

RESEARCH ARTICLE

Open Access

A combination of improved differential and global RNA-seq reveals pervasive transcription initiation and events in all stages of the life-cycle of functional RNAs in *Propionibacterium acnes*, a major contributor to wide-spread human disease

Yu-fei Lin¹, David Romero A¹, Shuang Guan¹, Lira Mamanova² and Kenneth J McDowall^{1*}

Abstract

Background: Sequencing of the genome of *Propionibacterium acnes* produced a catalogue of genes many of which enable this organism to colonise skin and survive exposure to the elements. Despite this platform, there was little understanding of the gene regulation that gives rise to an organism that has a major impact on human health and wellbeing and causes infections beyond the skin. To address this situation, we have undertaken a genome-wide study of gene regulation using a combination of improved differential and global RNA-sequencing and an analytical approach that takes into account the inherent noise within the data.

Results: We have produced nucleotide-resolution transcriptome maps that identify and differentiate sites of transcription initiation from sites of stable RNA processing and mRNA cleavage. Moreover, analysis of these maps provides strong evidence for 'pervasive' transcription and shows that contrary to initial indications it is not biased towards the production of antisense RNAs. In addition, the maps reveal an extensive array of riboswitches, leaderless mRNAs and small non-protein-coding RNAs alongside vegetative promoters and post-transcriptional events, which includes unusual tRNA processing. The identification of such features will inform models of complex gene regulation, as illustrated here for ribonucleotide reductases and a potential quorum-sensing, two-component system.

Conclusions: The approach described here, which is transferable to any bacterial species, has produced a step increase in whole-cell knowledge of gene regulation in *P. acnes*. Continued expansion of our maps to include transcription associated with different growth conditions and genetic backgrounds will provide a new platform from which to computationally model the gene expression that determines the physiology of *P. acnes* and its role in human disease.

Keywords: Pervasive transcription, Leaderless mRNA, Small non-protein-coding RNAs, RNA degradation and processing, Promoters and regulatory elements, Transcriptional landscape, Differential and global RNA-seq

* Correspondence: k.j.mcdowall@leeds.ac.uk

¹Astbury Centre for Structural Molecular Biology, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

Full list of author information is available at the end of the article

Background

Propionibacterium acnes is a member of the Actinobacteria, a phylum that includes *Streptomyces*, *Mycobacterium* and *Corynebacterium*. It is renowned through its association with acne vulgaris [1], the most common of all human skin diseases, as well as life-threatening diseases [2], such as meningitis [3] and endocarditis [4]. Most recently, evidence has emerged that *P. acnes* infection causes chronic back pain following physical injury [5]. More usually, it is a harmless commensal of the skin, residing ubiquitously in sebaceous follicles [6]. To adapt, colonise and survive, *P. acnes* needs to sense and respond to changes in its natural environment [7], yet almost nothing is known about how the physiology and growth of *P. acnes* is shaped by regulated gene expression. The 2.6 Mbp genome of *P. acnes* strain KPA171202, the first to be sequenced, contains 2,333 annotated genes [8]. However, the organisation and regulation of the transcription units within this clinically important organism remained to be determined experimentally. Indeed, prior to the work reported here, no transcription start sites (TSSs) had been mapped. Other key aspects of gene regulation for which no information was available were mRNA turnover [9], which ensures translation follows programs of transcription, the generation of RNA components of the translational machinery [10,11], and the prevalence of small regulatory RNAs [12]. Recently, the importance of understanding *P. acnes* gene regulation was exemplified by a study that found the ability of different strains of *P. acnes* to cause disease stems from divergence in the expression as well as the content of their genomes [13].

In the first part of this report, we describe improved global and differential RNA-sequencing (RNA-seq) approaches (Additional file 1) and their use in the construction of the first nucleotide-resolution transcriptome maps of *P. acnes* KPA171202, maps that show sites of RNA processing and degradation in addition to sites of transcription initiation. In the second part, we show how analysis of these maps can shed light on many of the key steps in bacterial gene expression and regulation. We describe 'pervasive' transcription, riboswitches, leaderless mRNAs, small non-protein-coding RNAs, post-transcriptional control and unusual tRNA processing. Expansion of these maps to incorporate transcriptional landscapes under different conditions and genetic backgrounds will ultimately reveal how regulated gene expression shapes the physiology and growth of *P. acnes*. As per the original differential (dRNA-seq) approach [14], we distinguished 5' ends corresponding to transcription initiation from those generated through RNA processing and degradation on the basis of their 5'-phosphorylation status. This was done here, however, using tobacco acid pyrophosphatase (TAP), an enzyme that converts a 5'

triphosphate to monophosphate [15], prior to constructing and sequencing cDNA libraries of native 5'-end segments (Additional file 1). The vast majority of primary transcripts are synthesised with a 5'-triphosphate group, while the 5' ends of most 'secondary' transcripts, *i.e.* those generated through RNA processing and degradation, have a monophosphate group [9,16,17]. Thus, an increased number of sequencing reads from a 5' end following TAP treatment is an identifier of a transcription start site (TSS). The original dRNA-seq approach differentiated 5' ends using Terminator™ exonuclease (TEX), which removes 5'-monophosphorylated transcripts [14]. To determine 3' as well as 5' boundaries and the abundance of transcripts, aliquots of some RNA samples were also analysed using a global RNA-sequencing (gRNA-seq) approach (Additional file 1) that avoids the use of PCR [18]. As indicated above, dRNA-seq identifies native 5'-ends only.

