7-CAAAAAAAAT	416187	416213	+	206	mC 22	R.1
PALG_E_964	MHP7448_0337	hypothetical protein	CTGGAAATT-N11-AATTCCAAG	416340	416368	+	51	mC 22	R.1
PALG_E_965	MHP7448_0339	c hypothetical protein	CCAAAAATT-N12-AAATTTTGG	420300	420329	-	211	TU 57 (I)	R
PALG_E_966	MHP7448_0342	c hypothetical protein	ACCCCTTTAA-N7-TCAAAGGAGT	426882	426908	-	283	TU 57 (I)	1.R
PALG_E_967	MHP7448_0343	pseudogene	TTTATACTA-N6-TAGTTTAAA	427303	427326	+	141	Class 2	1.R
PALG_E_968	MHP7448_0348	c hypothetical protein	ATATGAAAAAAC-N15-GTTTTTTCATAT	434320	434358	-	134	TU 59 (L)	R
PALG_E_969	MHP7448_0349	c hypothetical protein	AAATTAAGAA-N5-TTTTTAATTT	436060	436084	-	66	TU 59 (I)	R
PALG_E_970	MHP7448_0349	c hypothetical protein	GTATTTTTA-N5-TAAAAAAAC	436095	436117	-	101	TU 59 (I)	R
PALG_E_971	MHP7448_0352	hypothetical protein	AAAACTATAT-N6-ATATAATTTT	438630	438655	+	28	TU 60 (F)	R.1
PALG_E_972	MHP7448_0355	c hypothetical protein	ATTTAAAAA-N14-TTTTTTAAT	444505	444536	-	18	TU 61 (L)	1.R
PALG_E_973	MHP7448_0355	c hypothetical protein	AAATAAATTTA-N14-TAATTTTAGTT	444563	444598	-	76	TU 61 (L)	1.R
PALG_E_975	MHP7448_0360	P37-like ABC transporter substrate-binding lipoprotein	TTTGAATTA-N8-TAATCCAAA	450563	450588	+	284	TU 63 (F)	2.R
PALG_E_977	MHP7448_0365	hypothetical protein	AAAAAAACCTT-N7-AATGATTTTTT	456553	456581	+	169	TU 63 (I)	R
PALG_E_978	MHP7448_0365	hypothetical protein	TTAAAAATA-N13-TATTTTTAA	456629	456659	+	91	TU 63 (I)	R
PALG_E_979	MHP7448_0365	hypothetical protein	AAATCTTTTT-N10-AAAAATAATT	456715	456744	+	6	TU 63 (I)	R
PALG_E_980	MHP7448_0366	lipoprotein	ATAGATTTTTTT-N14-AAAAAAATGTAT	457389	457426	+	211	TU 63 (I)	R

PALG_E_981	MHP7448_0366	lipoprotein	TAAAAATTTT-N7-AAAAATTGTA	457536	457562	+	75	TU 63 (I)	R
PALG_E_982	MHP7448_0366	lipoprotein	TTTTAGAAA-N10-TTTTTAAAA	457607	457634	+	3	TU 63 (I)	R
PALG_E_983	MHP7448_0375	c PTS system enzyme IIB component	CTCTTTTTTA-N12-TAAAAATAAG	474294	474325	-	82	TU 64 (I)	R
PALG_E_984	MHP7448_0375	c PTS system enzyme IIB component	ATTTTTAAA-N5-TTAAAAAAT	474417	474439	-	205	TU 64 (I)	R
PALG_E_985	MHP7448_0375	c PTS system enzyme IIB component	AAAGTTTAAA-N7-TTTAAAAGTT	474462	474488	-	250	TU 64 (I)	R
PALG_E_987	MHP7448_0378	c lipoprotein	TTATTTTTT-N15-AAAAATTAA	479576	479608	-	15	TU 64 (F)	R.1
PALG_E_988	MHP7448_0378	c lipoprotein	AAAAATTTT-N12-AAAATTTTT	479610	479639	-	49	TU 64 (F)	R.1
PALG_E_989	MHP7448_0378	c lipoprotein	AAAATTTTAC-N9-GCAAAATTTT	479671	479699	-	110	TU 64 (F)	1.R
PALG_E_990	MHP7448_0379	ABC transport ATP-binding protein	AAAAATTTT-N12-AAAATTTTT	479610	479639	+	289	TU 65 (F)	1.R
PALG_E_991	MHP7448_0379	ABC transport ATP-binding protein	AAAATTTTGC-N9-GTAAAATTTT	479671	479699	+	229	TU 65 (F)	R.1
PALG_E_992	MHP7448_0384	c thioredoxin	TAAATTTTTT-N11-AAAAAATGTA	486053	486083	-	26	TU 66 (L)	1.R
PALG_E_993	MHP7448_0386	pseudogene	ATAATTTTT-N14-AAAAATTAT	487967	487998	+	35	Class 2	R.1
PALG_E_994	MHP7448_0391	conserved hypothetical protein	AAATTAACCA-N6-TATTTAATTT	492828	492853	+	182	TU 67 (F)	1.R
PALG_E_995	MHP7448_0391	conserved hypothetical protein	AAATAAAAAAAA-N11-TTTTTTTTCTTT	492961	492995	+	40	TU 67 (F)	1.R
PALG_E_996	MHP7448_0394	conserved hypothetical protein	ATTTTTTTTA-N11-TAAAAAATAT	495516	495546	+	64	TU 67 (I)	1.R
PALG_E_998	MHP7448_0397	c conserved hypothetical protein	TTTGAAAAAAAA-N14-TATTTTTTCAAA	499818	499855	-	27	TU 68 (L)	1.R
PALG_E_999	rnc	c ribonuclease III	AATTTTAAAAT-N15-ATTTTAAAATT	500698	500734	-	23	TU 68 (I)	1.R
PALG_ES_100	MHP7448_0306	c ABC transporter ATP-binding protein	CCTGATGTT-N7-AAGATGAGG	371785	371809	-	102	TU 48 (I)	R
PALG_ES_101	MHP7448_0312	c glycine cleavage system H protein	TTGAAATTA-N6-TAATTTATA	382771	382794	-	111	TU 50 (F)	R.1
PALG_ES_102	MHP7448_0312	c glycine cleavage system H protein	AATTTTTTA-N8-TAAAAAAGT	382868	382893	-	208	TU 50 (F)	R.1
PALG_ES_103	MHP7448_0312	c glycine cleavage system H protein	TCCATAATT-N12-AATTATTTA	382930	382959	-	270	TU 50 (F)	1.R
PALG_ES_104	MHP7448_0313	hypothetical protein	TATAAATTA-N6-TAATTTCAA	382771	382794	+	272	TU 51 (F)	2.R
PALG_ES_105	MHP7448_0313	hypothetical protein	ACTTTTTTA-N8-TAAAAAATT	382868	382893	+	173	TU 51 (F)	2.R
PALG_ES_106	MHP7448_0318	hypothetical protein	AATATTTTTT-N14-AAAAAATAGT	393093	393126	+	82	TU 51 (I)	1.R
PALG_ES_107	MHP7448_0318	hypothetical protein	AATTTTTTAA-N13-TAAAACAATT	393163	393195	+	13	TU 51 (I)	R.1
PALG_ES_109	MHP7448_0320	hypothetical protein	ACATACATCA-N8-TGATTTATTT	395599	395626	+	8	TU 51 (I)	R
PALG_ES_110	MHP7448_0322	c putative transposase	TTTCATATAA-N11-TTATATGATA	396614	396644	-	186	Class 1	1.R

PALG_ES_111	MHP7448_0323	transposase	GCAGAAATC-N13-GTTTTTTGC	396941	396971	+	73	Class 1	1.R
PALG_ES_112	MHP7448_0327	c conserved hypothetical protein	AAAATTTTA-N5-TACAATTCT	403884	403906	-	80	TU 53 (F)	1.R
PALG_ES_113	MHP7448_0327	c conserved hypothetical protein	ACGAAGTTG-N8-CAATTACGT	403950	403975	-	146	TU 53 (F)	2.R
PALG_ES_114	MHP7448_0330	hypothetical protein	TTTTGATTTT-N15-AAAGACAAAA	406154	406188	+	65	TU 54 (L)	R
PALG_ES_115	MHP7448_0331	c hypothetical protein	ATTAACAAT-N13-ATTCTTACT	409509	409539	-	11	TU 55 (L)	R
PALG_ES_116	MHP7448_0336	c hypothetical protein	TATTTAAAA-N13-TTTTAAACA	415940	415970	-	25	mC 21	1.R
PALG_ES_117	MHP7448_0336	c hypothetical protein	AAAATTTAT-N7-ATGAACTTT	416100	416124	-	185	mC 21	1.R
PALG_ES_118	MHP7448_0337	hypothetical protein	TGTTTAAAG-N12-CTCTAAATA	416227	416256	+	163	mC 22	R.1
PALG_ES_119	MHP7448_0337	hypothetical protein	TTTAAAACA-N5-TGACTTAAA	416262	416284	+	135	mC 22	R.1
PALG_ES_120	MHP7448_0337	hypothetical protein	ATTAGATGA-N6-TAATGTAAT	416288	416311	+	108	mC 22	R.1
PALG_ES_121	MHP7448_0339	c hypothetical protein	AAAAATTTTT-N5-AAAAATTTAT	420160	420184	-	71	TU 57 (I)	R
PALG_ES_122	MHP7448_0339	c hypothetical protein	AATATGTTG-N6-CAACCTTTT	420202	420225	-	113	TU 57 (I)	R
PALG_ES_123	MHP7448_0342	c hypothetical protein	AATATAATTA-N15-TAAGGATATT	426601	426635	-	2	TU 57 (I)	R.1
PALG_ES_124	MHP7448_0343	pseudogene	TTTATGATT-N11-ATCCATAAA	427257	427285	+	182	Class 2	1.R
PALG_ES_125	MHP7448_0344	c hypothetical protein	TAGATTTCT-N5-AGAAATCCA	429370	429392	-	155	TU 57 (F)	2.R.1
PALG_ES_126	MHP7448_0344	c hypothetical protein	AAATTATTT-N13-AAATAATGT	429403	429433	-	188	TU 57 (F)	2.R
PALG_ES_127	MHP7448_0347	hypothetical protein	AAATTCAGA-N15-TATGATTTT	432258	432290	+	144	TU 58 (L)	R
PALG_ES_128	MHP7448_0348	c hypothetical protein	AAATTTATA-N11-TATTATTTT	434254	434282	-	68	TU 59 (L)	R
PALG_ES_129	MHP7448_0348	c hypothetical protein	GGATCGGGG-N7-CCCCGCACC	434365	434389	-	179	TU 59 (L)	R
PALG_ES_130	MHP7448_0348	c hypothetical protein	ATAAATTTAA-N13-TTAATTTTTT	434396	434428	-	210	TU 59 (L)	R
PALG_ES_131	MHP7448_0349	c hypothetical protein	ATTAAATTA-N12-TAATTTAGT	436152	436181	-	158	TU 59 (I)	R
PALG_ES_133	glpK	c glycerol kinase	GTGTTTTAT-N12-AAAAAAGAC	450807	450836	-	272	TU 62 (F)	1.R
PALG_ES_134	MHP7448_0360	P37-like ABC transporter substrate-binding lipoprotein	TTAGTTTAC-N14-GAAAACTGA	450838	450869	+	3	TU 63 (F)	R.2
PALG_ES_135	MHP7448_0364	conserved hypothetical protein	TAAATTTAAT-N8-AATAATTTTA	455218	455245	+	0	TU 63 (I)	R
PALG_ES_136	MHP7448_0368	lipoprotein	AATTATTTA-N5-TTAAAAATT	461987	462009	+	0	TU 63 (I)	R
PALG_ES_137	MHP7448_0377	c conserved hypothetical protein	ACTTAAAAAA-N9-TTTTTTAATT	477710	477738	-	149	TU 64 (I)	R
PALG_ES_138	MHP7448_0378	c lipoprotein	TAAGAGAAAA-N7-TTTAATCTTA	479727	479753	-	166	TU 64 (F)	1.R

PALG_ES_139	MHP7448_0379	ABC transport ATP-binding protein	TAAGATTAAA-N7-TTTTCTCTTA	479727	479753	+	175	TU 65 (F)	R.1
PALG_ES_140	MHP7448_0379	ABC transport ATP-binding protein	AATAATCTT-N15-AAAATTACT	479779	479811	+	117	TU 65 (F)	R.1
PALG_ES_141	MHP7448_0389	pseudogene	AAAAAATTA-N13-TAATTCTCT	489838	489868	+	9	Class 2	R
PALG_ES_142	MHP7448_0390	c hypothetical protein	AATACTTAA-N6-TTAAGGTTT	492402	492425	-	137	TU 66 (F)	1.R.1
PALG_ES_143	MHP7448_0390	c hypothetical protein	AAAAATTTAG-N10-CTAAATTTGT	492449	492478	-	184	TU 66 (F)	1.R.1
PALG_ES_144	MHP7448_0391	conserved hypothetical protein	GAAATTAAT-N15-ATTAATTAC	492761	492793	+	242	TU 67 (F)	1.R
PALG_ES_145	MHP7448_0391	conserved hypothetical protein	ACTAAAAAT-N5-ATTTTTTCT	492802	492824	+	211	TU 67 (F)	1.R
PALG_ES_146	MHP7448_0399	c hypothetical protein	TAATTGCCT-N11-AGGAATTTA	501266	501294	-	0	TU 68 (F)	R.1
PALG_ES_147	MHP7448_0399	c hypothetical protein	ATAAATTTA-N10-TAATTATAT	501404	501431	-	138	TU 68 (F)	1.R
PALG_ES_148	MHP7448_0400	hypothetical protein	AGCATAAGTAAT-N12-AATAGTTATGCT	501649	501684	+	33	TU 69 (F)	R.1
PALG_ES_149	MHP7448_0400	hypothetical protein	TGCAAATCA-N9-TGATTTAAA	501690	501716	+	1	TU 69 (F)	R.1
PALG_ES_150	MHP7448_0409	c hypothetical protein	TGAAATTTTT-N14-AAAAATTTAA	515231	515264	-	12	TU 70 (I)	R
PALG_ES_151	MHP7448_0410	pseudogene	TTTTTTGGTTA-N14-TAGCCCAAAAA	517253	517288	-	246	Class 2	R
PALG_ES_152	MHP7448_0411	c hypothetical protein	TTCCGTTAG-N12-CTACCGGCA	518008	518037	-	299	TU 70 (I)	2.R
PALG_ES_153	MHP7448_0412	c conserved hypothetical protein	GATAAAAAA-N5-TTTTTTAGC	519118	519140	-	61	TU 70 (I)	R
PALG_ES_154	MHP7448_0413	c conserved hypothetical protein	TATTATTGTC-N6-GACAATCACA	519687	519712	-	70	TU 70 (I)	1.R
PALG_ES_155	MHP7448_0414	c putative ICEF-II	ATAACTTTG-N10-CAATGATAT	523232	523259	-	49	TU 70 (I)	R
PALG_ES_156	trsE	c TRSE-like protein	AAAATCGCAA-N15-TTGCCTTTTT	526724	526758	-	4	TU 70 (I)	R
PALG_ES_157	trsE	c TRSE-like protein	ATTTTTTGC-N15-GCAATTAAT	526780	526812	-	60	TU 70 (I)	R
PALG_ES_158	trsE	c TRSE-like protein	ATTGCGGAT-N11-ATTTGCAAT	526817	526845	-	97	TU 70 (I)	R
PALG_ES_159	MHP7448_0422	c hypothetical protein	TGATTTTGA-N9-TGAAAATGA	534592	534618	-	10	TU 71 (I)	R
PALG_ES_160	MHP7448_0431	c conserved hypothetical protein	TTTTTTGGTTA-N14-TAGCCCAAAAA	546957	546992	-	275	TU 73 (L)	R
PALG_ES_161	MHP7448_0432	c hypothetical protein	TTAAAAAAA-N8-TTAATTTAA	548472	548497	-	5	TU 73 (I)	R
PALG_ES_162	sgaH	c 3-hexulose-6-phosphate synthase	AACTACGTT-N14-AATTTAGTT	554580	554611	-	229	TU 73 (I)	1.R
PALG_ES_164	MHP7448_0447	hypothetical protein	TTAAGATCC-N7-GCATTTTAA	585510	585534	+	23	TU 76 (I)	R
PALG_ES_165	MHP7448_0453	c ABC transporter permease protein	AGCCAAAAAA-N6-TTTTTTAGTT	598935	598960	-	78	TU 77 (F)	R
PALG_ES_166	acpD	acyl carrier protein phosphodiesterase	AACTAAAAAA-N6-TTTTTTGGCT	598935	598960	+	87	TU 78 (F)	1.R

PALG_ES_167	rpmE	50S ribosomal protein L31	TTTTTACTT-N7-AAGCATAAA	600285	600309	+	0	TU 78 (I)	R.1
PALG_ES_168	rplK	c 50S ribosomal protein L11	TTTAAAAGAA-N6-TTGTTCTAAA	603087	603112	-	0	TU 79 (F)	R.2
PALG_ES_169	MHP7448_0461	hypothetical protein	TAATAAATT-N7-AATTTATAA	603617	603641	+	74	TU 80 (F)	1.R
PALG_ES_170	MHP7448_0462	hypothetical protein	ATCCTAAAA-N14-TCTTAAGAT	604699	604730	+	1	TU 80 (I)	R
PALG_ES_171	MHP7448_0463	hypothetical protein	AGTTGATAAA-N10-TTGATCAAAT	605495	605524	+	109	TU 80 (L)	R
PALG_ES_172	pepA	c leucyl aminopeptidase	TGCTATTAA-N6-TTACTAGGA	609367	609390	-	14	TU 81 (L)	R.1
PALG_ES_173	MHP7448_0466	c hypothetical protein	AGGTTAAAA-N5-TTTTAAATT	610954	610976	-	49	TU 81 (I)	R
PALG_ES_174	MHP7448_0484	c hypothetical protein	TATTTCATCTTT-N12-AAAGGAGAAATA	632978	633013	-	0	TU 83 (F)	R.1
PALG_ES_176	MHP7448_0484	c hypothetical protein	ATTATATAA-N13-TAATTTAAT	633149	633179	-	171	TU 83 (F)	R
PALG_ES_177	MHP7448_0485	conserved hypothetical protein	TATTTCTCCTTT-N12-AAAGATGAAATA	632978	633013	+	285	mC 29	2.R
PALG_ES_179	MHP7448_0485	conserved hypothetical protein	ATTAAATTA-N13-TTATATAAT	633149	633179	+	119	mC 29	R.1
PALG_ES_180	MHP7448_0488	c conserved hypothetical protein	ATATAATTT-N7-AAATTATTT	637070	637094	-	38	TU 84 (I)	R
PALG_ES_181	pgk	phosphoglycerate kinase	AATTTTAAAGGA-N12-TGCGTTAAAATT	640144	640179	+	85	TU 85 (F)	1.R.2
PALG_ES_182	MHP7448_0495	c ISMHp1 transposase	GCAGAAATC-N13-GTTTTTTGC	647141	647171	-	73	Class 1	1.R
PALG_ES_184	pta	c phosphate acetyltransferase	TTTAAAATT-N8-AATTTTCTA	676657	676682	-	17	TU 88 (I)	R.1
PALG_ES_185	MHP7448_0511	c conserved hypothetical protein	AATTTTATC-N10-GATACAAAT	678607	678634	-	30	TU 88 (I)	1.R
PALG_ES_186	MHP7448_0512	c hypothetical protein	AAAATACAA-N7-TAATATTTT	679998	680022	-	114	TU 88 (F)	R.1
PALG_ES_187	MHP7448_0513	46K surface antigen precursor	AAAATATTA-N7-TTGTATTTT	679998	680022	+	249	TU 89 (F)	3.R
PALG_ES_188	MHP7448_0519	c hypothetical protein	TTTTAATTT-N5-AAATTAGGA	688094	688116	-	7	TU 90 (I)	R
PALG_ES_189	MHP7448_0520	c hypothetical protein	TCAAATTTTT-N14-AAAAATTTCA	688529	688562	-	60	TU 90 (I)	1.R.1
PALG_ES_190	MHP7448_0520	c hypothetical protein	AAAAATTTT-N14-AAAATTTAT	688601	688632	-	132	TU 90 (I)	2.R
PALG_ES_191	MHP7448_0520	c hypothetical protein	AATTTTATT-N6-AATAAAAGT	688727	688750	-	258	TU 90 (I)	2.R
PALG_ES_192	pepF	c oligoendopeptidase F	ATTGATTAT-N12-ATAATGATT	691593	691622	-	183	TU 90 (F)	1.R
PALG_ES_193	MHP7448_0522	conserved hypothetical protein	ATTTTAAAAT-N11-ATTTTAAATT	691502	691532	+	105	mC 30	R.1
PALG_ES_194	MHP7448_0522	conserved hypothetical protein	AATCATTAT-N12-ATAATCAAT	691593	691622	+	15	mC 30	R.2
PALG_ES_195	lon	c heat shock ATP-dependent protease	ATATAAAACA-N10-TTTTTTAAAT	701780	701809	-	0	TU 91 (I)	R
PALG_ES_196	lon	c heat shock ATP-dependent protease	ACAATTAAA-N10-TTTTATGGT	701856	701883	-	76	TU 91 (I)	R