Results

Overview of approach

We analysed the transcriptomes of duplicate cultures of *P. acnes* strain KPA171202 grown exponentially in Holland's Synthetic Medium (Additional file 2). This defined medium was chosen as it supports highly reproducible growth, which reduces experimental variation and in turn allows underlying features to be identified with increased certainty. It can also be modified easily to investigate the contribution of individual components to growth and cellular physiology. To assess the sensitivity of our dRNA-seq approach to changes in gene expression, samples were taken following subculture with and without potassium downshift (*i.e.* removal of potassium sources from the medium). We choose this change because of its long-established relevance to skin, *e.g.* potassium levels fluctuate as a result of perspiration [19] and potassium deficiency is associated with dermatoses [20]. Moreover, *P. acnes* genes inducible by potassium downshift were predictable from prior studies of other bacteria [21]. Thus, the differential approach described here used eight cDNA libraries; 2 replicates × 2 conditions × 2 treatments (minus or plus TAP treatment). 3 to 6 million reads were obtained for each library and mapped onto the *P. acnes* genome. Next we counted, for each library, the number of times each genome position was the first nucleotide in sequence reads, *i.e.* associated with a 5' end *in vivo*. Then the reads corresponding to minus and plus TAP treatment were compared using M-A scatterplots for each replicate and condition. Samples from one of the cultures were also analysed by 5' RACE, high-density microarrays and gRNA-seq to allow comparison with more widely used approaches.

Transcription start sites

For each replicate and condition, M-A scatterplots [where $M = \text{Log}_2(\text{reads plus/minus TAP treatment})$, and $A = (\text{log}_2 \text{ plus} + \text{log}_2 \text{ minus})/2$] revealed a population of values that centred close to an M value of 0, corresponding to sites of processing and degradation, and another with higher M values, corresponding to transcription start sites (Figure 1). The envelope of values corresponding to sites of processing and degradation was defined by using the standard deviation of M values in windows of ascending A values [22,23]. By this simple approach we were able to take into account the inherent noise within the dRNA-seq data (Figure 1). 5' ends with M values above the envelope were then identified. This

corresponded to around 6 thousand of the 600 thousand positions in each comparison. To increase the power of our TSS analysis, we did not distinguish between samples from subcultures with and without the potassium downshift. Positions with M values above the envelope in each of the four experiments (2 duplicates \times 2 conditions) were designated positions of transcription initiation, and positions within 8 nt of each other were classified as belonging to the same TSS. With regard to the latter, it is well established that many promoters can initiate transcription at more than one nucleotide position [24]. For the few genes that showed a change in gene expression following potassium downshift, the stringency of the analysis was reduced to the condition

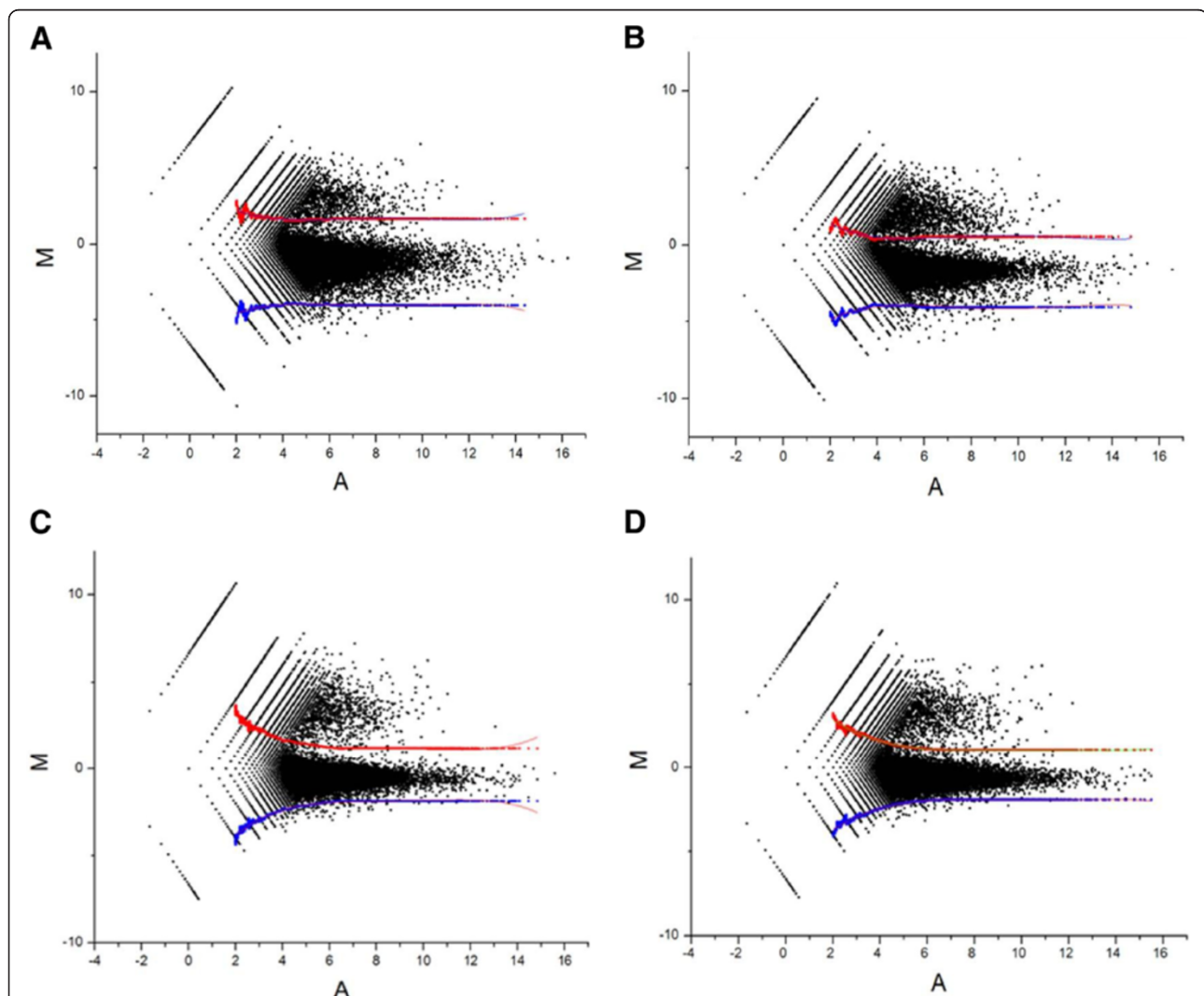


Figure 1 M-A scatterplots of values from the differential RNA-seq analysis. (A) and (B) show data for cells sub-cultured without and with a potassium downshift, respectively. (C) and (D), as (A) and (B), except data is for a duplicate starter culture. The M values correspond to $\text{Log}_2(\text{plus/minus})$ and A values to $(\text{Log}_2 \text{ plus} + \text{Log}_2 \text{ minus})/2$, where minus and plus refer to the number of reads before and after treatment with TAP. For further details, see *Methods*. The red and blue points represent the upper and lower boundaries, respectively, of an envelope that contains sites of processing and degradation. The upper and lower boundaries were defined by the equation $\mu \pm 3\sigma$, where μ and σ are the average and standard deviation, respectively, of M in a moving window of 5,000 data-points sorted in ascending order of A [22,23].

under which transcription could be detected most readily. By this approach we identified 4,058 TSSs (Additional file 3). The probability of a position being outside the envelope in each of four comparisons by chance is less than 1 in 100 million ($6/600$ to the power 4). The latter number vastly exceeding the total of 5' ends identified. Moreover, the identified TSSs were associated with low p-values (on average $6.5E-07$, see Additional file 3) as determined using Rank Products Analysis [25,26], which is a statistical method specifically developed to analyse data that is based on replicated experiments and compares two outputs, in our case RNA-seq reads obtained with and without TAP treatment.

Next, using the UCSC Genome Browser [27], we viewed the positions of the TSSs alongside genes annotated in GenBank [28] and transcription units revealed by gRNA-seq. With regard to the latter, it should be noted that for every position in the genome, we determined the number of times it was read irrespective of its position in fragmented RNA. Moreover, we limited the upper range of the gRNA-seq reads, which gives transcription units a block appearance, to make it easier to decipher their 5' and 3' boundaries. The main gene(s) of interest to us in any view are always depicted left to right. For genes on the reverse strand, the RNA-seq data are given negative values and are shown in red instead of black, and the genome position numbering is reversed. For all classes of functional RNA, mono- and poly-cistronic mRNAs, ribosomal RNAs, transfer RNAs and ubiquitous small RNAs, we were able to identify TSSs that aligned with the 5' boundaries of transcription units revealed by gRNA-seq (for examples, see Figure 2). Such TSSs are represented by black vertical lines. Interestingly, the majority of TSSs we identified did not correspond to the 5' boundaries of obvious transcription units, or at least those of annotated genes (for several examples, see panel B). These TSSs are represented by blue vertical lines. The function of the corresponding transcripts is cryptic with a large proportion being found within coding regions on either the sense or antisense strand. Evidence for 'pervasive' transcription, a widespread phenomenon in eukaryotes [30,31], is emerging in bacteria [14,32-46]. Of the TSSs we identified, 1106 were associated with the 5' boundaries of transcription units corresponding to annotated genes, as illustrated in Figure 2, or produced discrete RNAs of high abundance relative to flanking regions (see section below). The TSSs associated with step increases in transcription are identified within Additional file 3.

Processing and degradation sites (PDSs)

The vast majority of the 5' ends detected by dRNA-seq (~640,000) corresponded to processing and degradation sites (PDSs), which were not enriched following TAP

treatment and were internal to transcription units identified by gRNA-seq. *P. acnes* encodes three endonucleases, RNase E [47], RNase Y [48] and RNase III [49], and a dual endonuclease/5' to 3' exonuclease, RNase J [50,51], (Additional file 4) that could account for the large number of PDSs. The coverage provided by PDSs is sufficient to allow the detection of transcription units and changes in gene expression. This is illustrated here using the *kdp* operon (Figure 3), which encodes an inducible high-affinity potassium transporter (Additional file 2) as well as an associated two-component sensory and regulatory system [21]. The vertical bars correspond to positions at which a 5' end was detected by dRNA-seq following TAP treatment, irrespective of whether or not it was enriched. The bar height indicates the number of times a 5' end was detected by sequencing. Shown are the results of analysing RNA extracted from cells with and without a potassium downshift. A horizontal arrow (in main panel) indicates the position of genes corresponding to the *kdp* operon. Following the downshift, it is clear from the dRNA-seq data that the expression across this entire operon increases substantial. The average number of reads increased by ~100 fold, which is similar to that detected using microarrays and gRNA-seq [52]. Genes with altered gene expression as a result of the downshift have been catalogued based on Rank product analysis of the microarray data (Additional file 5).

Transcription and maturation of stable RNAs

In stark contrast to what has been found for *B. subtilis* [53], which along with *E. coli* is one of the main model systems in which tRNA processing has been studied in detail [11], none of the *P. acnes* tRNA genes (Additional file 6) are part of the three rRNA operons. Another striking difference is that most *P. acnes* tRNA genes are transcribed individually: we only found one example of a tricistronic tRNA operon (Val, GAC; Cys, GCA; Gly, GCC), and two examples of dicistronic operons (Met, CAT; Thr, GGT; and Asp, GTC; Phe, GAA). Thus, co-transcription does not appear to be a major means of regulating stable RNA production in *P. acnes*, unlike the situation in *B. subtilis* [53].