PALG_ES_197	pgiB	c glucose-6-phosphate isomerase	ATATTAAAA-N6-TTTTATGAT	709462	709485	-	76	TU 92 (I)	R
PALG_ES_2	MHP7448-tRNA-Trp2	tRNA-Trp	ACGCATTTC-N11-GAAATTAGT	185970	185998	-	132	Class 3	R
PALG_ES_200	potA	c spermidine/putrescine ABC transporter ATP-binding	GAAGCATCA-N8-TGATGCTGC	724012	724037	-	177	TU 94 (F)	R.1
PALG_ES_201	MHP7448_0547	c conserved hypothetical protein	AAAAAGCAT-N7-ATGCTTTTT	733728	733752	-	295	TU 96 (L)	R
PALG_ES_202	MHP7448_0556	c hypothetical protein	ATTCACAATT-N11-AATTATAAAT	741553	741583	-	134	TU 97 (I)	R
PALG_ES_203	MHP7448_0557	c d-ribulose-5-phosphate 3 epimerase	CAAAATTGCGATT-N12-ATTTGCAATTTTG	742596	742633	-	235	TU 97 (I)	R
PALG_ES_204	MHP7448_0564	c hypothetical protein	AAATATCAG-N10-CTGAATTTT	752472	752499	-	82	TU 99 (F)	R
PALG_ES_205	MHP7448_0564	c hypothetical protein	TTTCTATTGT-N13-AAATTAGAAA	752581	752613	-	191	TU 99 (F)	R
PALG_ES_206	MHP7448_0565	pseudogene	TTTTATAAA-N14-TCTGTAAAA	753650	753681	-	276	Class 2	1.R
PALG_ES_207	MHP7448_0567	c hypothetical protein	AAATATCAG-N10-CTGAATTTT	756408	756435	-	82	TU 100 (F)	R
PALG_ES_208	MHP7448_0567	c hypothetical protein	TTTCTATTGT-N13-AAATTAGAAA	756517	756549	-	191	TU 100 (F)	R
PALG_ES_210	MHP7448_0570	conserved hypothetical protein	AAATTATTT-N13-AAATAATGT	757691	757721	+	224	TU 101 (L)	1.R
PALG_ES_211	MHP7448_0570	conserved hypothetical protein	TGTTATTTA-N11-TAAATAAAA	757917	757945	+	0	TU 101 (L)	R.1
PALG_ES_212	rpsD	30S ribosomal protein S4	TTTTAGTCT-N15-AAATTAAAA	765214	765246	+	8	TU 103 (F)	R.3
PALG_ES_213	MHP7448_0591	PTS system glucose-specific enzyme IIB component	AAGTTTTAAA-N11-TTTAAAACAT	784368	784398	+	4	TU 106 (F)	R
PALG_ES_214	MHP7448_0596	hypothetical protein	GTATTATCT-N6-ATAAAATAC	790655	790678	+	24	TU 106 (I)	R
PALG_ES_215	MHP7448_0612	c hypothetical protein	AAAAATTTG-N5-CTAATTATT	814876	814898	-	179	TU 107 (I)	2.R
PALG_ES_216	MHP7448_0612	c hypothetical protein	TTTCATCTA-N14-TAGTTTAAA	814915	814946	-	218	TU 107 (I)	2.R
PALG_ES_217	ddem	c cytosine specific DNA methyltransferase	CTAAAGAAA-N8-TTTATTGAG	816666	816691	-	12	TU 107 (I)	1.R
PALG_ES_218	rplJ	c 50S ribosomal protein L10	ATAATAACT-N14-ATTTGTTAT	826465	826496	-	0	TU 107 (F)	R.4
PALG_ES_219	MHP7448_0621	putative lipoprotein	ACAAAAAACT-N11-AGTATTTTTT	829385	829415	+	11	TU 108 (L)	R.1
PALG_ES_220	MHP7448_0626	c conserved hypothetical protein	ACTTTAATT-N7-AATTAAATT	840606	840630	-	105	mC 40	1.R.1
PALG_ES_221	MHP7448_0627	ABC transporter ATP-binding - Pr1-like protein	AATTTAATT-N7-AATTAAAGT	840606	840630	+	203	TU 110 (F)	R.1
PALG_ES_222	MHP7448_0630	c 5'-nucleotidase precursor	TTTAAGTTA-N11-TAACAAAAA	848189	848217	-	14	TU 111 (F)	R
PALG_ES_223	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase	TAAAAAAGT-N5-AATTTTTAA	848492	848514	+	12	TU 112 (F)	1.R.3
PALG_ES_224	MHP7448_0634	conserved hypothetical protein	GAACAAAAA-N7-TTTTTAATC	850636	850660	+	98	TU 112 (L)	R
PALG_ES_225	MHP7448_0636	c ISMHp1 transposase	GCAGAAATC-N13-GTTTTTTGC	856745	856775	-	73	Class 1	1.R

PALG_ES_229	MHP7448_0663	c adhesin like-protein P146	TCAAGAAAT-N8-ATTTTTTTA	894485	894510	-	112	TU 117 (F)	2.R.1
PALG_ES_230	MHP7448_0664	ABC transporter ATP-binding - Pr1	AGTTGACAA-N9-TTGCCAAAT	894614	894640	+	291	TU 118 (F)	1.R
PALG_ES_231	MHP7448_0664	ABC transporter ATP-binding - Pr1	TCTTTTGAA-N5-TTTAAAAAA	894814	894836	+	95	TU 118 (F)	R.1
PALG_ES_233	MHP7448_0667	hypothetical protein	TTATGTTTTT-N5-AAAAACATTA	903718	903742	+	165	TU 120 (I)	R.1
PALG_ES_234	MHP7448_0667	hypothetical protein	TTCTTTTTTT-N14-AAAAAAAGTA	903848	903881	+	26	TU 120 (I)	1.R
PALG_ES_235	MHP7448_0673	c ISMHp1 transposase	GCAGAAATC-N13-GTTTTTTGC	913892	913922	-	73	Class 1	1.R
PALG_ES_236	MHP7448_0688	c ISMHp1 transposase	TTTTTATTT-N12-AAACAAAAA	266892	266922	-	73	Class 1	1.R
PALG_ES_43	dnaN	DNA polymerase III beta subunit	AAAAGTACT-N9-AGTGGTTTT	1623	1649	+	93	TU 01 (I)	1.R
PALG_ES_44	MHP7448_0016	c hypothetical protein	TAAAAGATTT-N9-AAATTTTTCA	17123	17151	-	139	mC 01	R
PALG_ES_45	MHP7448_0016	c hypothetical protein	AGCTTAAGA-N9-TCATAAGGT	17175	17201	-	191	mC 01	R
PALG_ES_46	MHP7448_0017	hypothetical protein	TGAAAAATTT-N9-AAATCTTTTA	17123	17151	+	208	mC 02	R
PALG_ES_47	MHP7448_0019	c ABC transporter ATP-binding protein	TTTTTAGATA-N14-TATATACAAA	21303	21336	-	166	TU 02 (I)	R
PALG_ES_48	MHP7448_0022	c hypothetical protein	TTTAAGTTA-N14-TAACTTATA	26809	26840	-	22	TU 02 (I)	R.2
PALG_ES_5	MHP7448-tRNA-Ser2	tRNA-Ser	TTATAAATT-N7-AATTTATTA	603617	603641	-	126	Class 3	1.R
PALG_ES_50	MHP7448_0025	c hypothetical protein	ATATAAAATT-N12-AATTTGATTT	32710	32741	-	240	TU 02 (F)	1.R.1
PALG_ES_51	sipS	signal peptidase I	AAATCAAATT-N12-AATTTTATAT	32710	32741	+	101	TU 03 (F)	2.R.2
PALG_ES_52	sipS	signal peptidase I	TGATAAAAT-N13-ATTTAATAA	32810	32840	+	2	TU 03 (F)	1.R.2
PALG_ES_53	gatC	Asp-tRNAAsn/Glu-tRNAGln amidotransferase C subunit	TTAAATTTT-N12-AAAATTTTA	33349	33378	+	53	TU 03 (I)	1.R
PALG_ES_57	MHP7448_0054	conserved hypothetical protein	ATTTTAAAT-N12-ATTGAAATT	65533	65562	+	185	TU 06 (I)	R
PALG_ES_59	MHP7448_0065	conserved hypothetical protein	ATAACTTTTT-N7-AAAAAGTTGT	82704	82730	+	49	TU 06 (L)	1.R
PALG_ES_60	uvrC	c excinuclease ABC subunit C	TAAAATTTTT-N6-AAAAATTTCA	87212	87237	-	19	mC 06	R.1
PALG_ES_61	uvrC	c excinuclease ABC subunit C	TCTAATTTA-N15-TTAATTATA	87295	87327	-	102	mC 06	2.R
PALG_ES_63	dnaK	chaperone protein dnaK - heat shock protein 70	CTCAAAATT-N5-AAATTAGAG	87801	87823	+	15	TU 07 (I)	R.1
PALG_ES_65	himA	bacterial nucleoid DNA-binding protein	AAGTAAAAT-N8-ATCTTAGTT	93057	93082	+	0	TU 07 (I)	R
PALG_ES_66	MHP7448_0081	amino acid permease	TTGTTAAATA-N10-TAATTAACTA	101422	101451	+	108	TU 09 (F)	2.R
PALG_ES_67	deoD	c purine-nucleoside phosphorylase	ATATTTTTT-N13-AACAAAAAT	106721	106751	-	49	TU 10 (I)	1.R
PALG_ES_68	MHP7448_0085	c hypothetical protein	TTAAATTTG-N6-CAATATTAA	107342	107365	-	141	TU 10 (I)	1.R

PALG_ES_69	MHP7448_0088	c hypothetical protein	ATTAACAAT-N13-ATTCTTACT	114022	114052	-	11	TU 11 (L)	R
PALG_ES_70	MHP7448_0088	c hypothetical protein	AAATTTAATT-N12-AAACAAATTT	114070	114101	-	59	TU 11 (L)	R
PALG_ES_71	MHP7448_0090	conserved hypothetical protein	AAAAATATGT-N13-AATTATTTTT	115401	115433	+	198	TU 12 (L)	R
PALG_ES_72	MHP7448_0090	conserved hypothetical protein	GATTAAACA-N15-TCTTTATTC	115577	115609	+	22	TU 12 (L)	R.1
PALG_ES_73	gyrB	c DNA gyrase subunit B	AAATTCCAA-N13-TTTAAATTT	150974	151004	-	87	TU 16 (I)	1.R
PALG_ES_74	MHP7448_0110	c ISMHp1 transposase	GCAGAAATC-N13-GTTTTTTGC	152916	152946	-	73	Class 1	1.R
PALG_ES_76	pyk	c pyruvate kinase	TTTGCACTTT-N12-AAAATACAAA	168562	168593	-	34	TU 19 (I)	R
PALG_ES_77	MHP7448_0127	c conserved hypothetical protein	TATTAAGTAA-N10-TTAGTTAAAA	169446	169475	-	22	TU 19 (F)	R
PALG_ES_78	MHP7448_0128	hypothetical protein	TTTTAACTAA-N10-TTACTTAATA	169446	169475	+	4	mC 11	R
PALG_ES_81	MHP7448_0135	hypothetical protein	TATATATTT-N6-AAATAGGTA	173597	173620	+	9	TU 22 (F)	R
PALG_ES_82	MHP7448_0145	chromate transport protein	ACTAATTTC-N11-GAAATGCGT	185970	185998	+	2	TU 24 (F)	R
PALG_ES_83	rpsJ	c 30S ribosomal protein S10	TAAATATTT-N13-AAATAAATA	220868	220898	-	267	TU 27 (F)	1.R
PALG_ES_84	MHP7448_0197	conserved hypothetical protein	TATTTATTT-N13-AAATATTTA	220868	220898	+	230	TU 28 (F)	1.R
PALG_ES_85	MHP7448_0197	conserved hypothetical protein	AATCTGAGA-N13-TGTCAGCTT	220924	220954	+	174	TU 28 (F)	1.R
PALG_ES_86	MHP7448_0202	conserved hypothetical protein	AGTGTAATT-N5-AATTGCATT	231804	231826	+	4	TU 29 (I)	R
PALG_ES_87	lysS	lysyl-tRNA synthetase	TTACAAAAAA-N6-TTTTTACTAA	236187	236212	+	87	TU 29 (I)	R
PALG_ES_88	MHP7448_0218	ISMHp1 transposase	GAAAATTCA-N14-TGAATATGC	256989	257020	+	102	Class 1	1.R
PALG_ES_90	MHP7448_0219	c hypothetical protein	TACTGAATA-N12-TATCCAGGA	259718	259747	-	57	mC 14	R.1
PALG_ES_91	MHP7448_0219	c hypothetical protein	TGTTGCTCA-N6-TGAGGAAAA	259851	259874	-	190	mC 14	R.1
PALG_ES_92	glyA	c glycine hydroxymethyltransferase	ATTAAAAAAA-N11-TTTTTTTAGT	268462	268492	-	20	TU 32 (F)	1.R
PALG_ES_93	MHP7448_0236	transposase	GCAGAAATC-N13-GTTTTTTGC	281887	281917	+	73	Class 1	1.R
PALG_ES_94	MHP7448_0248	c hypothetical protein	TTGCCCTTG-N6-CAACGGTAA	303091	303114	-	43	TU 37 (I)	R
PALG_ES_95	MHP7448_0270	nuclease, lipoprotein	TGATCTTTT-N12-AAAAGATGA	324842	324871	+	217	TU 41 (L)	R
PALG_ES_96	MHP7448_0289	hypothetical protein	TAATGTAAT-N8-ATTTTATTA	347528	347553	+	175	TU 45 (F)	R
PALG_ES_98	MHP7448_0299	hypothetical protein	TTGTATAATT-N9-AATTATACCA	360253	360281	+	63	TU 47 (F)	R
PALG_S_1157	gidA	glucose inhibited division protein A	TATTTTATA-N6-TTTAAAAAA	2934	2957	+	25	TU 01 (I)	R
PALG_S_1158	ftsY	cell division protein ftsY	TTTATAATT-N15-AATTATAGG	7983	8015	+	8	TU 01 (I)	R.1

PALG_S_1161	MHP7448_0025	c hypothetical protein	TCTGATATT-N12-AATATCATT	32659	32688	-	189	TU 02 (F)	1.R.3
PALG_S_1162	sipS	signal peptidase I	AATGATATT-N12-AATATCAGA	32659	32688	+	154	TU 03 (F)	2.R.2
PALG_S_1163	MHP7448_0030	hypothetical protein	TGATTTTTT-N10-CAAAAATAA	36625	36652	+	14	TU 03 (L)	1.R
PALG_S_1167	dinP	DNA polymerase IV	AAAATTAAG-N10-TTTGATTTT	57694	57721	+	0	TU 06 (I)	R
PALG_S_1169	nox	NADH oxidase	AATTTTTTA-N6-TAATAAATG	102917	102940	+	180	TU 09 (L)	2.R
PALG_S_1170	secA	c preprotein translocase SecA subunit	AGATTTTAG-N7-CTAAATTTT	110433	110457	-	194	TU 10 (F)	1.R
PALG_S_1171	secA	c preprotein translocase SecA subunit	TTTTAAATTAAAAT-N14-TTTTTAATTTAAAA	110491	110532	-	252	TU 10 (F)	1.R
PALG_S_1172	MHP7448_0089	c hypothetical protein	GTTATAATT-N9-AATTTTAAT	114810	114836	-	3	TU 11 (F)	R
PALG_S_1173	MHP7448_0090	conserved hypothetical protein	AAGGTAGTA-N5-TTTTACCTT	115444	115466	+	165	TU 12 (L)	R.1
PALG_S_1175	MHP7448_0095	c conserved hypothetical protein	ATTAAAATT-N7-AATTTTACT	125173	125197	-	49	mC 09	1.R
PALG_S_1176	MHP7448_0095	c conserved hypothetical protein	TGCTTTTTAAT-N15-TTTAAAAAGTA	125203	125239	-	79	mC 09	2.R
PALG_S_1180	clpB	c ATP-dependent protease binding protein	AATTGGCAC-N13-GTGCCAAAA	136868	136898	-	53	mC 10	1.R
PALG_S_1182	MHP7448_0105	pseudogene	TGAAAAATA-N5-TATTTTTTG	138679	138701	+	1	Class 2	R
PALG_S_1184	MHP7448_0108	c protein P97 - copy 2	ACTAGTTTA-N6-TAAACTTTT	148939	148962	-	232	TU 16 (I)	R
PALG_S_1186	MHP7448_0112	hypothetical protein	AGTTAAAAA-N8-TTTTTAAAT	154679	154704	+	14	TU 16 (F)	R
PALG_S_1187	infC	c translation initiation factor IF-3	ATAATATCT-N7-AGATGTTAG	167001	167025	-	142	TU 19 (I)	R
PALG_S_1188	pyk	c pyruvate kinase	TTAATTTTA-N8-TATTATTAA	168598	168623	-	70	TU 19 (I)	R
PALG_S_1189	MHP7448_0129	c aminopeptidase	TGTATAATT-N5-AATTATATC	170894	170916	-	21	TU 20 (L)	R
PALG_S_1190	MHP7448_0155	c conserved hypothetical protein	CCAAAAATA-N14-TATTTTTTT	194821	194852	-	115	TU 25 (I)	R
PALG_S_1191	MHP7448_0159	c hypothetical protein	ATTATTAAT-N12-ATTAATTTT	197257	197286	-	1	TU 25 (I)	R
PALG_S_1192	MHP7448_0162	hypothetical protein	TAAATTTTA-N5-TAAAATTGA	200668	200690	+	248	TU 26 (I)	R
PALG_S_1193	MHP7448_0162	hypothetical protein	TATTTTAGC-N15-ACAAAAATA	200849	200881	+	57	TU 26 (I)	R
PALG_S_1195	MHP7448_0198	protein P97 - copy 1	TGATTAGAA-N11-TTCAAATCT	221716	221744	+	224	TU 28 (I)	1.R
PALG_S_1197	MHP7448_0198	protein P97 - copy 1	TATCTAAAT-N11-ATTGAGAAA	221895	221923	+	45	TU 28 (I)	R
PALG_S_1198	MHP7448_0210	c ABC transporter ATP-binding protein	TATAATATT-N9-AAAATTAGA	241601	241627	-	16	TU 30 (F)	R
PALG_S_1200	oppB-1	oligopeptide ABC transporter system permease	GTTAAAAAA-N5-TTTTTTATT	243217	243239	+	30	TU 31 (I)	R
PALG_S_1201	nrdF	c ribonucleoside-diphosphate reductase beta chain	AGGTAAAAA-N6-TTTTTAATT	264548	264571	-	194	TU 32 (I)	1.R.1

PALG_S_1204	MHP7448_0248	c hypothetical protein	ACTTGTTGCA-N14-TGCAACAATG	303200	303233	-	152	TU 37 (I)	R
PALG_S_1205	smf	c DNA processing protein SMF	TTTAAATAA-N11-TGATTTCAA	304225	304253	-	90	TU 37 (F)	R
PALG_S_1207	eno	enolase	GTATTAATT-N9-AATGAATAA	304379	304405	+	22	TU 38 (F)	R.1
PALG_S_1208	MHP7448_0255	c hypothetical protein	TAATAAAAT-N10-ATTTTATAA	309876	309903	-	15	TU 39 (F)	R.1
PALG_S_1209	MHP7448_0270	nuclease, lipoprotein	ATAAAAAAC-N6-GTTTTTGAA	324900	324923	+	165	TU 41 (L)	R
PALG_S_1210	truB	tRNA pseudouridine synthase B	TTTTTTTTT-N5-GAAAAAAAG	335930	335952	+	0	TU 43 (I)	R
PALG_S_1211	MHP7448_0279	c transcriptional regulator	CTATAAAAA-N8-TTTTTATTA	339485	339510	-	217	mC 18	2.R.1
PALG_S_1212	pyrG	CTP synthase	TAAAAATCA-N12-TGATTTTAT	339574	339603	+	182	mC 19	2.R
PALG_S_1213	pyrG	CTP synthase	TGAAATTAA-N6-TTAATTTTT	339711	339734	+	51	mC 19	1.R.2
PALG_S_1215	MHP7448_0287	c hypothetical protein	GAATTTTTGTCA-N15-AAACAAAAATTC	346757	346795	-	98	TU 44 (F)	R.1
PALG_S_1216	MHP7448_0287	c hypothetical protein	TAATAATAT-N15-GTTTTATTA	346913	346945	-	254	TU 44 (F)	R.1
PALG_S_1217	MHP7448_0289	hypothetical protein	TTAATTGCT-N8-AGAAATTAC	347433	347458	+	270	TU 45 (F)	R
PALG_S_1218	MHP7448_0289	hypothetical protein	TGTGATTTT-N5-AAAGTCAAA	347495	347517	+	211	TU 45 (F)	R
PALG_S_1219	MHP7448_0290	hypothetical protein	TAGCTTTGG-N11-CAAAAGTTA	348941	348969	+	96	TU 45 (I)	R
PALG_S_1220	rpsF	c 30S ribosomal protein S6	TTTTTCTAA-N15-GGAGAAAAA	357451	357483	-	0	TU 46 (I)	R.1
PALG_S_1221	MHP7448_0297	c hypothetical protein	TTAATATTC-N6-GTATATTAA	359359	359382	-	14	TU 46 (I)	R
PALG_S_1222	MHP7448_0298	c hypothetical protein	TGGTATAAT-N11-ATTATACAA	360253	360281	-	27	TU 46 (F)	R
PALG_S_1223	MHP7448_0305	c ABC transporter ATP-binding protein	CTAAAATCA-N6-TGATTTTTT	369375	369398	-	52	TU 48 (L)	R
PALG_S_1224	MHP7448_0312	c glycine cleavage system H protein	TTATCTATA-N9-TATAGATTT	382729	382755	-	69	TU 50 (F)	R.1
PALG_S_1225	MHP7448_0313	hypothetical protein	AAAAAAATAT-N15-TTATTTTTTA	383019	383053	+	13	TU 51 (F)	R.2
PALG_S_1227	MHP7448_0316	pseudogene	CTTTAATAG-N15-ATATTAAAA	388360	388392	+	23	Class 2	R
PALG_S_1228	MHP7448_0322	c putative transposase	ATATGAATT-N9-AATTTATAC	396447	396473	-	19	Class 1	R.1
PALG_S_1229	MHP7448_0323	transposase	ACAAGTTAA-N12-TTAACTTTA	396718	396747	+	297	Class 1	1.R
PALG_S_1234	MHP7448_0335	conserved hypothetical protein	TTTTTAATT-N8-AATTAAAGG	414226	414251	+	6	TU 56 (L)	R.1
PALG_S_1235	MHP7448_0337	hypothetical protein	ATTTCGTTA-N6-TAATGAAAA	416370	416393	+	26	mC 22	R.1
PALG_S_1236	MHP7448_0339	c hypothetical protein	CCAAAAAAT-N11-ATTTTTTTA	420265	420293	-	176	TU 57 (I)	R
PALG_S_1237	MHP7448_0339	c hypothetical protein	CTACTTTTT-N5-AAAGAGTAT	420359	420381	-	270	TU 57 (I)	R