Analysis of our dRNA-seq data revealed processing sites (*i.e.* 5'-end sequences of the downstream products of endonucleolytic cleavage) close to the 3' ends of most *P. acnes* tRNAs (Figure 4), about half of which are encoded by genes with a 3' CCA triplet (Additional file 6). As in the previous figure, all genes are drawn left to right irrespective of whether they are on the forward or reverse strand, and the RNA-seq data for genes on the forward and reverse strand are shown in black using positive values and in red using negative values, respectively. The dRNA-seq data incorporates an additional track showing the results obtained before TAP

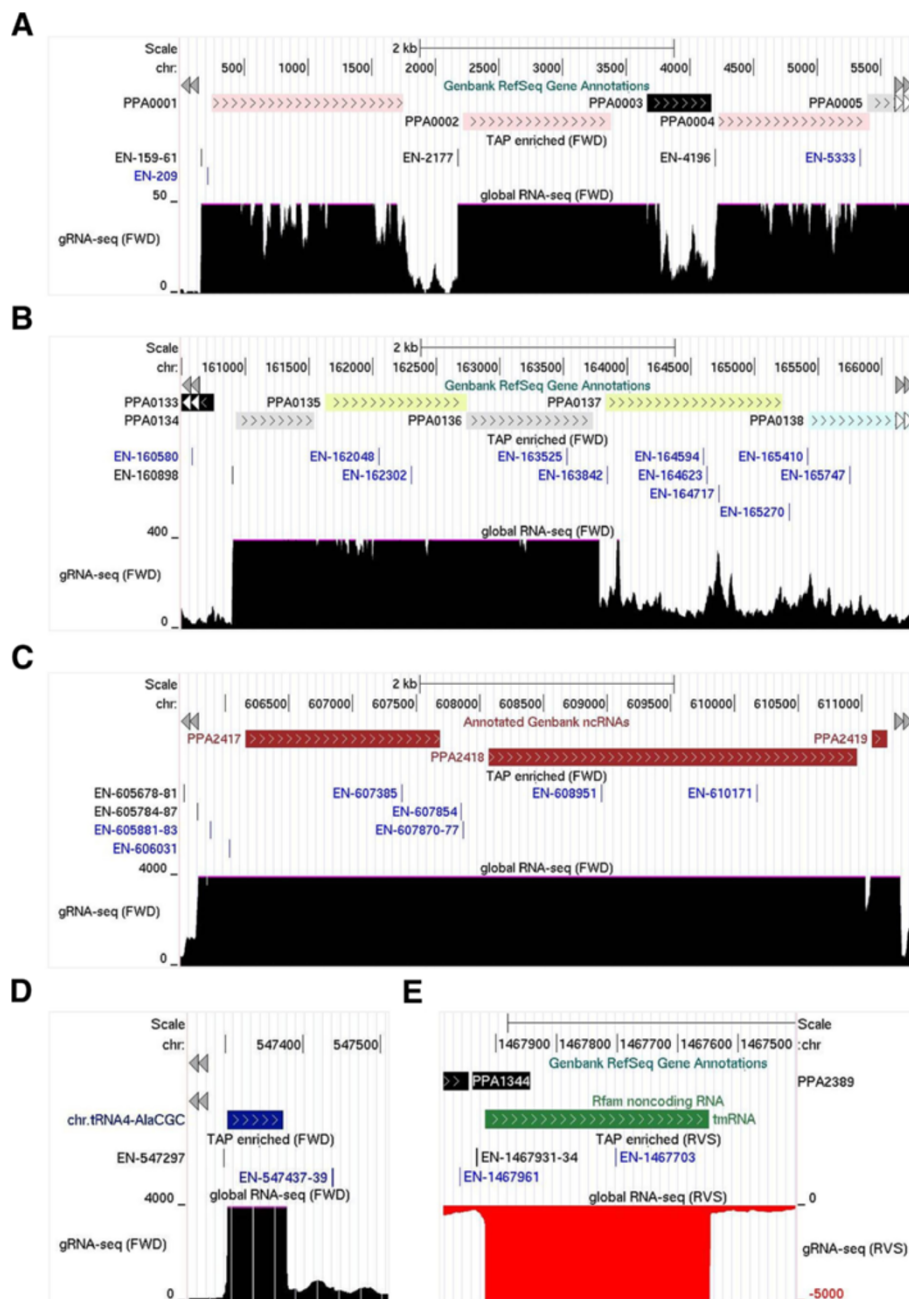
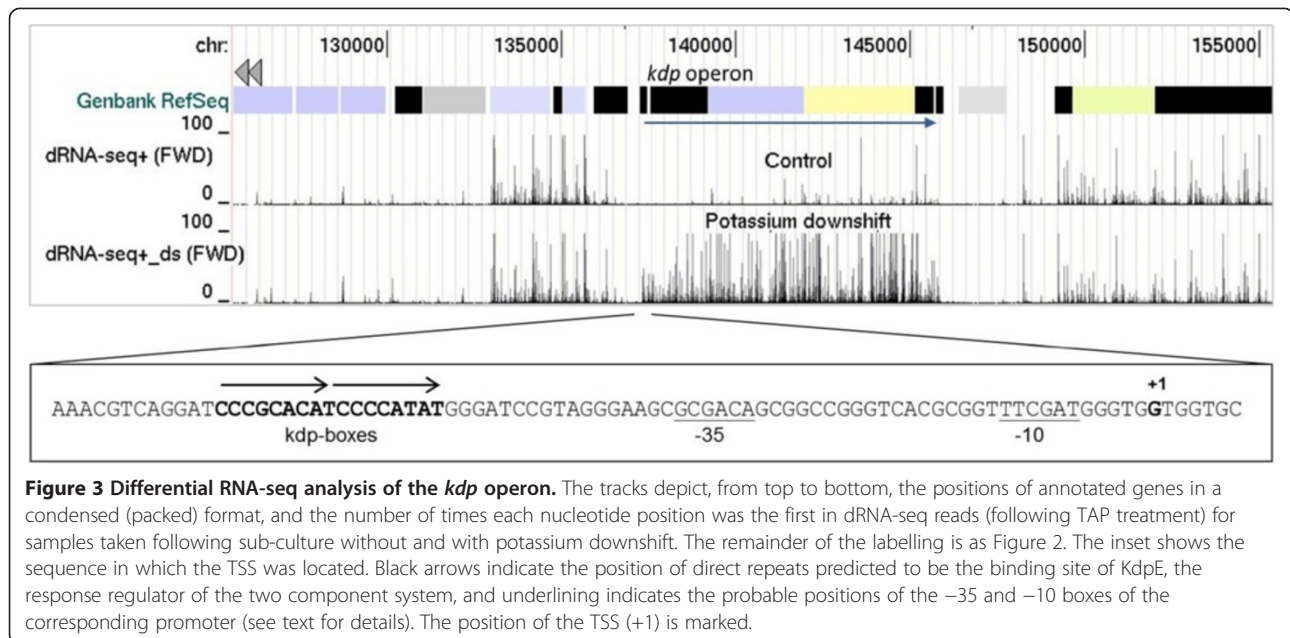


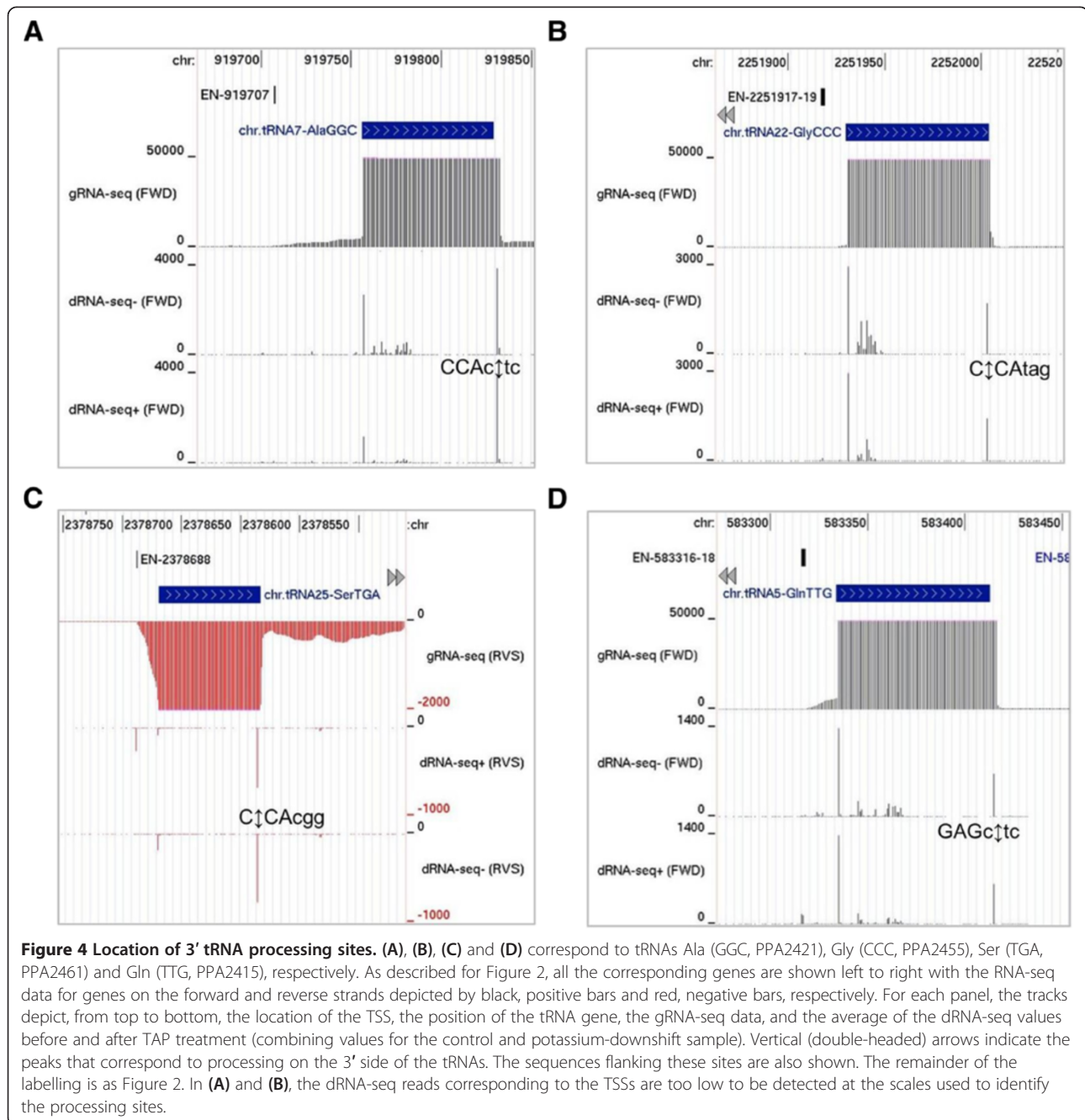
Figure 2 Examples of TSSs associated with different classes of RNA. All the corresponding genes are shown left to right regardless of whether they are encoded on the forward (A), (B), (C) and (D) or the reverse (E) strand. The global RNA-seq data for genes on the forward and reverse strand is coloured black and red, respectively. Furthermore, reads corresponding to the reverse strand have been given negative values. (A), (B), (C), (D) and (E) show data corresponding to monocistronic mRNAs (PPA001, PPA002 and PPA004), polycistronic mRNA (PPA0134-6), ribosomal RNA (PPA2417-2419), transfer RNA (AlaCGC) and sRNA (tmRNA), respectively. The panels are screenshots from the UCSC Microbial Genome Browser [29]. In each panel the tracks depict from top to bottom, the position of annotated genes (protein or RNA coding, as appropriate), the positions of TSSs identified by the analysis of M-A scatterplots (Additional file 3), and the number of times each position on the corresponding strand was sequenced following fragmentation of the transcriptome (gRNA-seq). The numbers at the ends of the RNA-seq tracks indicate the scale of the sequencing reads, while the numbers at the top of each panel indicate the genome position. TSSs in black text were judged by viewing of the gRNA-seq data to be clearly associated with leading edges of transcription, while those in blue text were not. EN indicates that a TSS was enriched following TAP treatment.