PALG_S_1238	MHP7448_0342	c hypothetical protein	GTTTTGATA-N5-TATCATAAA	426792	426814	-	193	TU 57 (I)	R
PALG_S_1239	MHP7448_0349	c hypothetical protein	ACAATAATT-N8-AATTATTAA	436023	436048	-	29	TU 59 (I)	R
PALG_S_1240	MHP7448_0366	lipoprotein	TGAACACAA-N11-TTGTGGTTA	457442	457470	+	167	TU 63 (I)	R
PALG_S_1241	MHP7448_0373	c conserved hypothetical protein	ATTTTAATT-N14-AATCAAAAA	473253	473284	-	19	TU 64 (I)	R
PALG_S_1243	sgaA	c PTS system enzyme IIB component	CCAAAAAAT-N5-ATTTTTTTA	473898	473920	-	1	TU 64 (L)	R
PALG_S_1244	MHP7448_0375	c PTS system enzyme IIB component	TTTAATATT-N6-AATATTAGA	474507	474530	-	295	TU 64 (I)	R
PALG_S_1246	MHP7448_0386	pseudogene	ATTTCCTTA-N14-AAAGGAAAA	487896	487927	+	106	Class 2	1.R
PALG_S_1247	MHP7448_0391	conserved hypothetical protein	AAAATAAGA-N12-TCTTTTTTC	492715	492744	+	291	TU 67 (F)	1.R
PALG_S_1248	MHP7448_0399	c hypothetical protein	TGTCATTTA-N6-TCAATAACA	501436	501459	-	170	TU 68 (F)	1.R
PALG_S_1249	MHP7448_0400	hypothetical protein	ATATAATTA-N10-TAAATTTAT	501404	501431	+	286	TU 69 (F)	R
PALG_S_1250	MHP7448_0400	hypothetical protein	TGTTATTGA-N6-TAAATGACA	501436	501459	+	258	TU 69 (F)	R.1
PALG_S_1251	asnS	asparaginyl-tRNA synthetase	GTATAATAA-N5-TTTTTATAT	503728	503750	+	13	TU 69 (I)	R
PALG_S_1253	MHP7448_0411	c hypothetical protein	TTAAAGTTT-N5-AAACTTTTG	517771	517793	-	62	TU 70 (I)	R.2
PALG_S_1254	MHP7448_0413	c conserved hypothetical protein	TGATATTAA-N6-TTGATATCT	519817	519840	-	200	TU 70 (I)	1.R
PALG_S_1255	trsE	c TRSE-like protein	AAAAATTAA-N14-AGAATTTTT	526848	526879	-	128	TU 70 (I)	R
PALG_S_1256	MHP7448_0422	c hypothetical protein	TATTTTTTG-N8-TAAAAAACA	534693	534718	-	111	TU 71 (I)	R
PALG_S_1257	MHP7448_0431	c conserved hypothetical protein	TAAAAAACT-N14-GGTTTTTAA	546730	546761	-	48	TU 73 (L)	R
PALG_S_1258	MHP7448_0433	c hypothetical protein	ATTTTTTAT-N10-ATAAAAATA	550615	550642	-	17	TU 73 (I)	R
PALG_S_1259	MHP7448_0433	c hypothetical protein	GGTATTATT-N8-AATAATATT	550669	550694	-	71	TU 73 (I)	R
PALG_S_1260	MHP7448_0440	c hypothetical protein	TCTTCAAGTA-N15-AACTTGAAAA	557854	557888	-	28	TU 73 (F)	R.1
PALG_S_1262	MHP7448_0441	hypothetical protein	TTTTCAAGTT-N15-TACTTGAAGA	557854	557888	+	112	mC 27	R.1
PALG_S_1263	MHP7448_0443	c hypothetical protein	TTTAGAAAA-N10-TTTTATAAT	562702	562729	-	58	TU 74 (L)	1.R
PALG_S_1264	MHP7448_0443	c hypothetical protein	CGCTAAATT-N7-AATTAAGCA	562833	562857	-	189	TU 74 (L)	2.R
PALG_S_1265	MHP7448_0443	c hypothetical protein	AAATTTTTG-N15-AAAAAATAT	562936	562968	-	292	TU 74 (L)	2.R
PALG_S_1266	MHP7448_0711	hypothetical protein	AAATTAGTTC-N14-GAACTAATGG	582009	582042	-	3	TU 76 (I)	R
PALG_S_1267	MHP7448_0711	hypothetical protein	TGAATTTAA-N11-ATAAATTTA	582052	582080	-	46	TU 76 (I)	R
PALG_S_1268	MHP7448_0711	hypothetical protein	TAAAAAAAC-N7-GTAGTTTTA	582128	582152	-	122	TU 76 (I)	R

PALG_S_1269	MHP7448_0447	hypothetical protein	TAAAAATTT-N6-AAAGTATTA	585354	585377	+	180	TU 76 (I)	R
PALG_S_1270	MHP7448_0447	hypothetical protein	TGGATCTAA-N5-TTAGATCAA	585434	585456	+	101	TU 76 (I)	R
PALG_S_1271	MHP7448_0452	c ABC transporter ATP-binding protein	ACAAAAAAA-N7-TTTTTTTAC	590808	590832	-	18	TU 77 (I)	1.R
PALG_S_1272	MHP7448_0461	hypothetical protein	TTAAAGTTT-N5-AAACTTTTG	603644	603666	+	49	TU 80 (F)	R.1
PALG_S_1273	MHP7448_0463	hypothetical protein	AAAATAATT-N10-AATTATTAA	605527	605554	+	79	TU 80 (L)	R
PALG_S_1274	MHP7448_0465	c methlytransferase	TTTAAAACA-N5-TTTTTTAAT	610102	610124	-	0	TU 81 (I)	R
PALG_S_1275	MHP7448_0470	c conserved hypothetical protein	TAGTAAAAA-N7-ATTTTACAA	616609	616633	-	45	TU 81 (F)	R.1
PALG_S_1276	MHP7448_0470	c conserved hypothetical protein	AATAAAAAA-N7-TTTTTGATA	616703	616727	-	139	TU 81 (F)	1.R
PALG_S_1277	MHP7448_0471	conserved hypothetical protein	TTGTAAAAT-N7-TTTTTACTA	616609	616633	+	111	TU 82 (F)	1.R.1
PALG_S_1278	MHP7448_0471	conserved hypothetical protein	ACTGTTTTT-N11-AAAAAAATT	616715	616743	+	1	TU 82 (F)	1.R.2
PALG_S_1279	MHP7448_0486	c hypothetical protein	AGACAATTTC-N14-GAAATTGTAA	634870	634903	-	27	TU 84 (L)	1.R
PALG_S_1281	fruA	PTS system, fructose-specific IIABC component	GCATTATTT-N8-AAATAATCT	641978	642003	+	17	TU 85 (I)	R
PALG_S_1286	xylG	xylose ABC transporter ATP-binding protein	AGTTAAAAA-N9-ATTTTAATT	681542	681568	+	89	TU 89 (I)	1.R
PALG_S_1287	MHP7448_0520	c hypothetical protein	TTTTTGAAT-N5-ATTGAAAAG	688477	688499	-	8	TU 90 (I)	1.R.1
PALG_S_1289	upp	c uracil phosphoribosyltransferase	TTTTTAAAATTT-N14-CTATTTTAAAAA	703310	703347	-	47	TU 91 (I)	R
PALG_S_1290	pmsR	c peptide methionine sulfoxide reductase	TTATTTTTT-N5-AAGAAATAC	707502	707524	-	1	TU 92 (L)	R.1
PALG_S_1291	pgiB	c glucose-6-phosphate isomerase	TAAAAAATT-N5-AATTTCTAA	709399	709421	-	13	TU 92 (I)	R
PALG_S_1292	MHP7448_0543	hypothetical protein	TTCTAATTT-N10-AAATGAGAT	724589	724616	+	217	TU 95 (I)	1.R
PALG_S_1295	MHP7448_0556	c hypothetical protein	TCTTAAAAA-N5-TTTTCAAAA	741496	741518	-	77	TU 97 (I)	R
PALG_S_1296	MHP7448_0557	c d-ribulose-5-phosphate 3 epimerase	TTTTATAAA-N14-TCTGTAAAA	742636	742667	-	275	TU 97 (I)	R
PALG_S_1297	MHP7448_0561	c hypothetical protein	TCAAATTGT-N5-ATAATTTGC	748532	748554	-	23	TU 97 (I)	R
PALG_S_1298	MHP7448_0564	c hypothetical protein	TCAAATTGT-N5-ATAATTTGC	752413	752435	-	23	TU 99 (F)	R
PALG_S_1300	MHP7448_0567	c hypothetical protein	TCAAATTGT-N5-ATAATTTGC	756349	756371	-	23	TU 100 (F)	R
PALG_S_1304	MHP7448_0570	conserved hypothetical protein	ACTTTTAGG-N10-CCTTAAATT	757889	757916	+	29	TU 101 (L)	R
PALG_S_1306	nagB	c glucosamine-6-phosphate isomerase	AATTTTAAT-N14-ATTAAAAAG	764894	764925	-	6	TU 102 (I)	R
PALG_S_1308	rpiB	ribose 5-phosphate isomerase B	ATATAATTT-N5-AAATTCTAG	765956	765978	+	7	TU 103 (I)	R
PALG_S_1309	tdk	thymidine kinase	AAATCTAAT-N14-TTTAGGTTT	807873	807904	+	12	TU 106 (I)	R

PALG_S_1312	MHP7448_0612	c hypothetical protein	TTTTTTTGAAA-N15-CTGCAAAAAAA	814780	814816	-	83	TU 107 (I)	2.R
PALG_S_1313	MHP7448_0620	conserved hypothetical protein	AATTATATC-N9-AATATAAAT	826773	826799	+	40	TU 108 (F)	R.2
PALG_S_1314	dam	c DNA adenine methylase	TTATAATTATAC-N15-GTATAATTATTT	833762	833800	-	254	mC 39	2.R.1
PALG_S_1315	MHP7448_0623	ABC transporter ATP-binding - Pr1	CTGTTATTA-N9-TAACAACAA	833853	833879	+	282	TU 109 (F)	1.R.1
PALG_S_1316	MHP7448_0626	c conserved hypothetical protein	TATTCCTTT-N14-TAATGAATA	840638	840669	-	137	mC 40	1.R
PALG_S_1317	MHP7448_0627	ABC transporter ATP-binding - Pr1-like protein	TATTCATTA-N14-AAAGGAATA	840638	840669	+	164	TU 110 (F)	1.R
PALG_S_1321	MHP7448_0667	hypothetical protein	ATAAATACT-N7-TGTATTTTT	903745	903769	+	138	TU 120 (I)	R.1
PALG_S_1323	MHP7448_0680	c hypothetical protein	TTATAGTAT-N8-ATATTATTA	50766	50791	-	16	TU 05 (I)	R
PALG_S_1327	MHP7448_0683	hypothetical protein	TAAAAATTATA-N14-TTTAATTTTAA	114824	114859	+	45	TU 12 (F)	R
PALG_S_1328	MHP7448_0683	hypothetical protein	TCAGACACA-N10-GGTGTCTCA	114875	114902	+	2	TU 12 (F)	R
PALG_S_1331	MHP7448_0684	hypothetical protein	CGGTTAATA-N10-TATCAACCT	163064	163091	+	68	TU 18 (I)	R
PALG_S_1334	MHP7448_0687	hypothetical protein	ATAATAAAT-N14-ATTTTTTTT	260037	260068	+	201	mC 15	R
PALG_S_1335	MHP7448_0687	hypothetical protein	TAAATAAAAAA-N15-TTTTTTATTAT	260087	260123	+	146	mC 15	R
PALG_S_1336	MHP7448_0687	hypothetical protein	CGAATAAAAAAA-N15-TTTTTTTATTAT	260182	260220	+	49	mC 15	R
PALG_S_1338	MHP7448_0688	c ISMHp1 transposase	CATCAAACT-N13-AAATTGATG	266955	266985	+	199	Class 1	1.R
PALG_S_1340	MHP7448_0688	c ISMHp1 transposase	TAAAAAAAT-N15-CTTTTTTAA	267149	267181	+	3	Class 1	1.R
PALG_S_1349	MHP7448_0694	c hypothetical protein	CTAAATTTT-N14-AATATTTTG	421384	421415	+	212	TU 57 (F)	1.R
PALG_S_1357	MHP7448_0697	c hypothetical protein	ATGATTATA-N8-TATAATTTT	527252	527277	+	48	TU 70 (I)	R
PALG_S_1358	MHP7448_0697	c hypothetical protein	CAAAGTCAA-N8-TGGACTTTA	527278	527303	+	22	TU 70 (I)	R
PALG_S_1359	MHP7448_0698	c hypothetical protein	TCAATTCTT-N12-AAAAATTTA	529082	529111	-	2	TU 70 (I)	R
PALG_S_1360	MHP7448_0698	c hypothetical protein	TCTTGTAAA-N14-TTAACAAAA	529114	529145	-	34	TU 70 (I)	R
PALG_S_1361	MHP7448_0698	c hypothetical protein	TTCTTTTTT-N11-AAAAAAGTA	529175	529203	-	95	TU 70 (I)	R
PALG_S_1364	MHP7448_0703	hypothetical protein	GTTCATTTT-N10-AATATGAAA	568938	568965	+	192	TU 75 (I)	R
PALG_S_1366	MHP7448_0703	hypothetical protein	TTTTAGATA-N11-TATCAAAAG	569113	569141	+	16	TU 75 (I)	R
PALG_S_1373	MHP7448_0707	hypothetical protein	GCTCAAGTT-N9-AATTTGATC	571125	571151	+	257	TU 75 (I)	R
PALG_S_1375	MHP7448_0707	hypothetical protein	GTTTTGGAT-N15-GTTCAAAAC	571371	571403	+	5	TU 75 (I)	R
PALG_S_1381	MHP7448_0712	hypothetical protein	ACTTTAAAT-N8-ATTTAAATA	584142	584167	+	282	TU 76 (I)	R

PALG_S_1385	MHP7448_0714	hypothetical protein	GCAGCATCA-N8-TGATGCTTC	724012	724037	+	234	TU 95 (F)	2.R
PALG_S_1391	MHP7448_0719	hypothetical protein	AAAATTCAG-N10-CTGATATTT	752472	752499	+	251	mC 35	1.R
PALG_S_1394	MHP7448_0720	hypothetical protein	AAAATTCAG-N10-CTGATATTT	756408	756435	+	251	TU 101 (F)	1.R
PALG_S_1395	MHP7448_0720	hypothetical protein	GTTTTAAGT-N12-AATTAAAAA	756460	756489	+	197	TU 101 (F)	1.R.1
PALG_S_1396	MHP7448_0720	hypothetical protein	TTTCTAATT-N15-CAATAGAAA	756517	756549	+	137	TU 101 (F)	R.1
PALG_S_1398	MHP7448_0722	c conserved hypothetical protein	ATAATAATA-N7-TATAATTAG	777761	777785	+	281	TU 104 (I)	2.R.2
PALG_S_1399	MHP7448_0722	c conserved hypothetical protein	ATATAATAA-N9-TTTTTATGT	777807	777833	+	233	TU 104 (I)	2.R.2
PALG_S_1403	rRNA-5s	5S ribosomal RNA	TATTTTTTG-N13-CAAAATAAA	799353	799383	-	41	Class 4	1.R
PALG_S_1404	rRNA-5s	5S ribosomal RNA	CACTAAAAA-N6-TTTTTACTA	799404	799427	-	92	Class 4	1.R
PALG_S_36	MHP7448-tRNA-Trp1	tRNA-Trp	TCCTTAATT-N9-AATTAAGAT	125018	125044	+	5	Class 3	1.R.1
PALG_S_38	MHP7448-tRNA-Tyr1	tRNA-Tyr	TTTTCCTTT-N14-TAAGGAAAT	487896	487927	-	156	Class 3	R.1
PALG_S_39	MHP7448-tRNA-Ser2	tRNA-Ser	TTATTTTAT-N7-TTAAAATCA	603684	603708	-	193	Class 3	1.R
SSR_01_I	MHP7448_0386	pseudogene	(ATT)$_{9}$	487810	487836	+	197	Class 2	1.R
SSR_01_L	MHP7448_0236	transposase	(ATTAATTACA)$_{2}$	281937	281956	+	34	Class 1	1.R
SSR_02_LI	MHP7448_0016	hypothetical protein	(TATT)$_{6}$	902991	903014	-	82	mC 01	R
SSR_03_LI	MHP7448_0017	hypothetical protein	(ATAA)$_{6}$	17066	17089	+	270	mC 02	R
SSR_05_I	MHP7448-tRNA-Tyr1	tRNA-tyr	(TAA)$_{9}$	432244	432270	-	70	Class 3	1.R.1
SSR_05_LI	MHP7448_0087	GTP-binding protein	(TTAAAAAAA)$_{2}$	111128	111145	+	458	mC 07	1.R
SSR_06_LI	MHP7448_0108	protein P97 - copy 2	(TTGTTTTTT)$_{2}$	771177	771194	-	179	TU 16 (I)	1.R
SSR_09_LI	MHP7448_0138	hypothetical protein	(AAAATTTTT)$_{2}$	177486	177503	+	230	TU 23 (F)	1.R
SSR_11_LI	nrdF	ribonucleoside-diphosphate reductase beta chain	(AAT)$_{6}$	655633	655650	-	76	TU 32 (I)	R.2
SSR_13_LI	glyA	glycine hydroxymethyltransferase	(TTTTTGTTC)$_{2}$	651320	651337	-	301	TU 32 (F)	2.R
SSR_14_LI	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	(AAAAACAAG)$_{2}$	268743	268760	+	178	TU 33 (F)	1.R
SSR_15_LI	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	(TAATTTTTT)$_{2}$	268480	268497	+	441	TU 33 (F)	1.R
SSR_16_LI	MHP7448_0252	hypothetical protein	(ATTT)$_{10}$	307111	307150	+	19	TU 38 (I)	R
SSR_17_LI	MHP7448_0257	hypothetical protein	(TCTAATAAA)$_{2}$	312151	312168	+	48	TU 40 (I)	1.R
SSR_19_LI	pyrG	CTP synthase	(ATTTTTAAAA)$_{2}$	339287	339306	+	479	mC 19	3.R

SSR_23_LI	MHP7448_0366	lipoprotein	(GATAATTTT)$_2$	457501	457518	+	119	TU 63 (I)	R.1
SSR_26_LI	MHP7448_0443	hypothetical protein	(GAAAAAAAAA)$_2$	357278	357297	-	139	TU 74 (I)	1.R.2
SSR_28_LI	MHP7448_0461	hypothetical protein	(TTA)$_6$	603222	603239	+	476	TU 80 (F)	1.R
SSR_29_LI	MHP7448_0463	hypothetical protein	(GA)$_{11}$	605422	605443	+	190	TU 80 (I)	R.2
SSR_30_LI	MHP7448_0484	hypothetical protein	(TAT)$_9$	286935	286961	-	141	TU 83 (F)	R.1
SSR_31_LI	MHP7448_0495	ISMHp1 transposase	(AAAAAAAAT)$_2$	272803	272820	-	192	Class 1	1.R
SSR_35_LI	tufA	elongation factor EF-Tu	(TAATA)$_4$	221157	221176	-	35	TU 91 (L)	1.R.4
SSR_38_LI	MHP7448_0565	pseudogene	(ATAAAT)$_4$	166673	166696	-	10	Class 2	R.1
SSR_39_LI	MHP7448_0598	hypothetical protein	(TTA)$_4$	793873	793884	+	297	TU 106 (I)	R
SSR_41_LI	dam	DNA adenine methylase	(GT)$_5$	86162	86171	-	401	mC 39	2.R
SSR_42_LI	MHP7448_0623	ABC transporter ATP-binding - Pr1	(CA)$_5$	833909	833918	+	243	TU 109 (F)	1.R.1
SSR_45_LI	MHP7448_0688	ISMHp1 transposase	(TTAACATTAA)$_2$	653211	653230	-	31	Class 1	1.R
SSR_46_LI	MHP7448_0698	hypothetical protein	(ATTTA)$_4$	390807	390826	-	174	TU 70 (I)	R
SSR_49_LI	MHP7448_0485	hypothetical protein	(ATA)$_9$	633119	633145	+	153	mC 29	1.R.1
SSRM_05_TL	MHP7448_0022	hypothetical protein	(T)$_{21}$	893196	893216	-	77	TU 02 (I)	2.R
SSRM_06_TL	MHP7448_0025	hypothetical protein	(T)$_{21}$	887469	887489	-	121	TU 02 (F)	3.R.1
SSRM_07_TL	sipS	signal peptidase I	(A)$_{21}$	32591	32611	+	231	TU 03 (F)	2.R.1
SSRM_08_TL	dnaJ	heat shock protein DnaJ	(A)$_{21}$	89807	89827	+	30	TU 07 (I)	1.R
SSRM_09_L	parE	topoisomerase IV subunit B	(A)$_8$	874944	874951	-	149	TU 04 (F)	2.R
SSRM_09_TL	rnpB	rnpB	(T)$_{25}$	138111	138135	+	76	TU 15 (I)	1.R
SSRM_10_L	gap	glyceraldehyde 3phosphate dehydrogenase	(T)$_8$	45129	45136	+	53	mC 03	R.2
SSRM_103_L	MHP7448_0308	hypothetical protein	(A)$_8$	372605	372612	+	424	TU 49 (I)	3.R
SSRM_104_L	MHP7448_0308	hypothetical protein	(A)$_8$	372642	372649	+	387	TU 49 (I)	3.R
SSRM_105_L	MHP7448_0312	glycine cleavage system H protein	(T)$_8$	537055	537062	-	358	TU 50 (F)	1.R
SSRM_109_L	MHP7448_0322	putative transposase	(A)$_8$	523314	523321	-	331	Class 1	1.R
SSRM_11_TL	MHP7448_0117	hypothetical protein	(T)$_{16}$	160968	160983	+	311	TU 18 (F)	1.R
SSRM_116_L	MHP7448_0334	hypothetical protein	(T)$_8$	412739	412746	+	378	TU 56 (F)	2.R