treatment. No increase in bar height following TAP treatment is an identifier of a site of processing or degradation. *P. acnes* has a homologue of tRNA nucleotidyltransferase (Additional file 4), which is known to attach CCA to tRNAs that are not transcribed with this motif [11]. A 3' CCA on mature tRNA is essential for several functions related to translation [54]. Closer inspection of tRNAs encoded by genes with a 3' CCA triplet revealed endonucleolytic cleavage a few nucleotides downstream of the tRNA (for an example, see panel A). This fits with current models in which the precursors of these tRNAs are trimmed back by 3' exonucleases to produce a 3' CCA end to which amino acids can be attached [11]. *P. acnes* contains homologues of RNase PH and RNase D, two 3'-5' exonucleases known to be involved in trimming tRNA (Additional file 4). Interestingly, cleavage sites were also identified *within* 3' CCA triplets (for examples, see panel B and C). In these examples, the cleavage was always between the Cs. To our knowledge such processing has not been described previously in bacteria. However, it is well established that at least some members of the tRNA nucleotidyltransferase family can recognise partial CCA ends and add only the residues that are missing [54]. Should *P. acnes* tRNA nucleotidyltransferase mediate this function *in vivo*, cleavages within 3' CCA triplets would not result in terminal inactivation of the tRNA. The identity of the *P. acnes* endoribonuclease that cuts within 3' CCA triplets is open to speculation. However, it is unlikely that it will be the homologue of RNase E as the *E. coli* enzyme has been found to have a preference for cutting bonds within sites rich in A and/or U nucleotides [55]. Whatever the actual identity of the endonuclease, it might also cut

3' to tRNAs encoded by genes that lack a 3' CCA triplet (Figure 4, panel D). *P. acnes* lacks an obvious homologue of tRNase Z, the endonuclease that generates the 3' end for the post-transcriptional attachment of the CCA in *E. coli*, *B. subtilis* and other bacteria [11]. Differential RNA-seq also identified mature tRNA 5' ends (Figure 4), which are generated ubiquitously by RNase P [11].

For each of the three rRNA operons in *P. acnes*, two TSSs were identified upstream of the 16S rRNA gene, an arrangement reported previously for *E. coli* [56-58] and *B. subtilis* [59]. For each operon, we also identified staggered cleavages in each of two pairs of complementary regions that flank 16S and 23S rRNA (Figure 5). These cleavages are likely mediated by the *P. acnes* homologue of RNase III (Additional file 4), which is a well-characterised, double-strand-specific endoribonuclease [60] with a widespread role in the maturation of ribosomal RNA [10]. In addition to sites of RNase III cleavage, we identified sites corresponding to the mature 5' ends of all three ribosomal RNAs and the mature 3' end of 16S rRNA. We also identified sites within one or two nucleotides downstream of the mature 3' end of 23S and 5S rRNA. This indicates that, as found for other bacteria, the maturation of rRNA in *P. acnes* requires the action of multiple endoribonucleases. All of the sites described above were associated with a step change in transcript level as revealed by gRNA-seq (data not shown). Following endonucleolytic cutting, the mature 3' ends of *P. acnes* 23S and 5S rRNA are likely produced by trimming of short 3' tails. Regarding the generation of the mature 5' end of 16S rRNA, *P. acnes* has homologues of both RNase J and RNase E (Additional file 4), ribonucleases that mediate this function in *B. subtilis* and *E. coli*,

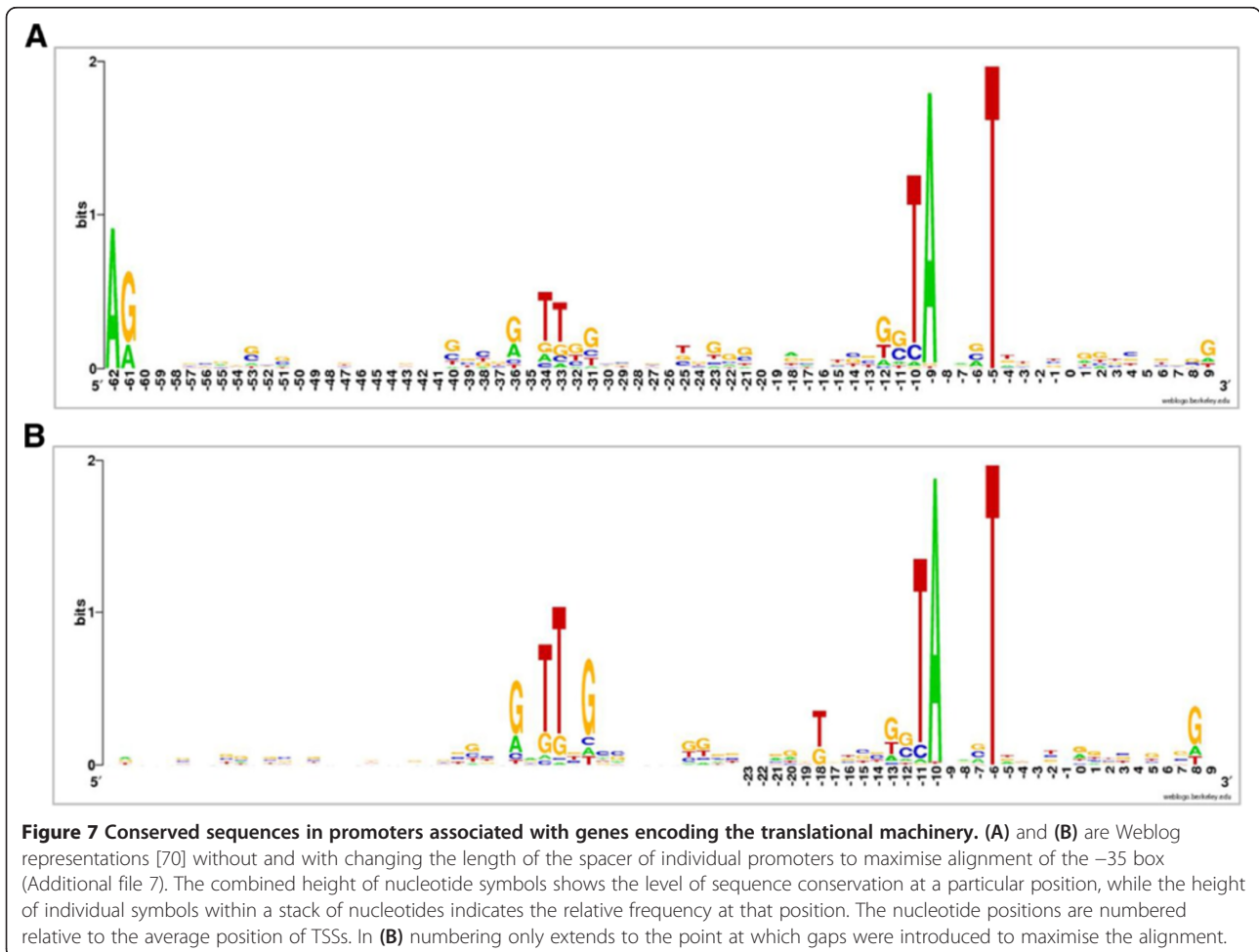
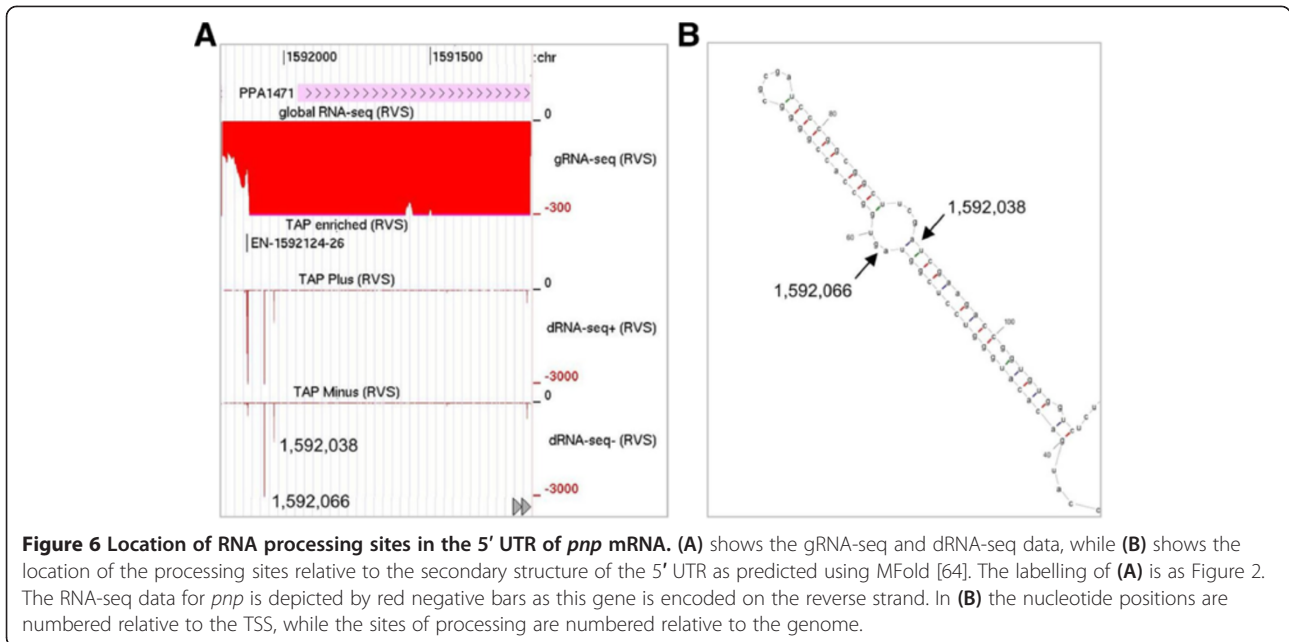


respectively [10]. Interestingly, we also identified multiple cleavage sites within 16S, 23S and 5S rRNA. These may represent steps in controlling the quality of rRNA (and ribosomes) [61], preventing rRNA accumulating in excess over ribosomal proteins [62,63] or simply mediating the normal turnover of the RNA. The nature of cleavages within processed rRNA is being investigated.

Processing within mRNA

We were also able to identify specific cleavage sites in other classes of RNA. For example, we detected staggered

cuts within a base-paired region of the 5' leader of *pnp* mRNA (Figure 6), which encodes a 3'-5' exonuclease (Additional file 4). Cleavage at the equivalent sites in *E. coli* by RNase III has been shown to produce 3' ends that are accessible by PNPase. This facilitates an autoregulatory mechanism that ensures any overproduction of PNPase is only transitory as it leads to increased degradation of *pnp* mRNA [65-67]. This autoregulatory mechanism would appear to be evolutionarily conserved: experimental evidence for its existence in *Streptomyces* spp. has been obtained [68].