SSRM_12_L	MHP7448_0042	hypothetical protein	$(A)_8$	55630	55637	+	29	TU 06 (I)	R
SSRM_121_L	MHP7448_0347	hypothetical protein	$(T)_8$	432215	432222	+	212	TU 58 (I)	1.R
SSRM_13_TL	rpsT	30S ribosomal protein S20	$(A)_{21}$	294017	294037	+	20	TU 36 (I)	R.2
SSRM_131_L	MHP7448_0366	lipoprotein	$(T)_8$	457330	457337	+	300	TU 63 (I)	1.R
SSRM_132_L	MHP7448_0366	lipoprotein	$(T)_8$	457527	457534	+	103	TU 63 (I)	R.1
SSRM_137_L	MHP7448_0386	pseudogene	$(T)_8$	487772	487779	+	254	Class 2	1.R
SSRM_14_L	MHP7448_0064	hypothetical protein	$(T)_8$	77947	77954	+	11	TU 06 (I)	R
SSRM_14_TL	MHP7448_0263	ABC transporter ATPbinding protein	$(A)_{19}$	315885	315903	+	8	TU 40 (I)	R
SSRM_143_L	MHP7448_0399	hypothetical protein	$(T)_9$	418452	418460	-	354	TU 68 (F)	1.R
SSRM_144_L	MHP7448_0400	hypothetical protein	$(A)_9$	501620	501628	+	89	TU 69 (F)	R.1
SSRM_145_L	asnS	asparaginyltRNA synthetase	$(A)_{21}$	503371	503391	+	372	TU 69 (I)	3.R
SSRM_149_L	MHP7448_0414	putative ICEFII	$(T)_8$	396452	396459	-	438	TU 70 (I)	1.R
SSRM_15_TL	MHP7448_0272	P97like protein	$(T)_{19}$	588105	588123	-	217	TU 42 (F)	2.R
SSRM_150_L	MHP7448_0426	hypothetical protein	$(A)_8$	379377	379384	-	2	TU 71 (F)	R
SSRM_151_L	efp	elongation factor EFP	$(T)_8$	540696	540703	+	81	TU 72 (F)	1.R.1
SSRM_152_L	MHP7448_0431	hypothetical protein	$(A)_8$	373136	373143	-	255	TU 73 (I)	1.R.1
SSRM_154_L	MHP7448_0431	hypothetical protein	$(T)_8$	373077	373084	-	314	TU 73 (I)	2.R
SSRM_16_TL	pheS	phenylalanyltRNA synthetase alpha chain	$(A)_{19}$	331957	331975	+	264	TU 43 (F)	1.R.1
SSRM_161_L	MHP7448_0443	hypothetical protein	$(A)_9$	357251	357259	-	177	TU 74 (I)	1.R.2
SSRM_162_L	MHP7448_0443	hypothetical protein	$(A)_9$	357036	357044	-	392	TU 74 (I)	3.R
SSRM_165_L	MHP7448_0453	ABC transporter permease protein	$(T)_9$	321058	321066	-	157	TU 77 (F)	R
SSRM_169_L	MHP7448_0463	hypothetical protein	$(A)_9$	605143	605151	+	482	TU 80 (I)	2.R
SSRM_179_L	MHP7448_0484	hypothetical protein	$(T)_8$	286884	286891	-	211	TU 83 (F)	1.R
SSRM_184_L	MHP7448_0485	hypothetical protein	$(A)_8$	633189	633196	+	102	mC 29	R.2
SSRM_195_L	MHP7448_0505	lipoprotein	$(A)_8$	249413	249420	-	246	TU 86 (F)	R.1
SSRM_196_L	pdhC	dihydrolipoamide acetyltransferase	$(A)_8$	671046	671053	+	398	TU 87 (F)	3.R
SSRM_198_L	pta	phosphate acetyltransferase	$(A)_8$	243372	243379	-	61	TU 88 (I)	1.R

SSRM_199_L	MHP7448_0512	hypothetical protein	(T)$_{9}$	239870	239878	-	318	TU 88 (F)	1.R
SSRM_20_L	MHP7448_0088	hypothetical protein	(T)$_{8}$	805959	805966	-	103	TU 11 (I)	R
SSRM_20_TL	rpsF	30S ribosomal protein S6	(T)$_{25}$	562494	562518	-	111	TU 46 (I)	1.R
SSRM_201_L	MHP7448_0520	hypothetical protein	(A)$_{8}$	231185	231192	-	419	TU 90 (I)	1.R
SSRM_21_TL	MHP7448_0300	hypothetical protein	(A)$_{23}$	362068	362090	+	42	TU 47 (I)	R.2
SSRM_216_L	MHP7448_0556	hypothetical protein	(T)$_{8}$	178630	178637	-	24	TU 97 (I)	R
SSRM_217_L	MHP7448_0559	hypothetical protein	(A)$_{9}$	174492	174500	-	108	TU 97 (I)	R.1
SSRM_22_TL	MHP7448_0302	permease	(A)$_{18}$	364483	364500	+	116	TU 47 (I)	2.R
SSRM_224_L	MHP7448_0579	hypothetical protein	(T)$_{8}$	152026	152033	-	15	TU 104 (I)	R
SSRM_227_L	MHP7448_0596	hypothetical protein	(T)$_{23}$	790610	790632	+	70	TU 106 (F)	R
SSRM_23_L	MHP7448_0093	hypothetical protein	(T)$_{8}$	797381	797388	-	73	TU 13 (F)	1.R
SSRM_23_TL	MHP7448_0342	hypothetical protein	(T)$_{18}$	493125	493142	-	339	TU 57 (I)	1.R
SSRM_230_L	xylF	ABC transporter xylosebinding lipoprotein	(T)$_{8}$	802003	802010	+	190	TU 106 (I)	1.R
SSRM_234_L	rplJ	50S ribosomal protein L10	(A)$_{21}$	93486	93506	-	109	TU 107 (F)	3.R.1
SSRM_235_L	rplJ	50S ribosomal protein L10	(A)$_{23}$	93509	93531	-	84	TU 107 (F)	3.R.1
SSRM_236_L	MHP7448_0620	hypothetical protein	(T)$_{21}$	826574	826594	+	245	TU 108 (F)	3.R
SSRM_237_L	MHP7448_0620	hypothetical protein	(T)$_{23}$	826549	826571	+	268	TU 108 (F)	3.R
SSRM_238_L	MHP7448_0621	putative lipoprotein	(A)$_{8}$	829355	829362	+	64	TU 108 (I)	R.1
SSRM_24_TL	MHP7448_0373	hypothetical protein	(T)$_{21}$	446712	446732	-	114	TU 64 (I)	1.R
SSRM_240_L	MHP7448_0623	ABC transporter ATPbinding Pr1	(A)$_{8}$	833781	833788	+	373	TU 109 (F)	3.R
SSRM_247_L	nusG	transcription antitermination protein	(A)$_{8}$	62608	62615	-	13	TU 113 (I)	R
SSRM_248_L	nusG	transcription antitermination protein	(T)$_{8}$	62548	62555	-	73	TU 113 (I)	R
SSRM_249_L	rplM	50S ribosomal protein L13	(T)$_{17}$	47460	47476	-	32	TU 115 (I)	R.1
SSRM_25_L	MHP7448_0095	hypothetical protein	(A)$_{9}$	794917	794925	-	261	mC 09	R.1
SSRM_25_TL	MHP7448_0386	pseudogene	(T)$_{21}$	487789	487809	+	224	Class 2	1.R
SSRM_253_L	MHP7448_0670	hypothetical protein	(T)$_{8}$	906617	906624	+	63	TU 120 (I)	2.R
SSRM_256_L	MHP7448_0094	hypothetical protein	(A)$_{8}$	122692	122699	+	18	mC 08	1.R.1

SSRM_259_L	MHP7448_0687	hypothetical protein	$(A)_8$	260187	260194	+	75	mC 15	R
SSRM_26_L	tpx	thiol peroxidase	$(T)_9$	125155	125163	+	100	TU 14 (F)	R.2
SSRM_26_TL	asnS	asparaginyltRNA synthetase	$(T)_{25}$	503319	503346	+	417	TU 69 (I)	3.R
SSRM_263_L	MHP7448_0694	hypothetical protein	$(A)_8$	498834	498841	-	111	TU 57 (I)	3.R.1
SSRM_264_L	MHP7448_0699	hypothetical protein	$(A)_8$	388936	388943	-	82	TU 70 (F)	1.R
SSRM_265_L	MHP7448_0700	hypothetical protein	$(T)_8$	531137	531144	+	2	mC 26	R
SSRM_266_L	MHP7448_0714	hypothetical protein	$(A)_8$	723908	723915	+	356	TU 95 (F)	2.R
SSRM_27_TL	MHP7448_0426	hypothetical protein	$(A)_{15}$	379342	379356	-	30	TU 71 (F)	R
SSRM_270_L	rRNA-16s	RNA ribossomal 16S	$(A)_8$	16222	16229	-	490	Class 4	3.R
SSRM_273_L	rRNA-23s	RNA ribossomal 23S	$(A)_8$	18516	18523	-	216	Class 4	1.R.1
SSRM_274_L	rRNA-5s	RNA ribossomal 5S	$(T)_8$	799453	799460	+	16	Class 4	1.R.1
SSRM_279_L	MHP7448-tRNA-His1	tRNAhis	$(T)_8$	759351	759358	-	58	Class 3	1.R
SSRM_28_TL	efp	elongation factor EFP	$(T)_{15}$	540724	540738	+	46	TU 72 (F)	1.R.1
SSRM_280_L	MHP7448-tRNA-His1	tRNAhis	$(T)_8$	759245	759252	-	164	Class 3	1.R
SSRM_284_L	MHP7448-tRNA-Asn1	tRNAAsp	$(A)_{20}$	335382	335401	+	61	Class 3	2.R
SSRM_285_L	MHP7448-tRNA-Tyr1	tRNAtyr	$(A)_8$	432301	432308	-	32	Class 3	1.R.1
SSRM_29_L	tpiA	triosephosphate isomerase	$(T)_8$	136897	136904	+	92	TU 15 (F)	1.R
SSRM_29_TL	MHP7448_0439	hypothetical protein	$(T)_{17}$	363153	363169	-	37	TU 73 (I)	1.R
SSRM_30_TL	MHP7448_0461	hypothetical protein	$(T)_{23}$	603570	603592	+	123	TU 80 (F)	1.R
SSRM_31_L	MHP7448_0110	ISMHp1 transposase	$(T)_8$	767012	767019	-	218	TU 16 (I)	1.R
SSRM_31_TL	MHP7448_0477	hypothetical protein	$(A)_{19}$	296167	296185	-	106	TU 83 (I)	3.R.1
SSRM_32_L	MHP7448_0110	ISMHp1 transposase	$(A)_9$	767028	767036	-	201	TU 16 (I)	1.R
SSRM_35_TL	MHP7448_0557	dribulose5phosphate 3 epimerase	$(A)_{20}$	177660	177679	-	40	TU 97 (I)	R
SSRM_36_TL	MHP7448_0559	hypothetical protein	$(A)_{15}$	174537	174551	-	57	TU 18 (F)	1.R
SSRM_38_TL	MHP7448_0656	prolipoprotein p65	$(T)_{17}$	876503	876519	+	53	TU 116 (I)	1.R.1
SSRM_39_TL	MHP7448-tRNA-His1	tRNAhis	$(A)_{16}$	759097	759112	-	304	Class 3	1.R
SSRM_41_TL	MHP7448-tRNA-Tyr1	tRNAtyr	$(A)_{21}$	432271	432291	-	49	Class 3	1.R.1

SSRM_42_TL	MHP7448-tRNA-Ser2	tRNAser	(A)$_{23}$	316488	316510	-	79	Class 3	1.R
SSRM_43_TL	MHP7448-tRNA-Gly1	tRNAgly	(T)$_{25}$	808472	808496	+	43	Class 3	1.R
SSRM_46_L	ldh	Llactate dehydrogenase	(A)$_{8}$	176470	176477	+	49	TU 22 (I)	R
SSRM_52_L	rpsM	30S ribosomal protein S13	(A)$_{8}$	712623	712630	-	55	TU 27 (I)	R
SSRM_58_L	trmE	tRNA modification GTPase ThdF/TrmE	(T)$_{8}$	680135	680142	-	13	TU 30 (I)	R
SSRM_60_L	oppB1	oligopeptide ABC transporter system permease	(A)$_{8}$	242859	242866	+	403	TU 31 (I)	1.R.2
SSRM_62_L	MHP7448_0219	hypothetical protein	(A)$_{8}$	660297	660304	-	115	mC 14	R.1
SSRM_64_L	nrdF	ribonucleosidediphosphate reductase beta chain	(T)$_{8}$	655386	655393	-	333	TU 32 (I)	2.R
SSRM_65_L	glyA	glycine hydroxymethyltransferase	(A)$_{8}$	651581	651588	-	50	TU 32 (F)	2.R
SSRM_66_L	glyA	glycine hydroxymethyltransferase	(T)$_{8}$	651294	651301	-	337	TU 32 (F)	2.R
SSRM_67_L	glyA	glycine hydroxymethyltransferase	(T)$_{8}$	651256	651263	-	375	TU 32 (F)	2.R
SSRM_68_L	glyA	glycine hydroxymethyltransferase	(T)$_{8}$	651175	651182	-	456	TU 32 (F)	2.R
SSRM_69_L	glyA	glycine hydroxymethyltransferase	(T)$_{10}$	651435	651444	-	194	TU 32 (F)	2.R
SSRM_70_L	MHP7448_0225	methylmalonatesemialdehyde dehydrogenase	(A)$_{8}$	268779	268786	+	152	TU 33 (F)	1.R
SSRM_74_L	MHP7448_0225	methylmalonatesemialdehyde dehydrogenase	(A)$_{10}$	268636	268645	+	293	TU 33 (F)	1.R
SSRM_78_L	rpsT	30S ribosomal protein S20	(A)$_{8}$	294048	294055	+	2	TU 36 (I)	R.2
SSRM_91_L	trmD	tRNA (guanineN1)methyltransferase	(T)$_{8}$	574394	574401	-	13	TU 44 (I)	R
Tandem_01_T	MHP7448_0397	hypothetical protein	(TTAAATAAAAATT)$_{2}$	420099	420124	-	165	TU 68 (I)	1.R
Tandem_01_TE	rpsP	30S ribosomal protein S16	(ATAATCCGTGTCCAA)$_{2}$	574064	574095	-	31	TU 44 (I)	R.3
Tandem_03_T	MHP7448_0695	hypothetical protein	(TATTATATAATAAA)$_{2}$	429259	429286	+	200	mC 23	3.R
Tandem_06_E	MHP7448_0197	hypothetical protein	(AAAAATTTTTTTA)$_{2}$	220702	220727	+	401	TU 28 (F)	1.R

*REP_ID was determined in palindrome elements as follows: class of repeat (PAL: palindrome without gap or PALG: palindrome with gap), software that predicted the element (E: palindrome (EMBOSS), S: Palindrome Search or ES: both software predicted that repeat). Tandem repeats REP_ID was classified as: class of repeat (SSR: Simple Sequence Repeat, SSRM: Simple Sequence Repeats of Mononucleotide), aleatory number and software that predicted the element (E: etandem, I: SSRRIT, L: SSRLocator, T: TRF), combination of letters represent that more than one algorithm predicted the same element.

**Sequence: PALG: number (N) of nucleotides that form the gap. SSR, SSRM and tandem are represented by (motif)$_n$ where n represents times of repetitive motif.

***Distance in nucleotide from DNA repeat localized in the 5' upstream region to start codon (ATG) of respective gene.

***Genome organization (GOZ): Repeats classification in transcriptional units polycistronics (TU) and monocistronics (mC). In GOZ, letters in parentheses comprises the position of gene within the respective TU: first (F), internal (I) or last (L). Class 1 represents repeats found in the 5' upstream region of a transposase coding genes. Class 2 represent elements found upstream pseudogenes. Class 3 comprises repeats found upstream tRNA coding genes. Class 4 represent repeats upstream rRNA coding genes.
****Promoter organization (POZ): Repeats were classified in relation to proximal putative promoter sequences. Nomenclature of classification follows the pattern n_1.R.n_2, whereas n_1 represent number of promoters found downstream the repeat and n_2 represent number of promoters present upstream the repetitive element.

Supplementary Table S3: Tandem comparison among related mycoplasmas

Mycoplasma hyopneumoniae 7448					*Mycoplasma hyopneumoniae 7422*						*Mycoplasma hyopneumoniae J*						*Mycoplasma flocculare*					
REP ID	**Gene name**	**Product**	**Copy n°***	**ATG Distance**	**Gene name**	**Product**	**Copy n°***	**Localization**	**ATG Distance**	**GOZ**	**Gene Name**	**Product**	**Copy n°***	**Localization**	**ATG Distance**	**GOZ**	**Gene name**	**Product**	**Copy n°***	**Localization**	**ATG Distance**	**GOZ**
Tandem_01_T	MHP7448_0397	hypothetical protein	2	165	MX02869	hypothetical protein	2	c 522521..522546	165	TU 70 (F)	MHJ_0410	hypothetical protein	2	c 499379..499404	165	TU 70 (L)	MF00009	hypothetical protein	0	-	-	TU 57 (L)
Tandem_03_T	MHP7448_0695	hypothetical protein	2	200	MX02674	hypothetical protein	2	437723..437750	201	TU 56 (F)	-	-	-	-	-	-	-	-	-	-	-	
Tandem_01_TE	rpsP	30S ribosomal protein S16	2	31	rpsP	30S ribosomal protein S16	2	c 358835..358864	31	TU 44 (I)	rpsP	30S ribosomal protein S16	2	c 332479..332508	31	TU 45 (I)	rpsP	30S ribosomal protein S16	0	-	-	TU 02 (F)
Tandem_06_E	MHP7448_0197	hypothetical protein	2	401	MX02806	hypothetical protein	1	239660..239672	310	TU 29 (F)	MHJ_0193	hypothetical protein	1	213292..213304	407	TU 30 (F)	MYF_RS00545	hypothetical protein	0	-		TU 16 (L)
SSR_49_LI	MHP7448_0485	hypothetical protein	9	153	MX03418	hypothetical protein	10	630880..630990	153	mC 36	MHJ_0482	hypothetical protein	11	609819..609851	152	mC 23	MF00865	hypothetical protein	1	c 24649..24651	153	mC 31
SSR_02_LI	MHP7448_0016	hypothetical protein	6	82	MX02838	hypothetical protein	8	c 17080..17111	82	mC 01	MHJ_0016	hypothetical protein	1	c 17134..17137	107	TU 2 (F)	-	-	-	-	-	-
SSR_03_LI	MHP7448_0017	hypothetical protein	6	270	MX02685	hypothetical protein	8	17080..17111	270	TU 02 (F)	-	-	-	-	-	-		-	-	-	-	-
SSR_05_LI	MHP7448_0087	GTP-binding protein	2	458	MX02930	GTP-binding protein	2	111023..111040	458	mC 04	rbgA	ribosomal biogenesis GTPase	0	-	-	mC 02	MF01266	GTP-binding protein	0	-	-	TU 68 (F)
SSR_09_LI	MHP7448_0138	hypothetical protein	2	230	MX03045	hypothetical protein	1	196604..196613	248	TU 25 (I)	MHJ_0134	hypothetical protein	0	-	-	TU 25 (F)	-	-	-	-	-	-
SSR_13_LI	glyA	glycine hydroxymethyltransferase	2	301	glyA	glycine hydroxymethyltransferase	2	c 283605..283622	284	TU 33 (I)	glyA	glycine hydroxymethyltransferase	0	-	-	TU 34 (F)	glyA	glycine hydroxymethyltransferase	0	-	-	mC 19
SSR_14_LI	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	2	178	MX02920	methylmalonate-semialdehyde dehydrogenase	0	-	-	TU 34 (F)	MHJ_0219	methylmalonate-semialdehyde dehydrogenase	2	257215..257232	187	TU 35 (F)	-	-	-	-	-	-
SSR_15_LI	MHP7448_0225	methylmalonate-semialdehyde dehydrogenase	2	441	MX02920	methylmalonate-semialdehyde dehydrogenase	0	-	-	TU 34 (F)	MHJ_0219	methylmalonate-semialdehyde dehydrogenase	2	256955..256972	447	TU 35 (F)	-	-	-	-	-	-
SSR_19_LI	pyrG	CTP synthase	2	479	pyrG	CTP synthase	2	352147..352167	469	mC 20	pyrG	CTP synthase	2	325769..325788	490	mC 12	MYF_RS01890	CTP synthase	0	-	-	mC 18
SSR_28_LI	MHP7448_0461	hypothetical protein	6	476	MX02822	hypothetical protein	-	-	-	TU 77 (I)	MHJ_0458	hypothetical protein	2	580253..580258	128	TU 78 (F)	-	-	-	-	-	-
SSR_30_LI	MHP7448_0484	hypothetical protein	9	141	MX03238	hypothetical protein	10	c 630880..630990	76	TU 81 (F)	MHJ_0481	hypothetical protein	11	c 609819..609851	107	TU 81 (F)	MYF_RS00935	hypothetical protein	2	247462..247467	205	TU 33 (F)
SSR_41_LI	dam	DNA adenine methylase	5	401	dam	DNA adenine methylase	5	c 816153..816162	401	mC 47	dam	DNA adenine methylase	5	c 817222..817231	399	mC 33	-	-	-	-	-	-
SSR_42_LI	MHP7448_0623	ABC transporter ATP-binding - Pr1	5	243	MX03354	ABC transporter ATP-binding - Pr1	5	816153..816163	243	TU 104 (F)	MHJ_0624	putative ABC transporter ATP-binding - Pr1	5	817222..817231	240	TU 105 (F)	MYF_RS02820	ABC transporter ATP-binding - Pr1-like protein	5	833909..833918	243	TU 94 (F)
SSRM_06_TL	MHP7448_0025	hypothetical protein	21	121	MX02912	hypothetical protein	21	c 32613..32633	121	TU 04 (F)	MHJ_0021	hypothetical protein	19	c 25792..25810	221	TU 02 (F)	-	-	-	-	-	-
SSRM_07_TL	sipS	signal peptidase I	21	231	MX00115	signal peptidase I	21	32613..32633	231	TU 05 (F)	sipS	signal peptidase I	19	25792..25810	263	TU 03 (F)	MF01338	hypothetical protein	3	31..33	207	TU 17 (F)
SSRM_09_L	parE	topoisomerase IV subunit B	8	149	parE	topoisomerase IV subunit B	8	c 45151..45158	149	TU 06 (F)	parE	topoisomerase IV subunit B	7	c 38413..38419	149	TU 04 (F)	parE	topoisomerase IV subunit B	5	c 15160..15164	107	TU 19 (F)
SSRM_10_L	gap	glyceraldehyde 3phosphate dehydrogenase	8	53	gap	glyceraldehyde 3phosphate dehydrogenase	8	45151..45159	53	TU 07 (F)	gap	glyceraldehyde 3phosphate dehydrogenase	7	38413..38420	53	TU 05 (F)	gap	glyceraldehyde 3phosphate dehydrogenase	5	15151..15155	56	TU 20 (F)
SSRM_105_L	MHP7448_0312	glycine cleavage system H protein	8	358	MX01193	glycine cleavage system H protein	8	c 395168..395175	357	TU 49 (F)	MHJ_0304	glycine cleavage system H protein	7	c 367130..367136	361	TU 49 (F)	-	-	-	-	-	-
SSRM_11_TL	MHP7448_0117	hypothetical protein	16	311	MX02626	hypothetical protein	4	180476..180479	327	TU 21 (I)	MHJ_0113	hypothetical protein	21	153684..153704	311	TU 19 (F)	-	-	-	-	-	-
SSRM_116_L	MHP7448_0334	hypothetical protein	8	378	MX02928	hypothetical protein	8	419678..419685	378	TU 53 (F)	MHJ_0325	hypothetical protein	8	392192..392199	353	TU 54 (F)	-	-	-	-	-	-
SSRM_143_L	MHP7448_0399	hypothetical protein	9	354	MX03086	hypothetical protein	6	c 523890..523895	59	TU 70 (F)	MHJ_0412	hypothetical protein	5	c 500989..500990	262	TU 70 (F)	MYF_RS01585	hypothetical protein	5	c 418155..418159	130	TU 55 (F)
SSRM_144_L	MHP7448_0400	hypothetical protein	9	89	MX03384	hypothetical protein	6	523890..523895	123	TU 71 (F)	MHJ_0413	hypothetical protein	5	500986..500990	37	TU 71 (F)	MYF_RS01590	hypothetical protein	5	418155..418159	94	TU 56 (F)
SSRM_15_TL	MHP7448_0272	P97like protein	19	217	MX03225	P97like protein	18	c 344818..344835	217	TU 42 (F)	MHJ_0264	putative P97like protein	20	c318423..318442	240	TU 43 (F)	MYF_RS01805	P97-like protein	6	c 476952..476957	229	TU 64 (F)
SSRM_150_L	MHP7448_0426	hypothetical protein	8	2	MX02655	hypothetical protein	-	-	-	TU 18 (F)	-	-	-	-	-	-	-	-	-	-	-	-
SSRM_151_L	efp	elongation factor EFP	8	81	efp	elongation factor EFP	5	540272..540276	94	TU 73 (F)	efp	elongation factor EFP	5	517148..517152	94	TU 73 (F)	efp	elongation factor EFP	6	c 29717..29722	95	TU 59 (F)
SSRM_16_TL	pheS	phenylalanyltRNA synthetase alpha chain	19	264	pheS	phenylalanyltRNA synthetase alpha chain	18	344818..344835	264	TU 43 (F)	pheS	phenylalanyltRNA synthetase alpha chain	20	318423..318442	255	Tu 44 (F)	pheS	phenylalanyltRNA synthetase alpha chain	6	74364..74369	239	TU 67 (F)

SSRM_165_L	MHP7448_0453	ABC transporter permease protein	9	157	MX02986	ABC transporter permease protein	5	c 596727..596731	161	TU 75 (F)	MHJ_0450	ABC transporter permease protein	10	c 575736..5745	130	TU 75 (F)	MF01278	ABC transporter permease protein	2	c 2180..2181	151	TU 34 (F)
SSRM_179_L	MHP7448_0484	hypothetical protein	8	211	MX03238	hypothetical protein	8	c 630953..630960	149	TU 81 (F)	MHJ_0481	hypothetical protein	7	c 609895..609901	183	TU 81 (F)	MYF_RS00935	hypothetical protein	5	247458..247462	210	TU 33 (F)
SSRM_184_L	MHP7448_0485	hypothetical protein	8	102	MX03418	hypothetical protein	8	630953..630960	102	mC 36	MHJ_0482	hypothetical protein	7	609895..609901	102	mC 23	MF00865	hypothetical protein	5	c 24583..24587	86	mC 31
SSRM_195_L	MHP7448_0505	lipoprotein	8	246	MX03170	lipoprotein	8	c 665592..665599	264	TU 84 (F)	MHJ_0502	putative lipoprotein	8	c 665592..665599	264	TU 84 (F)	MF00046	lipoprotein	6	114842..114847	232	TU 85 (F)
SSRM_196_L	pdhC	dihydrolipoamide acetyltransferase	8	398	pdhC	dihydrolipoamide acetyltransferase	8	665969..665976	400	TU 85 (F)	pdhC	dihydrolipoamide acetyltransferase	8	647867..647874	389	TU 85 (F)	pdhC	dihydrolipoamide acetyltransferase	-	-	-	TU 84 (I)
SSRM_199_L	MHP7448_0512	hypothetical protein	9	318	MX03132	hypothetical protein	7	c 675187..675193	266	TU 86 (F)	MHJ_0510	hypothetical protein	8	c 656645..656652	253	TU 86 (F)	MF00857	hypothetical protein	6	20516..20521	152	TU 109 (I)
SSRM_227_L	MHP7448_0596	hypothetical protein	23	70	MX03007	hypothetical protein	16	773274..773289	72	TU 100 (F)	MHJ_0596	hypothetical protein	18	771025..771042	52	TU 102 (I)	-	-	-	-	-	-
SSRM_23_L	MHP7448_0093	hypothetical protein	8	73	MX02851	hypothetical protein	8	c 120803..12810	131	TU 15 (F)	MHJ_0090	hypothetical protein	8	c 117390..117397	102	TU 14 (F)	MYF_RS00485	hypothetical protein	16	c 107255..107270	130	TU 13 (F)
SSRM_234_L	rplJ	50S ribosomal protein L10	21	109	rplJ	50S ribosomal protein L10	19	c 808800..808818	84	TU 102 (F)	rplJ	50S ribosomal protein L10	20	c 809882..809901	157	TU 103 (F)	rplJ	50S ribosomal protein L10	7	53886..53892	155	TU 105 (F)
SSRM_235_L	rplJ	50S ribosomal protein L10	23	84	rplJ	50S ribosomal protein L10	16	c 808822..808837	106	TU 102 (F)	rplJ	50S ribosomal protein L10	19	c 809861..809879	136	TU 103 (F)	rplJ	50S ribosomal protein L10	7	53906..53912	135	TU 105 (F)
SSRM_236_L	MHP7448_0620	hypothetical protein	21	245	MX03025	hypothetical protein	16	808822..808837	246	TU 103 (F)	MHJ_0621	hypothetical protein	20	809882..809901	211	TU 104 (F)	MYF_RS02575	pseudogene	6	683498..683503	239	TU 86 (F)
SSRM_237_L	MHP7448_0620	hypothetical protein	23	268	MX03025	hypothetical protein	19	808800..808818	265	TU 103 (F)	MHJ_0621	hypothetical protein	19	809861..809879	233	TU 104 (F)	MYF_RS02805	hypothetical protein	5	683487..683491	251	TU 93 (I)
SSRM_240_L	MHP7448_0623	ABC transporter ATPbinding Pr1	8	373	MX03354	ABC transporter ATPbinding Pr1	8	816025..816032	373	TU 104 (F)	MHJ_0624	ABC transporter ATPbinding Pr1	8	817094..817101	370	TU 105 (F)	MYF_RS02820	ABC transporter ATP-binding - Pr1-like protein	4	689468..689471	295	TU 94 (F)
SSRM_25_L	MHP7448_0095	hypothetical protein	9	261	MX02858	hypothetical protein	9	c 123266..123274	261	mC 07	MHJ_0092	hypothetical protein	7	c 119650..119656	58	mC 04	MYF_RS00495	hypothetical protein	7	c 109368..109374	82	mC 05
SSRM_256_L	MHP7448_0094	hypothetical protein	8	18	MX03000	hypothetical protein	8	120803..12811	18	mC 06	MHJ_0091	hypothetical protein	8	117390..117398	0	mC 03	MYF_RS00490	hypothetical protein	16	107255..107270	18	mC 04
SSRM_26_L	tpx	thiol peroxidase	9	100	tpx	thiol peroxidase	9	123266..123274	100	TU 16 (F)	tpx	thiol peroxidase	10	119853..119862	97	TU 15 (I)	tpx	thiol peroxidase	6	11569..11574	114	TU 70 (F)
SSRM_264_L	MHP7448_0699	hypothetical protein	8	82	-	-	-	-	-	-	-	-	-	-			-	-	-	-	-	-
SSRM_265_L	MHP7448_0700	hypothetical protein	8	2	MX02857	hypothetical protein	8	c 149280..149287	2	mC 12	-	-	-	-			-	-	-	-	-	-
SSRM_266_L	MHP7448_0714	hypothetical protein	8	356	-	-	-	-	-	-	-	-	-	-			MYF_RS00655	hypothetical protein	5	c 154409..154413	306	TU 20 (F)
SSRM_27_TL	MHP7448_0426	hypothetical protein	15	30	MX02655	hypothetical protein	-	-	-	TU 18 (F)	-	-	-	-	-	-	-	-	-	-	-	-
SSRM_28_TL	efp	elongation factor EFP	15	46	efp	elongation factor EFP	20	540305..540324	46	TU 73 (F)	efp	elongation factor EFP	20	517181..517200	46	TU 73 (F)	efp	elongation factor EFP	6	c 29681..29686	59	TU 59 (F)
SSRM_29_L	tpiA	triosephosphate isomerase	8	92	tpiA	triosephosphate isomerase	3	135514..135516	21	TU 17 (I)	tpiA	triosephosphate isomerase	6	131654..131660	46	TU 16 (F)	tpiA	triosephosphate isomerase	4	22767..22770	31	TU 71 (I)
SSRM_30_TL	MHP7448_0461	hypothetical protein	23	123	MX02822	hypothetical protein	-	-	-	TU 77 (F)	MHJ_0458	hypothetical protein	22	580211..580232	167	TU 78 (F)	-	-	-	-	-	-
SSRM_62_L	MHP7448_0219	hypothetical protein	8	115	MX02884	hypothetical protein	8	c 276798..276805	190	TU 33 (L)	MHJ_0214	hypothetical protein	8	c 250425..250432	94	TU 34 (L)	-	-	-	-	-	-
SSRM_65_L	glyA	glycine hydroxymethyltransferase	8	50	glyA	glycine hydroxymethyltransferase	6	c 283386..283391	65	TU 33 (F)	glyA	glycine hydroxymethyltransferase	7	c 256967..256973	52	TU 34 (F)	glyA	glycine hydroxymethyltransferase	6	505707..505712	52	mC 19
SSRM_66_L	glyA	glycine hydroxymethyltransferase	8	337	glyA	glycine hydroxymethyltransferase	8	c 283641..283648	320	TU 33 (F)	glyA	glycine hydroxymethyltransferase	8	c 257260..257267	345	TU 34 (F)	glyA	glycine hydroxymethyltransferase	7	505467..505473	291	mC 19
SSRM_67_L	glyA	glycine hydroxymethyltransferase	8	375	glyA	glycine hydroxymethyltransferase	8	c 283679..283686	358	TU 33 (F)	glyA	glycine hydroxymethyltransferase	8	c 257298..257305	383	TU 34 (F)	glyA	glycine hydroxymethyltransferase	4	505456..505459	305	mC 19
SSRM_68_L	glyA	glycine hydroxymethyltransferase	8	456	glyA	glycine hydroxymethyltransferase	8	c 283760..283767	439	TU 33 (F)	glyA	glycine hydroxymethyltransferase	8	c 257379..257386	464	TU 34 (F)	glyA	glycine hydroxymethyltransferase	5	505352..505356	408	mC 19
SSRM_69_L	glyA	glycine hydroxymethyltransferase	10	194	glyA	glycine hydroxymethyltransferase	10	c 283497..283506	176	TU 33 (F)	glyA	glycine hydroxymethyltransferase	8	c 257110..257117	195	TU 34 (F)	glyA	glycine hydroxymethyltransferase	4	505510..505513	244	mC 19
SSRM_70_L	MHP7448_0225	methylmalonatesemialdehyde dehydrogenase	8	152	MX02920	methylmalonatesemialdehyde dehydrogenase	8	283641..283648	152	TU 34 (F)	MHJ_0219	methylmalonatesemialdehyde dehydrogenase	8	257260..257267	152	TU 35 (F)	-	-	-	-	-	-
SSRM_74_L	MHP7448_0225	methylmalonatesemialdehyde dehydrogenase	10	293	MX02920	methylmalonatesemialdehyde dehydrogenase	10	283497..283506	294	TU 34 (F)	MHJ_0219	methylmalonatesemialdehyde dehydrogenase	8	257110..257117	302	TU 35 (F)	-	-	-	-	-	-

*Copy n° comprises number of times that a repetitive element is present in the ortholog of analyzed genome. REP_ID, ATG distance and GOZ were classified as Supplementary Table S2.
Hairline (-) represent that no ortholog was found in respective genome, or no intergenic region was available for analysis.

Supplementary Table S4: Palindrome elements in adhesins coding genes comparison among related mycoplasmas.

MHP_7448					**MHP_7422**				**MHP_J**				**MFL**			
Gene	**Adhesin**	**REP_ID**	**Location Repeat**	**N° elements**	**Repeat[1]**	**Location Repeat[2]**	**N° elements**	**Conservation[3]**	**Repeat[1]**	**Location Repeat[2]**	**N° elements**	**Conservation[3]**	**Repeat[1]**	**Location Repeat[2]**	**N° elements**	**Conservation[3]**
MHP7448_0361 MHP7448_0362	P29 P69	PALG_E_472	TU 63	5	*	TU/MX02948	5	100%	*	TU/MHJ_0356	5	100%	x	TU/MYF_01405	3	60%
		PALG_E_473			*				*				*			
		PALG_E_474			*				*				*			
		PALG_E_975			*				*				x			
		PALG_ES_134			*				*				*			
MHP7448_0497	P76	PALG_E_1055	GENE	5	*	MX03203	5	100%	x	MHJ_0494	3	60%	x	MYF_00955	1	20%
		PALG_E_1056			*				x				x			
		PALG_E_572			*				*				*			
		PALG_E_573			*				*				x			
		PALG_ES_183			*				*				x			
MHP7448_0198	P97	PALG_E_338	GENE	6	*	MX03220	6	100%	x	MHJ_0194	3	50%	-	-	-	-
		PALG_E_339			*				x							
		PALG_E_895			*				x							
		PALG_E_896			*				*							
		PALG_S_1195			*				*							
		PALG_S_1197			*				*							
MHP7448_0108	P97	PALG_E_291	GENE	3	*	MX03142	3	100%	x	MHJ_0105	2	67%	x	MYF_00575	1	33%
		PALG_E_859			*				*				*			
		PALG_S_1184			*				*				x			
MHP7448_0272	P97	PALG_E_381	GENE	6	*	MX03225	5	83%	*	MHJ_0264	5	67%	*	MYF_01935	4	67%
		PALG_E_382			*				*				x			
		PALG_E_383			*				*				*			
		PALG_E_385			x				x				*			
		PALG_E_925			*				*				*			
		PALG_E_926			*				x				x			
MHP7448_0199	P102	PALG_E_336	TU 28	5	*	TU/MX02806	4	80%	*	TU/MHJ_0193	5	100%	x	TU/MYF_00605	1	20%
		PALG_E_337			x				*				x			
		PALG_E_894			*				*				x			
		PALG_ES_84			*				*				*			
		PALG_ES_85			*				*				x			
MHP7448_0107	P102	PALG_E_863	TU 16	1	*	TU/MX00499	1	100%	*	TU/MHJ_0107	1	100%	x	TU/MYF_00605	0	0%
MHP7448_0271	P102	PALG_E_924	GENE	1	*	TU/MX03225	1	100%	x	MHJ_0263	0	0%	x	TU/MYF_01935	0	0%

MHP7448_0105	P102	PALG_E_285	TU 15	3	*	MX03599	3	100%	*	MHJ_0102	1	25%	*	TU/MYF_00605	3	
		PALG_E_286			*				x				*			
		PALG_E_287			*				x				*			
MHP7448_0663	P146	PALG_E_1145	GENE/ TU 117	7	*	MX02965	7	100%	*	MHJ_0663	6	86%	x	MYF_03270	4	57%
		PALG_E_1146			*				*				*			
		PALG_E_1147			*				*				*			
		PALG_E_1148			*				*				*			
		PALG_E_708			*				*				*			
		PALG_E_709			*				x				x			
		PALG_ES_229			*				*				x			
MHP7448_0496	P216	PALG_E_570	GENE	2	*	3210	2	100%	x	MHJ_0493	1	50%	x	MYF_00955	1	100%
		PALG_E_571			*				*				*			
MHP7448_0373	LppS	PALG_E_478	GENE	2	x	MX03119	1	50%	x	MHJ_0369	0	0%	x	MYF_01525	1	50%
		PALG_S_1241			*				x				*			
MHP7448_0372	LppT	PALG_E_482	TU 64	5	x	TU/MX03119	1	20%	x	TU/MHJ_0374	2	40%	x	TU/MYF_01490	2	40%
		PALG_E_987			*				*				*			
		PALG_E_988			x				*				*			
		PALG_E_989			x				x				x			
		PALG_ES_138			x				x				*			
		PALG_E_237	TU 01	5	*	TU/MX03508	5	100%	*	TU/MHJ_0001	5	100%	*	TU/MYF_0005	3	60%
		PALG_E_238			*				*				*			
MHP7448_0006	MgPa	PALG_E_239			*				*				x			
MHP7448_0005	MgPa	PALG_E_240			*	MX02961			*				*			
		PAL_ES_241			*				*				x			

Comparison of palindrome among *M. hyopneumoniae* strain 7448 (MHP_7448), *M. hyopneumoniae* strain 7422 (MHP_7422), *M. hyopneumoniae* strain J (MHP_J) and *M. flocculare* (MFL) orthologs. [1]Asterisks (*) represent the presence of repetitive element and (x) comprises the absence of element. [2]Repeats could be found directly upstream CDS (represented by gene_name) or upstream the first gene of respective TU (represented by TU/name of first gene of TU). [3] Percent of conservation was done by diving number of palindrome elements found in ortholog gene by number of palindromes found in MHP_7448.

Supplementary Table S5: Presence of repeat elements on upstream regions of regulated genes in *M. hyopneumoniae* strain 232

MHP_232		**MHP_7448**	**MHP_7448 x MHP_232**	
Gene_name	**Product**	**Gene_name**	**Conservation**[1]	**ΔG**[2]
mhp004	hypothetical protein	MHP7448_004	5/5 (100%)	-8.06
mhp024	hypothetical protein	MHP7448_022	2/2 (100%)	-2.14
atpB	ATP synthase A chain	atpB	2/2 (100%)	-5.61
glyS	glycyl-tRNA synthetase	glyS		
mhp232	hypothetical protein	MHP7448_0685	0/1 (0%)	*
spoU	rRNA methylase	spoU	0/1 (0%)	*
mglA	ATP-binding protein	mglA	2/2 (100%)	-6.32
mhp325	hypothetical protein	MHP7448_0312	4/5 (80%)	-2.61
mhp326	hypothetical protein	MHP7448_0313	5/5 (100%)	-4.34
mhp355	hypothetical protein	MHP7448_0344	4/4 (100%)	-4.85
mhp366	hypothetical protein	MHP7448_0355	2/2 (100%)	-3.33
mhp516	hypothetical protein	MHP7448_0518	1/4 (25%)	-3.13
mhp623	hypothetical protein	MHP7448_0604	4/5 (80%)	-4.18
ftsY	cell division protein	ftsY	2/2 (100%)	-5.09
gatA	glutamyl-tRNA amidotransferase subunit A	gatA	0/5 (0%)	*
mhp115	ABC transporter ATP binding protein	MHP7448_0272	5/7 (71%)	-3.59
rpl4	50s ribosomal protein L4	rplD	6/6 (100%)	-3.82
rpl23	50s ribosomal protein L23	rplW		
mhp262	hypothetical protein	MHP7448_0118	4/4 (100%)	-6.09
pfkA	6-phosphofructokinase	pfkA	1/1 (100%)	-2.90
mhp293	hypothetical protein	MHP7448_0331	0/3 (0%)	*
mhp424	hypothetical protein	MHP7448_408	2/2 (100%)	-2.34
mhp451	hypothetical protein	MHP7448_452	3/3 (100%)	-6.91

mhp452	hypothetical protein	MHP7448_453	0/3 (0%)	*
napA	neutrophil activating protein	napA	3/3 (100%)	-4.83
gyrA	DNA gyrase subunit A	gyrA	6/6 (100%)	-3.79
secA	preprotein translocase subunit	MHP7448_0086	4/5 (80%)	-7.33
mhp019	hypothetical protein	MHP7448_017	0/3 (0%)	*
mhp027	hypothetical protein	MHP7448_025	4/4 (100%)	-3.08
mhp324	hypothetical protein	MHP7448_0311	4/4 (100%)	-2.55
aprE	subtilisin-like serine protease	MHP7448_0332	2/4 (50%)	-4.48
mhp354	hypothetical protein	MHP7448_0342	6/7 (86%)	-9.34
mhp360	hypothetical protein	MHP7448_0349	4/4 (100%)	-2.76
gap	glyceraldehyde 3-phosphate dehydrogenase	MHP7448_035	3/3 (100%)	-4.70
mhp630	hypothetical protein	MHP7448_0611	4/4 (100%)	-4.03
rpsG	30s ribosomal protein S7	MHP7448_076	1/1 (100%)	-2.30
rps17	30s ribosomal protein S17	MHP7448_185	6/6 (100%)	-3.82
mhp093	hypothetical protein	MHP7448_287	5/6 (83%)	-7.85
yx1	hypothetical protein	MHP7448_498	5/6 (83%)	-5.62
mhp508	hypothetical protein	MHP7448_511	2/2 (100%)	-0.67
mhp561	hypothetical protein	MHP7448_544	1/4 (25%)	-2.74
mhp662	hypothetical protein	MHP7448_642	3/4 (75%)	-7.42
mhp371	putative high affinity transport system protein p37 precursor	MHP7448_0360	3/5 (60%)	-0.87
potA	ABC transporter ATP-binding protein	MHP7448_0369		
ugpA	sn-glycerol-3-phosphate transport system permease	MHP7448_0370		
fusA	GTP-binding protein chain elongation factor ef-g	fusA	4/5 (80%)	-1.12
trsE	transfer complex protein	MHP7448_0247	5/5 (100%)	-10.43
mhp118	hypothetical protein	MHP7448_0261	2/2 (100%)	-2.99
mhp395	hypothetical protein	MHP7448_0383	1/5 (20%)	-3.59

mhp629	hypothetical protein	MHP7448_0610	3/3 (100%)	-7.07
mhp092	hypothetical protein	MHP7448_340	1/4 (25%)	-2.12
oppD	oligopeptide transport system permease protein	MHP7448_502	5/6 (83%)	-5.62
pgiB	phosphoglucose isomerase B	MHP7448_531	1/2 (50%)	-2.03
pyrH	uridylate kinase	MHP7448_535	2/2 (100%)	-6.29
mhp576	hypothetical protein	MHP7448_556	0/3 (0%)	*
mhp654	hypothetical protein	MHP7448_633	7/7 (100%)	-4.31
p115	p115 protein	MHP7448_657	1/1 (100%)	-3.88
nrdE	ribonucleoside-diphosphate reductase alpha chain	nrdE	2/5 (40%)	-2.23
mhp130	putative DNA processing protein	smf	5/5 (100%)	-10.43
truB	tRNA pseudouridine synthase B	truB	1/1 (100%)	-3.26
mhp170	hypothetical protein	MHP7448_0211	2/3 (67%)	-1.70
oppF	oligopeptide transport system permease protein	oppF-1		
rpl5	50s ribosomal protein L5	MHP7448_0182	6/6 (100%)	-3.82
rps8	30s ribosomal protein S8	rpsH		
adk	adenylate kinase	adk		
mhp247	hypothetical protein	MHP7448_0135	1/1 (100%)	-0.62
mhp273	hypothetical protein	MHP7448_0106	1/3 (33%)	-1.91
mhp303	hypothetical protein	MHP7448_0342	2/7 (29%)	-0.89
gtp1	GTP-binding protein	ychF	2/2 (100%)	-4.04
rps18	30s ribosomal protein S18	rpsR		
mhp517	hypothetical protein	MHP7448_0519	1/1 (100%)	-2.02
mhp309	hypothetical protein	MHP7448_297	0/1 (0%)	*
mhp312	hypothetical protein	MHP7448_299	2/2 (100%)	-3.28
mhp341	hypothetical protein	MHP7448_330	1/1 (100%)	-0.70
mhp362	hypothetical protein	MHP7448_351	3/3 (100%)	-1.39

rluD	ribosomal large subunit pseudouridine synthase D	MHP7448_458	4/4 (100%)	-4.83
mhp472	hypothetical protein	MHP7448_474	4/4 (100%)	-6.68
mhp480	hypothetical protein	MHP7448_482	6/6 (100%)	-3.87
mhp583	hypothetical protein	MHP7448_566	1/4 (25%)	-3.61
mhp634	hypothetical protein	MHP7448_0615	4/4 (100%)	-4.03
rplL	50s ribosomal protein L7/l12	MHP7448_0618		
rpoC	DNA-directed RNA polymerase beta chain	MHP7448_0616	1/1 (100%)	-1.72
cysS	cysteinyl-tRNA synthetase	MHP7448_641	3/4 (75%)	-7.42
mhp670	hypothetical protein	MHP7448_649	1/4 (25%)	-7.32
mhp140	hypothetical protein	MHP7448_0240	6/6 (100%)	-2.49
glpQ	glycerophosphoryl diester phosphodiesterase	ugpQ	2/2 (100%)	-3.28
ackA	acetate kinase	ackA	4/4 (100%)	-4.92
mhp510	hypothetical protein	MHP7448_0512		
atpG	ATP synthase gamma chain	atpG	2/2 (100%)	-5.61
tsf	elongation factor ts	tsf		
dnaN	DNA polymerase III, beta chain	dnaN	2/2 (100%)	-3.59
hpt	hypoxanthine phosphoribosyl transferase	hpt	2/2 (100%)	-2.99
mhp006	hypothetical protein	MHP7448_006	5/5 (100%)	-8.06
mhp022	ABC transporter ATP binding protein	MHP7448_020	1/1 (100%)	-1.32
mhp639	hypothetical protein	MHP7448_0620	4/4 (100%)	-4.32
ushA	5'-nucleotidase	MHP7448_0630	7/7 (100%)	-6.04
mhp666	hypothetical protein	MHP7448_0646	3/4 (75%)	-7.42
mhp087	hypothetical protein	MHP7448_079	0/4 (0%)	*
mhp411	hypothetical protein	MHP7448_397	2/2 (100%)	-5.40
metK	S-adenosylmethionine synthetase	MHP7448_451	1/1 (100%)	-9.21
nrdF	ribonucleoside-diphosphate reductase beta chain	nrdF	2/4 (50%)	-4.53

rpl2	50s ribosomal protein L2	rplB	1/1 (100%)	-3.77
rpsL	30s ribosomal protein S12	rpsN	2/2 (100%)	-3.26
smpB	SsrA-binding protein	smpB	2/2 (100%)	-1.16
topA	DNA topoisomerase I (omega-protein)	topA	5/6 (83%)	-7.85
rps19	30s ribosomal protein S19	rpsS	6/6 (100%)	-3.82
rpl14	50s ribosomal protein L14	rplN		
rpl6	50s ribosomal protein L6	rplF		
rps5	30s ribosomal protein S5	rpsE		
rpoA	DNA-directed RNA polymerase alpha chain	rpoA		
mhp377	putative lipoprotein	MHP7448_0366	4/5 (80%)	-4.52
mhp401	hypothetical protein	MHP7448_389	1/1 (100%)	-0.34
mhp429	hypothetical protein	MHP7448_411	4/7 (57%)	-5.47
mhp434	hypothetical protein	MHP7448_431	3/5 (60%)	-2.74
atpA	ATP synthase alpha chain	atpA	6/6 (100%)	-3.87
mhp478	hypothetical protein	MHP7448_480		
mhp481	hypothetical protein	MHP7448_483		
mhp482	hypothetical protein	MHP7448_484	6/6 (100%)	-3.87
mhp284	hypothetical protein	MHP7448_0095	1/4 (25%)	-5.86
mhp236	chromate transport protein, putative	MHP7448_0146	1/4 (25%)	-2.03
mhp350	hypothetical protein	MHP7448_0338	1/4 (25%)	+0.20
mhp077	hypothetical protein	MHP7448_071	1/1 (100%)	-0.32
p146	p146 adhesin like-protein, p97 paralog	MHP7448_663	7/7 (100%)	-4.79
pepP	xaa-pro aminopeptidase	pepP	2/2 (100%)	-3.24
polC	DNA polymerase III, alpha chain	polC	2/2 (100%)	-6.29
mhp405	hypothetical protein	MHP7448_391	7/7 (100%)	-3.98
mhp406	hypothetical protein	MHP7448_392		

ftsZ	cell division protein	MHP7448_393		
mhp461	hypothetical protein	MHP7448_463	5/5 (100%)	-3.83
ktrA	potassium uptake protein	MHP7448_545	5/5 (100%)	-3.89
nagE	pts system, n-acetylglucosamine-specific enzyme II, ABC component	MHP7448_574	0/1 (0%)	-1.76
pdhD	dihydrolipoamide dehydrogenase	pdhD	5/5 (100%)	-3.14
parC	topoisomerase IV subunit A	parC	3/3 (100%)	-4.40
lspA	lipoprotein signal peptidase (prolipoprotein signal peptidase) (signal peptidase II) (spase II)	lps		
mhp057	hypothetical protein	MHP7448_054	4/4 (100%)	-5.84
mhp064	hypothetical protein	MHP7448_061	2/2 (100%)	-5.61
mhp069	hypothetical protein	MHP7448_065	0/1 (0%)	*
mhp071	hypothetical protein	MHP7448_0681	2/2 (100%)	-3.33
lepA	30 kda GTP-binding protein lepA	MHP7448_072		
dnaK	chaperone protein dnaK	dnaK	3/2 (100%)	-4.74
serS	seryl-tRNA synthetase	serS	6/6 (100%)	-10.73
mhp144	putative dehydrogenase	MHP7448_0235	6/6 (100%)	-5.57
mhp145	putative D-ribose-binding protein mutant	rbsB		
mhp148	hypothetical protein	iolE		
iolD	myo-inositol catabolism	iolD		
mhp150	hypothetical protein	MHP7448_0228		
mhp151	hypothetical protein	iolB		
iolC	myo-inositol catabolism	iolC		
rbsA	ribose transport ATP-binding protein	mglA	1/2 (50%)	+0.17
mhp171	ABC transporter ATP binding protein	MHP7448_0210	3/4 (75%)	-2.35
map	methionine amino peptidase	map	6/6 (100%)	-3.82
rps11	30s ribosomal protein S11	rpsK		
rpl17	50s ribosomal protein L17	rplQ		

clpB	ATP-dependent serine proteinase, 24 protein	clpB	4/4 (100%)	-3.90
glpF	glycerol uptake facilitator protein	glpF	1/1 (100%)	-1.91
glpK	glycerol kinase (ATP:glycerol 3-phosphotransferase) (glycerokinase)	glpK	3/4 (75%)	-1.84
mhp289	hypothetical protein	MHP7448_0090	5/5 (100%)	-2.19
mhp275	hypothetical protein, p102 paralog	MHP7448_0105 (P)	0/1 (0%)	*
mhp271	p97 cilium adhesin paralog	MHP7448_0108	3/3 (100%)	-4.65
mhp254	hypothetical protein	MHP7448_0127	1/1 (100%)	-1.39
mhp334	hypothetical protein	MHP7448_0319	5/5 (100%)	-4.34
mhp374	hypothetical protein	MHP7448_0363	1/1 (100%)	-0.25
mhp399	hypothetical protein	MHP7448_386	4/4 (100%)	-6.17
mhp445	hypothetical protein	MHP7448_443	4/5 (80%)	-7.93
mhp479	hypothetical protein	MHP7448_481	6/6 (100%)	-3.87
asnS	asparaginyl-tRNA synthetase	MHP7448_500	2/3 (67%)	-3.02
rplK	50s ribosomal protein L11	rplK	2/2 (100%)	-3.24
rps13	30s ribosomal protein S13	rpsM	2/3 (67%)	-10.66
tpiA	triosephosphate isomerase	tpiA	2/5 (40%)	-3.73
oppC	oligopeptide transport system permease protein	oppC	5/6 (83%)	-5.62
oppB	oligopeptide transport system permease protein	oppB		
P46	46kd surface antigen precursor	MHP7448_0513	5/5 (100%)	-5.49
mhp005	hypothetical protein	MHP7448_005	5/5 (100%)	-8.06
mhp013	hypothetical protein	MHP7448_013		
rluC	ribosomal large subunit pseudouridine synthase	rluC		
acpD	acyl carrier protein phosphodiesterase	acpD-1	4/4 (100%)	-4.83
gatB	glutamyl-tRNA amidotransferase, subunit B	gatB	0/5 (0%)	*
metG	methionyl-tRNA synthetase	metS	7/7 (100%)	-3.98
mhp337	hypothetical protein	MHP7448_0089	2/2 (100%)	-4.18

mhp240	hypothetical protein	MHP7448_0142	0/1 (0%)	-
mhp235	hypothetical protein	MHP7448_0147	1/4 (25%)	-2.03
mhp023	ABC transporter ATP binding protein	MHP7448_021	4/4 (100%)	-3.08
mhp136	hypothetical protein	MHP7448_0244	1/1 (100%)	-3.71
mod	site-specific DNA-methyltransferase (adenine-specific)	MHP7448_0316	0/1 (0%)	-3.11
mhp353	hypothetical protein	MHP7448_0341	1/4 (25%)	-2.12
mhp373	putative transport system permease protein p69	MHP7448_0362	3/5 (60%)	-0.87
mhp393	hypothetical protein	MHP7448_0381	3/5 (60%)	-3.59
mhp310	hypothetical protein	MHP7448_297	0/1 (0%)	*
mhp419	hypothetical protein	MHP7448_404	2/6 (33%)	-1.14
proS	prolyl aminoacyl-tRNA synthetase	proS	2/5 (40%)	-2.12
rpsT	30s ribosomal protein S20	rpsT	2/2 (100%)	-1.39
mhp038	hypothetical protein	VacB	2/2 (100%)	-5.91
mhp465	hypothetical protein	MHP7448_467	4/4 (100%)	-6.01
mhp466	hypothetical protein	MHP7448_468		
bcrA	ABC transporter ATP binding protein	MHP7448_469		
mhp468	hypothetical protein	MHP7448_470		
mgtE	MG2+ ion transporter	mgtE	6/6 (100%)	-2.93
mhp509	hypothetical protein	MHP7448_0713	4/4 (100%)	-4.92
mhp474	hypothetical protein	MHP7448_476	2/2 (100%)	-2.82
mhp486	hypothetical protein	MHP7448_488	4/4 (100%)	-7.02
mhp584	hypothetical protein	MHP7448_567	1/4 (25%)	-0.90
polA	DNA polymerase I	MHP7448_581	6/6 (100%)	-7.98
mhp619	hypothetical protein	MHP7448_0600	0/1 (0%)	*
mhp620	hypothetical protein	MHP7448_0601		
mhp621	hypothetical protein	MHP7448_0602		

mhp674	hypothetical protein	MHP7448_653	2/3 (67%)	-3.24

[1] Conservation was determined by dividing number of elements found in MHP_232 by number of elements in MHP_7448. [2] ΔG represents the best values found among all palindrome elements conserved between MHP_7448 and MHP_232. Abbreviations: MHP_232 – *M. hyopneumoniae* strain 232; MHP_7448 – *M. hyopneumoniae* strain 7448.

Supplementary Table S6: Palindrome elements comparison analysis in regulated genes of *M. hyopneumoniae* strain 232

Gene_name (MHP_232)	**Product**	**Location Repeat**	**Repeat**	**N° of elments**	**Conservation**	**Oxidative[1]**	**Heat Shock[1]**	**Infection[1]**	**Iron[1]**	**Norepinephrine[1]**	**Ortholog in MHP_7448**	**Location Repeat**	**Repeat**	**N° of elements**	**ΔG**
mhp004	hypothetical protein	TU/ mhp001	*	5	100%	x					MHP7448_004	TU 01	PALG_E_237	5	-1.65
			*										PALG_E_238		-8.06
			*										PALG_E_239		-4.11
			*										PALG_E_240		-0.64
			*										PAL_ES_241		-0.61
mhp024	hypothetical protein	GENE	*	2	100%	x					MHP7448_022	Gene	PALG_ES_48	2	-2.14
			*										PALG_E_817		-1.85
atpB	ATP synthase A chain	TU/ mhp42	*	2	100%	x					atpB	TU 06	PALG_E_253	2	-0.37
glyS	glycyl-tRNA synthetase		*								glyS		PALG_E_827		-5.61
mhp232	hypothetical protein	GENE	x	0	0%	x					MHP7448_0685	Gene	PALG_E_735	1	-0.94
spoU	rRNA methylase	GENE	x	0	0%	x					spoU	TU 23	PALG_E_323	1	-5.20
mglA	ATP-binding protein	GENE	*	2	100%	x			x	x	mglA	Gene	PALG_E_412	2	-3.07
			*										PALG_S_1223		-6.32
mhp325	hypothetical protein	TU/ mhp325	*	4	80%	x				x	MHP7448_0312	Gene/ TU 50	PALG_E_415	5	+0.55
			*										PALG_S_1224		-2.55
			*										PALG_ES_102		-2.24
			x										PALG_ES_101		-1.81
			*										PALG_ES_103		-2.61
mhp326	hypothetical protein	TU/ mhp326	*	5	100%	x				x	MHP7448_0313	Gene/ TU 51	PALG_ES_104	5	-1.81
			*										PALG_ES_105		-1.85
			*										PALG_E_416		-2.57
			*										PALG_E_417		-2.49
			*										PALG_S_1225		-4.34
mhp355	hypothetical protein	GENE	*	4	100%	x					MHP7448_0344	Gene/ TU 57	PALG_E_455	4	-1.06
			*										PALG_E_456		-2.12
			*										PALG_E_457		+0.20
			*										PALG_ES_125		-4.85
mhp366	hypothetical protein	GENE	*	2	100%	x		x			MHP7448_0355	Gene	PALG_E_972	2	-3.33
			*										PALG_E_973		-0.23
mhp516	hypothetical protein	TU/ mhp518	x	1	25%	x					MHP7448_0518	TU 90	PALG_E_1067	4	-1.13
			x										PALG_E_593		+0.16
			*										PALG_E_594		-3.13
			x										PALG_ES_192		-1.80
mhp623	hypothetical protein	GENE	x	4	80%	x					MHP7448_0604	Gene	PAL_ES_420	**5**	+0.08
			*										PALG_E_659		-2.72
			*										PALG_E_1106		-3.02
			*										PALG_E_1107		-1.49
			*										PALG_E_1108		-4.18
ftsY	cell division protein	GENE	*	2	100%	x				x	ftsY	Gene	PALG_E_241	2	-2.98
			*										PALG_S_1158		-5.09
gatA	glutamyl-tRNA amidotransferase subunit A	TU/ mhp29	x	0	0%	x		x			gatA	TU 03	PALG_E_245	5	-2.19
			x										PALG_E_820		-1.21
			x										PALG_ES_51		-0.54
			x										PALG_ES_52		-1.33
			x										PALG_S_1162		-2.65

mhp115	ABC transporter ATP binding protein	TU/ mhp123	*	5	71%	x			MHP7448_0272 (P)	Gene/ TU 42	PALG_E_925	7	-3.59
			x								PALG_E_926		-2.98
			*								PALG_E_381		+0.19
			*								PALG_E_382		-2.13
			*								PALG_E_383		-1.27
			*								PALG_E_385		-3.13
			x								PAL_S_540		-0.72
rpl4 rpl23	50s ribosomal protein L4 50s ribosomal protein L23	TU/ mhp186	*	6	100%	x		x	rplD rplW	TU 27	PALG_E_334	6	-2.12
			*								PALG_E_333		-0.72
			*								PALG_E_892		-3.22
			*								PALG_E_893		-3.82
			*								PALG_ES_83		-0.28
			*								PAL_S_519		-2.62
mhp262	hypothetical protein	TU/ mhp263	*	4	100%	x			MHP7448_0118	TU 18	PALG_E_303	4	-0.45
			*								PALG_E_304		-0.78
			*								PALG_E_866		-6.09
			*								PALG_E_867		-2.10
pfkA	6-phosphofructokinase	GENE	*	1	100%	x			pfkA	Gene	PALG_E_863	1	-2.90
mhp293	hypothetical protein	-	x	0	0%	x			MHP7448_0331	Gene	PALG_ES_115	**3**	-0.64
			x								PALG_E_956		-3.24
			x								PALG_E_434		-0.67
mhp383	hypothetical protein	-	x	-	-	x			-	-	-	-	-
mhp403	hypothetical protein	-	x	-	-	x			-	-	-	-	-
mhp424	hypothetical protein	GENE	*	2	100%	x			MHP7448_408	Gene	PALG_E_503	2	-2.34
			*								PALG_E_1004		-1.41
mhp451	hypothetical protein	GENE	*	3	100%	x		x	MHP7448_452 (P)	Gene	PALG_S_1271	3	-5.23
			*								PALG_E_1027		-6.91
			*								PALG_E_1028		-1.07
mhp452	hypothetical protein	-	x	0	0%	x		x	MHP7448_453	Gene/ TU 77	PALG_E_1029	3	-2.34
			x								PALG_ES_165		-2.34
			x								PALG_E_1030		-4.84
napA	neutrophil activating protein	TU/ mhp453	*	3	100%	x			napA	TU 78	PALG_E_1032	3	-3.68
			*								PALG_E_534		-4.83
			*								PALG_E_535		-2.99
yx1	hypothetical protein	TU/ mhp502	x	5	83%	x	x		MHP7448_498	TU 86	PALG_E_1060	6	-12.61
			*								PALG_E_1061		-5.62
			*								PALG_E_576		-0.90
			*								PALG_E_577		-1.22
			*								PALG_E_578		-0.56
			*								PALG_E_579		-0.88
mhp508	hypothetical protein	GENE	*	2	100%	x			MHP7448_511 (P)	Gene	PALG_ES_185	**2**	-0.67
			*								PALG_E_584		+0.83
gyrA	DNA gyrase subunit A	GENE	*	6	100%	x			gyrA	Gene /mC31	PALG_E_1074	6	-2.85
			*								PALG_E_1075		-3.79
			*								PALG_E_603		-1.69
			*								PALG_E_604		-1.44
			*								PALG_E_605		-1.05
			*								PALG_E_606		-2.38
mhp561	hypothetical protein	TU/	*	1	25%	x			MHP7448_544 (P)	TU 95	PALG_E_781	4	-2.74

		mhp561	x							PALG_E_782		-1.67
			x							PALG_E_783		-0.50
			x							PALG_S_1385		-5.46
mhp630	hypothetical protein	TU/ mhp107	*	4	100%	x		MHP7448_0611 (P)	TU 107	PALG_E_1114	4	-1.33
			*							PALG_E_1115		-4.03
			*							PALG_E_673		-2.04
			*							PALG_ES_218		-2.91
mhp662	hypothetical protein	TU/ mhp668	*	3	75%	x		MHP7448_642	TU 113	PALG_E_1135	4	-7.42
			*							PALG_E_1136		-5.71
			*							PALG_E_696		-2.88
			x							PALG_E_697		-1.43
mhp019	hypothetical protein	-	x	0	0%		x	MHP7448_017	Gene/ mC02	PALG_E_242	3	+0.10
			x							PALG_E_812		-2.23
			x							PALG_ES_46		-0.25
mhp027	hypothetical protein	TU/ mhp27	*	4	100%		x	MHP7448_025	TU 02	PALG_E_818	4	-0.91
			*							PALG_E_819		-3.08
			*							**PALG_ES_50**		-0.84
			*							PALG_S_1161		-1.87
gap	glyceraldehyde 3-phosphate dehydrogenase	GENE	*	3	100%		x	MHP7448_035	Gene /mC03	PALG_E_251	3	-4.70
			*							PALG_E_824		-2.82
			*							PALG_E_825		-1.91
rpsG	30s ribosomal protein S7	GENE	*	1	100%		x	MHP7448_076	Gene	PALG_E_259	1	-2.30
mhp093	hypothetical protein	TU/ mhp93	*	5	83%		x	MHP7448_287	Gene/ TU 44	PALG_E_398	6	-5.32
			*							PALG_E_399		-2.67
			*							PALG_E_932		-3.13
			*							PALG_E_933		-4.13
			*							PALG_S_1215		-7.85
			x							PALG_S_1216		-0.20
mhp142	hypothetical protein	-	-	-	-		x	MHP7448_0238	TU 34	*	-	-
rps17	30s ribosomal protein S17	TU/ mhp186	*	6	100%		x	MHP7448_185	TU 27	PALG_E_334	6	-2.12
			*							PALG_E_333		-0.72
			*							PALG_E_892		-3.22
			*							PALG_E_893		-3.82
			*							PALG_ES_83		-0.28
			*							PAL_S_519		-2.62
mhp219	hypothetical protein	-	-	-			x	MHP7448_686	TU 26	*	-	-
ABC	ATP-dependent transport protein	-	-	-			x	MHP7448_160	TU 25	*	-	-
aprE	subtilisin-like serine protease	GENE	x	2	50%		x	MHP7448_0332	Gene	PALG_E_435	4	-1.59
			x							PALG_E_436		-0.45
			*							PALG_E_437		-1.75
			*							PALG_E_957		-4.48
secA	preprotein translocase subunit	GENE	x	4	80%		x	MHP7448_0086	GENE	PALG_E_271	5	-0.63
			*							PALG_E_272		-0.29
			*							PALG_E_845		-7.05
			*							PALG_S_1170		-0.93
			*							PALG_S_1171		-7.33
mhp324	hypothetical protein	TU/ mhp325	*	4	100%		x	MHP7448_0311	TU 50	PALG_E_415	4	+0.55
			*							PALG_S_1224		-2.55
			*							PALG_ES_102		-2.24
			*							PALG_ES_101		-1.81

mhp354	hypothetical protein	GENE	*	6	86%	x			MHP7448_0342	GENE	PALG_E_450	7	-9.34
			*								PALG_E_451		-0.88
			*								PALG_E_452		-0.59
			*								PALG_E_966		-2.56
			x								PALG_ES_123		-0.89
			*								PALG_S_1238		-1.86
			*								PAL_ES_334		-0.18
mhp357	hypothetical protein	-	-	-	-	x			MHP7448_0345	TU 58	*	-	-
mhp360	hypothetical protein	GENE	*	4	100%	x			MHP7448_0349	GENE	PALG_E_969	4	-2.76
			*								PALG_E_970		-2.55
			*								PALG_ES_131		-2.14
			*								PALG_S_1239		-2.09
		TU/ mhp371	*	3	60%	x				TU 63	PALG_E_472	5	-0.87
mhp371	putative high affinity transport system protein p37 precursor		*						MHP7448_0360		PALG_E_473		+0.48
potA	ABC transporter ATP-binding protein		*						MHP7448_0369		PALG_E_474		-0.38
ugpA	sn-glycerol-3-phosphate transport system permease		x						MHP7448_0370		PALG_E_975		+0.18
			x								PALG_ES_134		-1.86
mhp395	hypothetical protein	GENE	x	1	20%	x			MHP7448_0383	TU 65	PALG_E_483	5	-1.12
			x								PALG_E_990		-3.63
			*								PALG_E_991		-3.59
			x								PALG_ES_139		-2.12
			x								PALG_ES_140		-1.53
mhp460	hypothetical protein	-	-	-	-	x			-	-	-	-	-
oppD	oligopeptide transport system permease protein	TU/ mhp502	x	5	83%	x	x	x	MHP7448_502	TU 86	PALG_E_1060	6	-12.61
			*								PALG_E_1061		-5.62
			*								PALG_E_576		-0.90
			*								PALG_E_577		-1.22
			*								PALG_E_578		-0.56
			*								PALG_E_579		-0.88
pgiB	phosphoglucose isomerase B	GENE	x	1	50%	x			MHP7448_531	GENE	PALG_ES_197	**2**	-1.05
			*								PALG_S_1291		-2.03
pyrH	uridylate kinase	GENE	*	2	100%	x			MHP7448_535	GENE	PALG_E_1077	**2**	-0.45
			*								PALG_E_1078		-6.29
mhp575	hypothetical protein	-	-	-		x			MHP7448_555	TU 97	*	-	-
mhp576	hypothetical protein	-	-	0	0%	x			MHP7448_556	GENE	PALG_E_632	3	-0.65
			-								PALG_ES_202		-0.33
			-								PALG_S_1295		-1.85
mhp629	hypothetical protein	GENE	*	3	100%	x			MHP7448_0610	GENE	PALG_E_662	3	-0.43
			*								PALG_E_664		-4.57
			*								PALG_E_665		-7.07
mhp654	hypothetical protein	TU/ mhp654	*	7	100%	x			MHP7448_633	TU 112	PALG_E_1129	7	-1.82
			*								PALG_E_1130		-4.31
			*								PALG_E_687		-0.65
			*								PALG_E_688		-1.49
			*								PALG_E_689		-0.45
			*								PALG_E_690		-3.88
			*								PALG_ES_223		-1.54
p115	p115 protein	GENE	*	1	100%	x			MHP7448_657	GENE	PALG_E_707	1	-3.88
fusA	GTP-binding protein chain elongation	TU	*	4	80%	x	x		fusA	TU 08	PALG_E_261	5	-0.93

	factor ef-g	/mhp088	*							PALG_E_262		-1.12
			x							PALG_E_263		-2.52
			*							PALG_E_840		-0.83
			*							PAL_S_484		-0.44
mhp092	hypothetical protein	TU/ mhp089	x	1	25%	x	x	MHP7448_340 (P)	Gene/ TU 57	PALG_E_455	4	-1.06
			*							PALG_E_456		-2.12
			x							PALG_E_457		+0.20
			x							PALG_ES_125		-4.85
truB	tRNA pseudouridine synthase B	GENE	*	1	100%	x		truB	GENE	PALG_S_1210	1	-3.26
mhp118	hypothetical protein	TU/ mhp123	*	2	100%	x	x	MHP7448_0261	TU 40	PALG_E_379	2	-1.85
			*							PALG_E_917		-2.99
mhp130	putative DNA processing protein	TU/ mhp130	*	5	100%	x		smf	GENE/ TU 37	PALG_E_374	5	-10.43
			*							PALG_E_375		-2.01
			*							PALG_E_914		-0.35
			*							PALG_S_1205		+0.00
			*							PAL_S_530		-0.14
mhp131	hypothetical protein	-	-	-	-	x		-	-	-	-	-
trsE	transfer complex protein	TU/ mhp130	*	5	100%	x		MHP7448_0247	TU 37	PALG_E_374	5	-10.43
			*							PALG_E_375		-2.01
			*							PALG_E_914		-0.35
			*							PALG_S_1205		+0.00
			*							PAL_S_530		-0.14
nrdE	ribonucleoside-diphosphate reductase alpha chain	TU/ mhp155	*	2	0,4	x		nrdE	TU 32	PALG_E_354	5	-1.94
			x							PALG_E_355		-0.22
			*							PALG_E_903		-2.23
			x							PALG_E_904		-5.22
			x							PALG_ES_92		-3.76
mhp170	hypothetical protein		*				x	MHP7448_0211		PALG_E_343		-1.70
oppF	oligopeptide transport system permease protein	TU/ mhp171	*	2	67%	x		oppF-1	TU 31	PALG_E_344	3	-0.34
			x							PALG_E_345		-1.63
mhp181	putative 26.3 kda protein in cilium adhesion operon (orf5)	-	-	-		x	x	MHP7448_0200	mC13	*	-	-
			*	6	100%	x			TU 27	PALG_E_334	6	-2.12
			*							PALG_E_333		-0.72
rpl5	50s ribosomal protein L5	TU/ mhp186	*					MHP7448_0182		PALG_E_892		-3.22
rps8	30s ribosomal protein S8		*				x	rpsH		PALG_E_893		-3.82
adk	adenylate kinase		*				x	adk		PALG_ES_83		-0.28
			*							PAL_S_519		-2.62
mhp228	hypothetical protein	-	-	-	-	x		MHP7448_0154	TU 25	*	-	-
rpe	ribulose-5-phosphate-3-epimerase	-	-	-	-	x		rpe	TU 25	*	-	-
mhp247	hypothetical protein	GENE	*	1	100%	x		MHP7448_0135	GENE	PALG_ES_81	**1**	-0.62
mhp273	hypothetical protein	GENE	x	1	33%	x		MHP7448_0106	GENE	PALG_E_290	3	-1.13
			*							PALG_E_856		-1.91
			x							PALG_E_858		-3.08
mhp303	hypothetical protein	GENE	x	2	29%	x		MHP7448_0342 (P)	GENE	PALG_E_450	7	-9.34
			x							PALG_E_451		-0.88
			*							PALG_E_452		-0.59
			x							PALG_E_966		-2.56
			*							PALG_ES_123		-0.89
			x							PALG_S_1238		-1.86

			x								PAL_ES_334		-0.18
gtp1	GTP-binding protein	TU/ mhp307	*	2	100%	x			ychF	TU 46	PALG_E_937	2	-2.14
rps18	30s ribosomal protein S18		*						rpsR		PALG_S_1222		-4.04
mhp309	hypothetical protein	-	x	0	0%	x			MHP7448_297	GENE	PALG_S_1221	1	-2.19
mhp312	hypothetical protein	GENE	*	2	100%	x			MHP7448_299	GENE/ TU 47	PALG_E_938	2	-2.73
			*								PALG_ES_98		-3.28
mhp341	hypothetical protein	GENE	*	1	100%	x			MHP7448_330	GENE	PALG_ES_114	**1**	-0.70
mhp362	hypothetical protein	TU/ mhp362	*	3	100%	x			MHP7448_351	GENE/ TU 59	PALG_E_461	3	+0.31
			*								PALG_E_462		+0.28
			*								PALG_E_463		-1.39
rluD	ribosomal large subunit pseudouridine synthase D	TU/ mhp453	*	4	100%	x			MHP7448_458	TU 78	PALG_E_1032	4	-3.68
			*								PALG_E_534		-4.83
			*								PALG_E_535		-2.99
			*								PALG_ES_166		-3.25
mhp472	hypothetical protein	TU/ mhp469	*	4	100%	x			MHP7448_474	TU 82	PALG_E_1037	4	-6.68
			*								PALG_E_1038		-0.25
			*								PALG_S_1277		-3.13
			*								PALG_S_1278		-2.43
mhp480	hypothetical protein	TU/ mhp483	*	6	100%	x		x	MHP7448_482	TU 83	PALG_E_1042	6	-3.87
			*								PALG_E_1043		-3.24
			*								PALG_E_1044		-2.79
			*								PALG_E_554		-2.74
			*								PALG_ES_174		-2.82
			*								PALG_ES_176		+1.04
mhp517	hypothetical protein	GENE	*	1	100%	x			MHP7448_0519	GENE	PALG_ES_188	**1**	-2.02
mhp583	hypothetical protein	TU/ mhp585	x	1	0,25	x			MHP7448_566	TU 100	PALG_E_1094	4	-3.61
			*								PALG_ES_207		-0.90
			x								PALG_ES_208		-0.51
			x								PALG_S_1300		-2.94
		TU/ mhp107	*	4	100%	x				TU 107	PALG_E_1114	4	-1.33
			*								PALG_E_1115		-4.03
mhp634	hypothetical protein		*					x	MHP7448_0615		PALG_E_673		-2.04
rplL	50s ribosomal protein L7/l12		*				x		MHP7448_0618		PALG_ES_218		-2.91
rpoC	DNA-directed RNA polymerase beta chain	GENE	*	1	100%	x		x	MHP7448_0616	GENE	PALG_E_1110	1	-1.72
cysS	cysteinyl-tRNA synthetase	TU/ mhp668	*	3	75%	x			MHP7448_641	TU 113	PALG_E_1135	4	-7.42
			*								PALG_E_1136		-5.71
			*								PALG_E_696		-2.88
			x								PALG_E_697		-1.43
mhp670	hypothetical protein	TU/ mhp669	x	1	25%	x			MHP7448_649	TU114	PALG_E_1137	4	-7.32
			x								PALG_E_1138		-5.61
			*								PALG_E_699		-2.89
			x								PALG_E_700		-1.91
mhp140	hypothetical protein	GENE	*	6	100%		x		MHP7448_0240	GENE	PALG_E_364	6	-2.41
			*								PALG_E_365		-1.46
			*								PALG_E_907		-1.32
			*								PALG_E_908		-2.49
			*								PALG_E_909		-2.31
			*								PALG_E_910		+0.47
glpQ	glycerophosphoryl diester	TU/	*	2	100%		x		ugpQ	TU 47	PALG_E_938	2	-2.73

	phosphodiesterase	mhp312	*							PALG_ES_98		-3.28
ackA mhp510	acetate kinase hypothetical protein	TU/ mhp510	x	4	100%	x		ackA MHP7448_0512	TU 88	PALG_E_1065	4	-1.64
			*							PALG_E_1066		-4.92
			*							PALG_E_585		-2.61
			*							PALG_ES_186		-0.41
potB	spermidine/putrescine transport system permease	-	-	-		x		potB	GENE	*	-	-
mhp081	hypothetical protein	-	-	-		x		-	-	*	-	-
mhp087	hypothetical protein	GENE	x	0	0%	x	x	MHP7448_079	TU 08	PALG_E_261	4	-0.93
			x							PALG_E_262		-1.12
			x							PALG_E_840		-0.83
			x							PAL_S_484		-0.44
rpsL	30s ribosomal protein S12	GENE	*	2	100%	x		rpsN	GENE	PALG_E_260	2	-1.82
			*							PALG_E_839		-3.26
hpt	hypoxanthine phosphoribosyl transferase	TU/ mhp123	*	2	100%	x		hpt	TU 40	PALG_E_379	2	-1.85
			*							PALG_E_917		-2.99
rpl2	50s ribosomal protein L2	GENE	*	1	100%	x	x	rplB	GENE	PALG_E_332	1	-3.77
mhp411	hypothetical protein	GENE	*	2	100%	x		MHP7448_397	GENE	PALG_E_492	2	-1.92
			*							PALG_E_998		-5.40
metK	S-adenosylmethionine synthetase	GENE	*	1	100%	x		MHP7448_451	GENE	PALG_E_531	1	-9.21
mhp639	hypothetical protein	TU/ mhp639	*	4	100%	x		MHP7448_0620	GENE/ TU 108	PALG_E_1117	4	-1.33
			*							PALG_E_1118		-4.32
			*							PALG_E_674		-1.33
			*							PALG_S_1313		-2.24
ushA	5'-nucleotidase	TU/ mhp651	*	7	100%	x		MHP7448_0630	GENE/ TU 111	PALG_E_1126	7	-6.04
			*							PALG_E_1127		-3.89
			*							PALG_E_1128		-2.43
			*							PALG_E_684		-1.63
			*							PALG_E_685		-0.25
			*							PALG_E_686		-4.41
			*							PALG_ES_222		-1.53
mhp666	hypothetical protein	TU/ mhp668	*	3	75%	x		MHP7448_0646	TU 113	PALG_E_1135	4	-7.42
			*							PALG_E_1136		-5.71
			*							PALG_E_696		-2.88
			x							PALG_E_697		-1.43
dnaN	DNA polymerase III, beta chain	GENE	*	2	100%		x	dnaN	GENE	PALG_ES_43	2	-1.78
			*							PALG_E_805		-3.59
mhp006	hypothetical protein	TU/ mhp001	*	5	100%		x	MHP7448_006	TU 01	PALG_E_237	5	-1.65
			*							PALG_E_238		-8.06
			*							PALG_E_239		-4.11
			*							PALG_E_240		-0.64
			*							PAL_ES_241		-0.61
mhp022	ABC transporter ATP binding protein	GENE	*	1	100%		x	MHP7448_020	GENE	PALG_E_243	1	-1.32
atpG tsf	ATP synthase gamma chain elongation factor ts	TU/ mhp042	*	2	100%		x	atpG tsf	TU 06	PALG_E_253	2	-0.37
			*							PALG_E_827		-5.61
topA	DNA topoisomerase I (omega-protein)	TU/ mhp093	*	5	83%		x	topA	TU 44	PALG_E_398	6	-5.32
			*							PALG_E_399		-2.67
			*							PALG_E_932		-3.13
			*							PALG_E_933		-4.13
			*							PALG_S_1215		-7.85

			x						PALG_S_1216		-0.20
smpB	SsrA-binding protein	TU/ mhp110	*	2	100%	x	smpB	GENE/ TU 41	PALG_E_921	2	-1.12
			*						PALG_E_922		-1.16
nrdF	ribonucleoside-diphosphate reductase beta chain	GENE	*	2	50%	x	nrdF	GENE	PALG_E_352	4	-4.53
			*						PALG_S_1201		-3.42
			x						PALG_E_353		-3.25
			x						PALG_E_902		-3.51
rps19	30s ribosomal protein S19	TU/ mhp186	*	6	100%	x	rpsS	TU 27	PALG_E_333	6	-3.25
rpl14	50s ribosomal protein L14		*				rplN		PALG_E_334		-0.72
rpl6	50s ribosomal protein L6		*				rplF		PALG_E_892		-2.12
rps5	30s ribosomal protein S5		*				rpsE		PALG_E_893		-0.72
rpoA	DNA-directed RNA polymerase alpha chain		*				rpoA		PALG_ES_83		-3.22
			*						PAL_S_519		-3.82
rpl2	50s ribosomal protein L2	-	x	0	0%	x	rplB	GENE	PALG_E_332	1	-3.77
mhp230	hypothetical protein	-	-	-	-	x	MHP7448_0152	TU 25	*	-	-
mhp377	putative lipoprotein	GENE	*	4	0,8	x	MHP7448_0366	GENE	PALG_E_980	5	-4.52
			*						PALG_E_981		-2.14
			*						PALG_E_477		-1.55
			*						PALG_E_982		-0.73
			x						PALG_S_1240		-2.71
mhp401	hypothetical protein	GENE	*	1	100%	x	MHP7448_389 (P)	GENE	PALG_ES_141	1	-0.34
mhp429	hypothetical protein	GENE	*	4	57%	x	MHP7448_411	GENE	PALG_E_505	7	-2.64
			*						PALG_E_506		-1.59
			*						PALG_E_507		-0.93
			x						PALG_E_508		-1.24
			x						PALG_E_509		-2.15
			x						PALG_ES_152		-2.88
			*						PALG_S_1253		-5.47
mhp434	hypothetical protein	GENE	*	3	60%	x	MHP7448_431	GENE	PALG_E_514	5	-1.20
			x						PALG_E_1014		-2.08
			x						PALG_E_515		-5.05
			*						PALG_ES_160		-2.74
			*						PALG_S_1257		-2.60
		TU/ mhp483	*	6	100%	x		TU 83	PALG_E_1042	6	-3.87
			*						PALG_E_1043		-3.24
atpA	ATP synthase alpha chain		*				atpA		PALG_E_1044		-2.79
mhp478	hypothetical protein		*				MHP7448_480		PALG_E_554		-2.74
mhp481	hypothetical protein		*				MHP7448_483		PALG_ES_174		-2.82
			*						PALG_ES_176		+1.04
mhp482	hypothetical protein	TU/ mhp483	*	6	100%	x	MHP7448_484	GENE/ TU 83	PALG_ES_174	6	-3.87
			*						PALG_E_1042		-3.24
			*						PALG_E_1043		-2.79
			*						PALG_E_1044		-2.74
			*						PALG_E_554		-2.82
			*						PALG_ES_176		+1.04
mhp284	hypothetical protein	GENE	*	1	25%	x	MHP7448_0095	GENE/ mC09	PALG_E_850	4	-5.86
			x						PALG_S_1175		-3.90
			x						PALG_S_1176		-5.18
			x						PAL_E_99		-7.27
mhp527	hypothetical protein	-	-	-	-	x	-	-	-	-	-

mhp533	hypothetical protein	-	-	-	-		-	-	-	-	-	-
polC	DNA polymerase III, alpha chain	TU/ mhp552	*	2	100%		x	polC	TU 92	PALG_E_1077	2	-0.45
			*							PALG_E_1078		-6.29
pepP	xaa-pro aminopeptidase	TU/ mhp675	*	2	100%		x	pepP	TU 116	PALG_E_704	2	-2.68
			*							PALG_E_705		-3.24
p146	p146 adhesin like-protein, p97 paralog	GENE	*	7	100%		x	MHP7448_663	GENE	PALG_E_1145	7	-0.17
			*							PALG_E_1146		-4.40
			*							PALG_E_1147		-1.65
			*							PALG_E_1148		-1.83
			*							PALG_E_708		-4.79
			*							PALG_E_709		-0.18
			*							PALG_ES_229		-0.10
mhp077	hypothetical protein	GENE	*	1	100%		x	MHP7448_071	GENE	PALG_ES_65	1	-0.32
era	GTP-binding protein era homolog	-	-	-	-		x	era	-	*	-	-
mhp236	chromate transport protein, putative	TU/ mhp237	*	1	25%		x	MHP7448_0146	TU 24	PALG_E_325	4	-2.03
			x							PALG_E_880		-1.52
			x							PALG_E_881		-2.96
			x							PALG_ES_82		-1.64
mhp302	hypothetical protein	-	-	-	-		x	-	-	-	-	-
mhp350	hypothetical protein	TU/ mhp351	x	1	25%		x	MHP7448_0338	Gene/ TU 57	PALG_E_455	4	-1.06
			x							PALG_E_456		-2.12
			*							PALG_E_457		+0.20
			x							PALG_ES_125		-4.85
mhp405 mhp406 ftsZ	hypothetical protein hypothetical protein cell division protein	TU/ mhp405	*	7	100%		x	MHP7448_391 MHP7448_392 MHP7448_393	TU 67	PALG_E_489	7	-1.47
			*							PALG_E_994		-1.21
			*							PALG_E_995		-3.98
			*							PALG_ES_144		-1.88
			*							PALG_ES_145		-3.13
			*							PALG_S_1247		-1.46
			*							PAL_ES_354		+0.41
mhp461	hypothetical protein	GENE	*	5	100%		x	MHP7448_463		PALG_E_1034	5	-0.55
			*							PALG_E_1035		-3.83
			*							PALG_E_539		-1.23
			*							PALG_ES_171		-0.94
			*							PALG_S_1273		-2.48
pdhD	dihydrolipoamide dehydrogenase	TU/ mhp503	*	5	100%		x	pdhD	TU 87	PALG_E_1062	5	-3.14
			*							PALG_E_1063		-1.73
			*							PALG_E_580		-2.95
			*							PALG_E_581		-0.96
			*							PALG_E_582		-2.73
ktrA	potassium uptake protein	GENE	*	5	100%		x	MHP7448_545	GENE/ mC32	PALG_E_613	5	-2.91
			*							PALG_E_614		-3.89
			*							PALG_E_615		-2.15
			*							PALG_E_616		-0.28
			*							PALG_E_617		+0.47
mhp574	hypothetical protein	-	-	-	-		x	MHP7448_715	TU 97	*	-	-
nagE	pts system, n-acetylglucosamine-specific enzyme II, ABC component	GENE	x	0	0%		x	MHP7448_574	TU 102	PALG_E_797	1	-1.76
lspA	lipoprotein signal peptidase (prolipoprotein signal peptidase) (signal peptidase II) (spase II)	TU/ mhp35	*	3	100%	x		lps	TU 04	PALG_E_249	3	-2.00

parC	topoisomerase IV subunit A		*						x	parC		PALG_E_822		-4.40
			*									PALG_E_823		-2.12
mhp057	hypothetical protein	GENE	*	4	100%	x				MHP7448_054	GENE	PALG_E_829	4	-3.93
			*									PALG_ES_57		-1.25
			*									PALG_E_830		-2.90
			*									PAL_S_477		-5.84
ffh	signal recognition particle protein	-	-	-	-	x				MHP7448_057	GENE	*	-	-
mhp064	hypothetical protein	TU/ mhp42	*	2	100%	x				MHP7448_061	TU 06	PALG_E_253	2	-0.37
			*									PALG_E_827		-5.61
mhp069	hypothetical protein	GENE	x	0	0%	x				MHP7448_065 (P)	GENE	PALG_ES_59	1	-5.18
mhp071	hypothetical protein	TU/ mhp071	*	2	100%	x				MHP7448_0681	TU 07	PAL_E_77	2	-3.33
lepA	30 kda GTP-binding protein lepA		*						x	MHP7448_072		PAL_ES_442		+0.87
dnaK	chaperone protein dnaK	GENE	*	3	100%	x				dnaK	GENE	PALG_E_258	3	-3.50
			*									PALG_E_835		-4.74
			*									PALG_ES_63		-0.46
dnaJ	heat-shock protein	-	-	-		x				dnaJ	GENE	*	-	-
serS	seryl-tRNA synthetase	TU/ mhp129	*	6	100%	x			x	serS	TU 38	PALG_E_377	6	-10.73
			*									PALG_E_378		-2.72
			*									PALG_E_915		-2.05
			*									PALG_S_1207		-0.57
			*									PAL_S_534		+0.16
			*									PAL_S_536		-1.85
mhp144	putative dehydrogenase	TU/ mhp153	*	6	100%	x			x	MHP7448_0235	TU 33	PALG_E_356	6	-3.48
mhp145	putative D-ribose-binding protein mutant		*							rbsB		PALG_E_357		-5.57
mhp148	hypothetical protein		*							iolE		PALG_E_358		-0.37
iolD	myo-inositol catabolism		*							iolD		PALG_E_359		-0.45
mhp150	hypothetical protein		*							MHP7448_0228		PALG_E_905		-5.32
mhp151	hypothetical protein		*				x	x		iolB		PALG_E_906		+0.15
iolC	myo-inositol catabolism		-					x		iolC				-
rbsA	ribose transport ATP-binding protein	GENE	*	1	50%	x				mglA	GENE	PALG_E_360	2	+0.17
			x									PALG_E_362		-0.11
mhp171	ABC transporter ATP binding protein	GENE	*	3	75%	x			x	MHP7448_0210	GENE	PALG_E_342	4	-0.41
			*									PALG_E_897		-1.11
			*									PALG_E_898		-2.35
			x									PALG_S_1198		+0.00
thdF	thiophene and furan oxidation protein	-	-	-	-	x				trmE	GENE	*	-	-
		TU/ mhp186	*	6	100%	x					TU 27	PALG_E_333	6	-2.12
map	methionine amino peptidase		*						x	map		PALG_E_334		-0.72
rps11	30s ribosomal protein S11		*							rpsK		PALG_E_892		-3.22
rpl17	50s ribosomal protein L17		*							rplQ		PALG_E_893		-3.82
			*									PALG_ES_83		-0.28
			*									PAL_S_519		-2.62
rps13	30s ribosomal protein S13	GENE	*	2	67%	x	x			rpsM	GENE	PALG_E_889	3	-2.65
			*									PALG_E_330		-10.66
			x									PAL_S_516		-3.09
DeoB	phosphopentomutase	-	-	-	-	x				DeoB	TU 26	*	-	-
mhp224	hypothetical protein	-	-	-	-	x			x	MHP7448_0158	TU 25	*	-	-
mhp254	hypothetical protein	TU/ mhp254	*	1	100%	x				MHP7448_0127	GENE/ TU 19	PALG_ES_77	1	-1.39
mhp271	p97 cilium adhesin paralog	GENE	*	3	100%	x				MHP7448_0108	GENE	PALG_E_291	3	-3.67

			*										PALG_E_859		-3.27
			*										PALG_S_1184		-4.65
mhp275	hypothetical protein, p102 paralog	GENE	x	0	0%		x		x		MHP7448_0105 (P)	GENE	PALG_S_1182	1	-3.32
tpiA	triosephosphate isomerase	GENE	*	2	40%		x				tpiA	GENE	PALG_E_298	5	-3.10
			x										PALG_E_299		-8.11
			x										PALG_E_864		-2.65
			*										PALG_S_1186		-3.73
			x										PAL_E_3		-1.12
clpB	ATP-dependent serine proteinase, heat shock protein	GENE	*	4	100%		x				clpB	GENE	PALG_E_853	4	-3.33
			*										PALG_E_854		-3.90
			*										PALG_S_1180		-7.37
			*										PAL_S_495		-3.25
mhp289	hypothetical protein	GENE	*	5	100%		x				MHP7448_0090	GENE	PALG_E_847	5	-1.08
			*										PALG_E_848		-0.29
			*										PALG_ES_71		-1.59
			*										PALG_ES_72		+0.75
			*										PALG_S_1173		-2.19
mhp308	hypothetical protein	-	-	-	-		x				-	-	-	-	-
mhp334	hypothetical protein	TU/ mhp326	*	5	100%		x				MHP7448_0319	TU 51	PALG_ES_104	5	-1.81
			*										PALG_ES_105		-1.85
			*										PALG_E_416		-2.57
			*										PALG_E_417		-2.49
			*										PALG_S_1225		-4.34
mhp356	hypothetical protein	-	-	-	-		x	x			MHP7448_0696 (P)	mC24	*	-	-
glpF	glycerol uptake facilitator protein	TU/ mhp370	*	1	100%		x				glpF	TU 62	PAL_S_576	1	-1.91
glpK	glycerol kinase (ATP:glycerol 3-phosphotransferase) (glycerokinase)	GENE	*	3	75%		x			x	glpK		PALG_E_470	4	-1.84
			*										PALG_E_471		-1.73
			*										PALG_ES_133		-0.20
			x										PAL_E_158		-3.81
mhp374	hypothetical protein	GENE	*	1	100%		x			x	MHP7448_0363	GENE	PAL_S_577	1	-0.25
mhp399	hypothetical protein	GENE	*	4	100%		x				MHP7448_386	GENE	PALG_E_484	4	-2.96
			*										PALG_E_485		-4.21
			*										PALG_E_993		-3.09
			*										PALG_S_1246		-6.17
mhp445	hypothetical protein	GENE	*	4	80%	x	x				MHP7448_443	Gene	PALG_E_1023	5	-1.54
			*										PALG_S_1263		-1.67
			*										PALG_E_1024		-7.93
			x										PALG_S_1264		-3.35
			*										PALG_S_1265		-1.52
mhp446	hypothetical protein	-	-	-	-		x				MHP7448_444	TU 74	*	-	-
rplK	50s ribosomal protein L11	GENE	*	2	100%		x				rplK	GENE	PALG_E_1033	2	-3.24
			*										PALG_ES_168		-1.62
mhp479	hypothetical protein	TU/ mhp483	*	6	100%		x			x	MHP7448_481	TU 83	PALG_E_1042	6	-3.87
			*										PALG_E_1043		-3.24
			*										PALG_E_1044		-2.79
			*										PALG_E_554		-2.74
			*										PALG_ES_174		-2.82
			*										PALG_ES_176		+1.04
asnS	asparaginyl-tRNA synthetase	GENE	*	2	67%		x				MHP7448_500	GENE	PALG_E_1057	3	-2.65

			x										PALG_E_574		-2.57
			*										PALG_E_575		-3.02
		TU/ mhp502	x	5	83%		x		x			TU 86	PALG_E_1060	6	-12.61
			*										PALG_E_1061		-5.62
oppC	oligopeptide transport system permease protein		*								oppC		PALG_E_576		-0.90
oppB	oligopeptide transport system permease protein		*							x	oppB		PALG_E_577		-1.22
			*										PALG_E_578		-0.56
			*										PALG_E_579		-0.88
P46	46kd surface antigen precursor	TU/ mhp511	*	5	100%		x				MHP7448_0513	TU 89	PALG_E_586	5	-1.12
			*										PALG_E_587		-2.40
			*										PALG_E_588		-5.49
			*										PALG_E_589		-2.34
			*										PALG_ES_187		-0.44
mhp596	hypothetical protein	-	-	-			x				MHP7448_579	GENE	*	-	-
		TU/ mhp001	*	5	100%	x	x					TU 01	PALG_E_237	5	-1.65
mhp005	hypothetical protein		*								MHP7448_005		PALG_E_238		-8.06
mhp013	hypothetical protein		*							x	MHP7448_013		PALG_E_239		-4.11
rluC	ribosomal large subunit pseudouridine synthase		*								rluC		PALG_E_240		-0.64
			*										PAL_ES_241		-0.61
mhp007	hypothetical protein	-	-	-	-		x				MHP7448_007	GENE	*	-	-
mhp023	ABC transporter ATP binding protein	TU/ mhp27	*	4	100%		x				MHP7448_021	TU 02	PALG_E_818	4	-0.91
			*										PALG_E_819		-3.08
			*										**PALG_ES_50**		-0.84
			*										PALG_S_1161		-1.87
gatB	glutamyl-tRNA amidotransferase, subunit B	TU/ mhp29	x	0	0%		x		x		gatB	TU 03	PALG_E_245	5	-2.19
			x										PALG_E_820		-1.21
			x										PALG_ES_51		-0.54
			x										PALG_ES_52		-1.33
			x										PALG_S_1162		-2.65
mhp038	hypothetical protein	TU/ mhp41	*	2	100%	x	x				VacB	TU 05	PALG_E_252	2	-0.05
			*										PALG_E_826		-5.91
rpsT	30s ribosomal protein S20	GENE	*	2	100%		x				rpsT	GENE	PALG_E_372	2	-1.26
			*										PALG_E_913		-1.39
mhp136	hypothetical protein	TU/ mhp139	*	1	100%		x	x			MHP7448_0244	GENE	PALG_E_371	1	-3.71
mhp176	hypothetical protein	-	-	-	-		x		x		MHP7448_0205	TU 29	*	-	-
mhp235	hypothetical protein	TU/mhp237	*	1	25%		x	x			MHP7448_0147	TU 24	PALG_E_325	4	-2.03
			x										PALG_E_880		-1.52
			x										PALG_E_881		-2.96
			x										PALG_ES_82		-1.64
mhp240	hypothetical protein	GENE	x	0	0%		x			x	MHP7448_0142	TU 23	PALG_E_323	1	-5.20
gltX	glutamyl tRNA synthetase	-	-	-	-		x				gltX	GENE	*	-	-
mhp310	hypothetical protein	GENE	x	0	0%		x	x			MHP7448_297	GENE	PALG_S_1221	1	-2.19
mod	site-specific DNA-methyltransferase (adenine-specific)	GENE	x	0	0%		x				MHP7448_0316	GENE	PALG_S_1227	1	-3.11
mhp337	hypothetical protein	GENE	*	2	100%		x		x		MHP7448_0089	GENE	PALG_E_846	2	-4.18
			*										PALG_S_1172		-0.71
mhp353	hypothetical protein	TU/	x	1	25%		x				MHP7448_0341	Gene/	PALG_E_455	4	-1.06

		mhp354	*						TU 57	PALG_E_456		-2.12
			x							PALG_E_457		+0.20
			x							PALG_ES_125		-4.85
mhp373	putative transport system permease protein p69	TU/ mhp370	*	3	60%	x		MHP7448_0362	TU 63	PALG_E_472	5	-0.87
			*							PALG_E_473		+0.48
			*							PALG_E_474		-0.38
			x							PALG_E_975		+0.18
			x							PALG_ES_134		-1.86
mhp393	hypothetical protein	TU/ mhp391	*	3	60%	x		MHP7448_0381	TU 65	PALG_E_483	5	-1.12
			*							PALG_E_990		-3.63
			*							PALG_E_991		-3.59
			x							PALG_ES_139		-2.12
			x							PALG_ES_140		-1.53
proS	prolyl aminoacyl-tRNA synthetase	TU/ trna-tyr	x	2	40%	x	x	proS	TU 66	PALG_E_486	5	+0.07
			*							PALG_E_487		-1.13
			*							PALG_E_488		-2.12
			x							PALG_ES_142		-1.96
			x							PALG_ES_143		-4.40
metG	methionyl-tRNA synthetase	TU/ mhp405	*	7	100%	x		metS	TU 67	PALG_E_489	7	-1.47
			*							PALG_E_994		-1.21
			*							PALG_E_995		-3.98
			*							PALG_ES_144		-1.88
			*							PALG_ES_145		-3.13
			*							PALG_S_1247		-1.46
			*							PAL_ES_354		+0.41
mhp419	hypothetical protein	TU/ mhp416	x	2	33%	x		MHP7448_404	TU 69	PALG_E_496	6	-3.63
			*							PALG_E_497		-1.14
			x							PALG_ES_148		-4.88
			x							PALG_ES_149		-2.88
			*							PALG_S_1249		-0.91
			x							PALG_S_1250		-1.60
acpD	acyl carrier protein phosphodiesterase	TU/ mhp453	*	4	100%	x		acpD-1	TU 78	PALG_E_1032	4	-3.68
			*							PALG_E_534		-4.83
			*							PALG_E_535		-2.99
			*							PALG_ES_166		-3.25
mhp465	hypothetical protein	TU/ mhp468	*	4	100%	x		MHP7448_467	TU 81	PALG_E_1036	4	-6.01
mhp466	hypothetical protein		*				x	MHP7448_468		PALG_E_543		-0.98
bcrA	ABC transporter ATP binding protein		*					MHP7448_469		PALG_S_1275		-3.12
mhp468	hypothetical protein		*				x	MHP7448_470		PALG_S_1276		-1.16
mhp474	hypothetical protein	GENE	*	2	100%	x	x	MHP7448_476		PALG_ES_174	2	-2.82
			*							PALG_ES_176		+1.04
mgtE	MG2+ ion transporter	TU/ mhp487	*	6	100%	x		mgtE	TU 84	PALG_E_1049	6	-0.83
			*							PALG_E_558		-0.54
			*							PALG_E_559		+0.28
			*							PALG_E_560		-0.52
			*							PAL_E_184		-1.23
			*							PAL_S_614		-2.93
mhp486	hypothetical protein	GENE	*	4	100%	x		MHP7448_488	GENE	PALG_E_556	4	-0.83
			*							PALG_E_557		-7.02
			*							PALG_ES_180		-2.29

			*										PALG_S_1280		-2.12
mhp509	hypothetical protein	TU/ mhp510	*	4	100%		x				MHP7448_0713	TU 88	PALG_E_1065	4	-1.64
			*										PALG_E_1066		-4.92
			*										PALG_E_585		-2.61
			*										PALG_ES_186		-0.41
mhp526	hypothetical protein	-	-	-	-		x	x		x	MHP7448_0423 (P)	TU 71	*	-	-
mhp584	hypothetical protein	TU/ mhp585	x	1	25%		x				MHP7448_567	GENE	PALG_E_1094	4	-3.61
			*										PALG_ES_207		-0.90
			x										PALG_ES_208		-0.51
			x										PALG_S_1300		-2.94
polA	DNA polymerase I	TU/ mhp604	*	6	100%		x				MHP7448_581	TU 104	PALG_E_800	6	-1.28
			*										PALG_E_801		-7.98
			*										PALG_S_1398		-1.30
			*										PALG_S_1399		+0.21
			*										PAL_E_234		-2.54
			*										PAL_ES_458		-0.42
mhp619	hypothetical protein	TU/ mhp610	x	0	0%		x				MHP7448_0600	TU 106	PALG_ES_213	1	-4.72
mhp620	hypothetical protein										MHP7448_0601 (P)				
mhp621	hypothetical protein										MHP7448_0602				
mhp658	hypothetical protein	-	-	-	-		x				-	-	-	-	-
rplM	50s ribosomal protein L13	-	-	-	-	x	x				rplM	-	-	-	-
mhp674	hypothetical protein	GENE	x	2	67%		x	x			MHP7448_653	GENE	PALG_E_701	3	-0.84
			*										PALG_E_702		-3.24
			*										PALG_E_703		+0.35
mhp697	hypothetical protein	-	-	-	-		x				MHP7448_676	TU 121	-	-	-

[1] Grown culture conditions that *M. hyopneumoniae* strain 232 (MHP_232) was exposed. Location Repeat, Repeat and Conservation were classified as Supplementary Table S5.

8 CURRICULUM VITAE

CATTANI, A.M.

1. DADOS PESSOAIS

Nome: Amanda Malvessi Cattani

Nascimento: 08/03/1992, Garibaldi, RS, Brasil

E-mail: amanda.cattani@gmail.com

2. FORMAÇÃO

Mestrado em Biologia Celular e Molecular.

Universidade federal do Rio Grande do Sul, UFRGS, Porto Alegre, RS, Brasil.

Título: Elementos repetitivos na regulação da transcrição de *Mycoplasma hyopneumoniae*

Orientadora: Dra. Irene Silveira Schrank

Bolsista: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Período: 2014 – 2016

Graduação em Biotecnologia

Universidade federal do Rio Grande do Sul, UFRGS, Porto Alegre, RS, Brasil.

Título: Clonagem e expressão de genes relacionados com a biossíntese de NAD e FAD de *Mycoplasma hyopneumoniae*

Orientadora: Dra. Irene Silveira Schrank

Período: 2010 – 2014

Curso Técnico em Enologia e Viticultura

Instituto Federal do Rio Grande do Sul, IFRS, Bento Gonçalves, RS, Brasil

Período: 2007 – 2010

3. ESTÁGIOS

Estágio Obrigatório

Laboratório da empresa BEIFIUR Ltda, Garibaldi, RS, Brasil.

Projeto: Avaliar o comportamento dos micro-organismos *Azospirillum brasilense, Bradyrhizobium japonicum, Bacillus subtilis*, *Bacillus thuringiensis* e *Bacillus sphaericus*, importantes no melhoramento e desenvolvimento agrícola, quando combinados com herbicidas ou fungicidas comumente utilizados nas lavouras. Além de avaliar métodos para a produção e manutenção desses micro-organismos.

Período: 08/2013 – 12/2013

Estágio de Iniciação Cientifica

Laboratório de Micro-organismos Diazotróficos – Centro de Biotecnologia

Universidade federal do Rio Grande do Sul, UFRGS, Porto Alegre, RS, Brasil.

Projeto: Caracterização de proteínas envolvidas em rotas metabólicas de *Mycoplasma hyopneumoniae.*

Orientação: Dra. Irene Silveira Schrank

Período: 2012 – 2013

Bolsista: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Estágio de Iniciação Cientifica Voluntário

Laboratório de Micro-organismos Diazotróficos – Centro de Biotecnologia

Universidade federal do Rio Grande do Sul, UFRGS, Porto Alegre, RS, Brasil.

Projeto: Caracterização de proteínas envolvidas em rotas metabólicas de *Mycoplasma hyopneumoniae.*

Orientação: Dra. Irene Silveira Schrank

Período: 2011 – 2012

Estágio Obrigatório

Laboratório de análises da empresa Chandon Brasil Ltda, Garibaldi, RS, Brasil.

Projeto: Atividades gerais no controle laboratorial dos processos fermentativos da produção de espumantes. Período: 2009

4. ARTIGOS COMPLETOS PUBLICADOS

SIQUEIRA, F. M., WEBER, de S. S., CATTANI, A. M., SCHRANK, I. S. 2014, Genome organization in Mycoplasma hyopneumoniae: identification of promoter-like sequences, *Mol Biol Rep.*, 41 (8), 5395-402.

5. RESUMOS PUBLICADOS EM EVENTOS E CONGRESSO

CATTANI, A. M.; SIQUEIRA, F. M.; SCHRANK, I. S. Relevance of repetitive elements on transcription regulation in Mycoplasma hyopneumoniae. In: 6° Congresso de Biotecnologia, 2015, Brasília. Anais do 6° Congresso de Biotecnologia, 2015.

CATTANI, A. M.; SIQUEIRA, F. M.; SCHRANK, I. S. *In silico* identification of repetitive elements in Mycoplasma hyopneumoniae. In: XVI Encontro do PPGBCM, 2015, Porto Alegre. Livro de Resumos XVI Encontro do PPGBCM, 2015.

CATTANI, A. M.; SIQUEIRA, F. M.; SCHRANK, I. S. Clonagem e expressão de genes relacionados com a biossíntese de NAD e FAD de Mycoplasma hyopneumoniae. In: XXV Salão de Iniciação Científica da UFRGS, 2013, Porto Alegre. Livro de Resumos XXV Salão de Iniciação Científica da UFRGS, 2013.

CATTANI, A. M.; SIQUEIRA, F. M.; SCHRANK, I. S. Promoter-Like sequences localization in transcription units of Mycoplasma hyopneumoniae. In: 28ª Reunião de Genética de Micro-organismos, 2012, Foz do Iguaçu. Anais da 28ª Reunião de Genética de Micro-organismos, 2012.

CATTANI, A. M.; SIQUEIRA, F. M.; SCHRANK, I. S. Clonagem e expressão de genes relacionados a síntese de FAD e NAD de Mycoplasma hyopneumoniae. In: XXIV Salão de Iniciação Científica, 2012, Porto Alegre. Livro de Resumos XXV Salão De Iniciação Científica da UFRGS, 2012.