TSSs in *Propionibacterium* and related genera that utilise the promoters identified here as a learning set will be presented elsewhere. Consistent with the analysis of the promoters of rRNA, r-protein and tRNA genes (Additional file 7 and Figure 7), the overall level of sequence conservation at the -35 position was considerably lower. Nevertheless, promoter sequences were identified that also matched the single most common sequence variant of the -35 box consensus (GnTTnG). In addition to the promoters of rRNA, r-proteins and tRNA genes, this included the promoters of the genes of translation factors EF-Tu (PPA1873), IF-2 (PPA1493) and IF-3 (PPA1414) and central metabolism enzymes, e.g. alanine dehydrogenase (PPA2274), dihydrolipoamide acyltransferase (PPA0693), uridylate kinase (PPA1519), 3-oxoacyl-(acyl-carrier-protein) reductase (PPA1533), cytochrome d ubiquinol oxidase subunit I (PPA0176), fructose-1,6-bisphosphate aldolase (PPA2024), isopentenyl diphosphate delta isomerase (PPA2115), nitric-oxide reductase subunit B (PPA1975), and polynucleotide phosphorylase (PPA1471). Moreover, these promoters were associated with some of the highest transcript levels (data not shown), consistent with the well-established finding that promoters with matches to a consensus tend to be 'strong' [73].

Uncovering multiple layers of regulation

The identification of TSSs and promoter sequences within the context of high-resolution transcriptome maps provides a much improved platform for assessing the complexity of gene regulation. This is illustrated here using the *P. acnes* homologue of NrdR, a transcription factor that controls the expression of ribonucleotide reductases (RNRs) [74,75], key enzymes that catalyse the formation of deoxyribonucleotides from ribonucleotides [76]. By using MEME to compare sequences positioned -60 to +15 relative to TSSs mapped for *nrdR* and genes encoding components of RNRs, we were able to identify probable binding sites for NrdR (herein referred to as nrd-boxes). These binding sites overlapped some, but not all of the identified promoters: a pair of nrd-boxes overlapping the distal of the two promoters for the *nrdRJ* operon and a single nrd-box overlapping the *nrdAB* promoter (Figure 8). Moreover, after constructing a position-weight matrix and scanning the entire genome of *P. acnes* using PREDetector [77], we identified another pair of nrd-boxes far downstream of the *nrdDG* promoter active under the growth conditions used in this study. Thus, it appears that the transcription of *nrdR* and several of its targets are under the control of multiple promoters, only some of which are regulated by NrdR.

We also found evidence of post-transcriptional control: much of the transcription of the *nrdAB* operon (encoding subunits of a RNR) appears to terminate before the first structural gene (Figure 8). Consistent with this

interpretation, the 5' UTR region of *nrdAB* is annotated as containing a cobalamin riboswitch [78], a *cis*-regulatory element that is widely distributed in the 5' UTRs of cobalamin (vitamin B12)-related genes in bacteria [79-82]. Interestingly, the 5' UTR of *nrdDG* (encoding an anaerobic RNR and activating protein), but not the *nrdRJ* operon is also annotated as containing a cobalamin riboswitch. Furthermore, the activity of the RNR encoded by *nrdJ*, which is co-transcribed with *nrdR*, is cobalamin dependent [83], and we detected expression of cobalamin biosynthetic genes (data not shown). The above suggests that cobalamin is present at sufficient levels to activate the riboswitches, which in turn down regulates the expression of *nrdAB* and *nrdDG*. The result is that much of the RNR production is via *nrdJ*.

Our results also lead us to propose that the NrdR repressor is active under the conditions used for this study. The bulk of the transcription of *nrdRJ* appears to initiate at the distal promoter, not the proximal promoter overlapped by a pair of nrd-boxes, and we did not detect transcription initiation in the immediate vicinity of the pair of nrd-boxes located upstream of *nrdDG*. We did detect relatively high levels of transcription from an *nrdAB* promoter, but this is overlapped by only a single nrd-box. We speculate that should a promoter exist in the vicinity of the nrd-boxes located upstream of *nrdDG* inactivation of NrdR will produce a transcript lacking a functional riboswitch, thereby removing the cobalamin regulation. The RNR encoded by the *nrdDG* operon is thought to function under anaerobic conditions.

Comparison with standard 5' RACE

Prior to undertaking RNA-seq analysis, we had initiated a study of a two-component system that is likely involved in *P. acnes* quorum sensing [84]. Unusually, the genes of the histidine kinase (*pqsA*, PPA0945) and response regulator (*pqsC*, PPA0947) are divergently transcribed, and *pqsA* is preceded by a gene (*pqsB* PPA0946) that is predicted to encode an extracellular signalling peptide (Figure 9, panel A). Analysis of the transcription of this locus using standard 5' RACE revealed single promoters upstream of both *pqsB* and *pqsC* (Figure 9, panel B). RNA-seq revealed much more: it not only identified both of the TSSs identified by 5' RACE (to the precise nucleotides), it identified a second TSS that was associated with a clear step increase in transcription upstream of *pqsB* and identified a potential antisense RNA (with an associated TSS, EN-1027593) overlapping the 5' end of *pqsB* (Figure 9). Both of these new elements have now been incorporated into a continuing dissection of the *pqs* locus. Additional TSSs were also detected by RNA-seq upstream of *pqsC*; however, their contribution to the expression of this gene is more difficult to assess due to background transcription. The TSS of *pqsB*

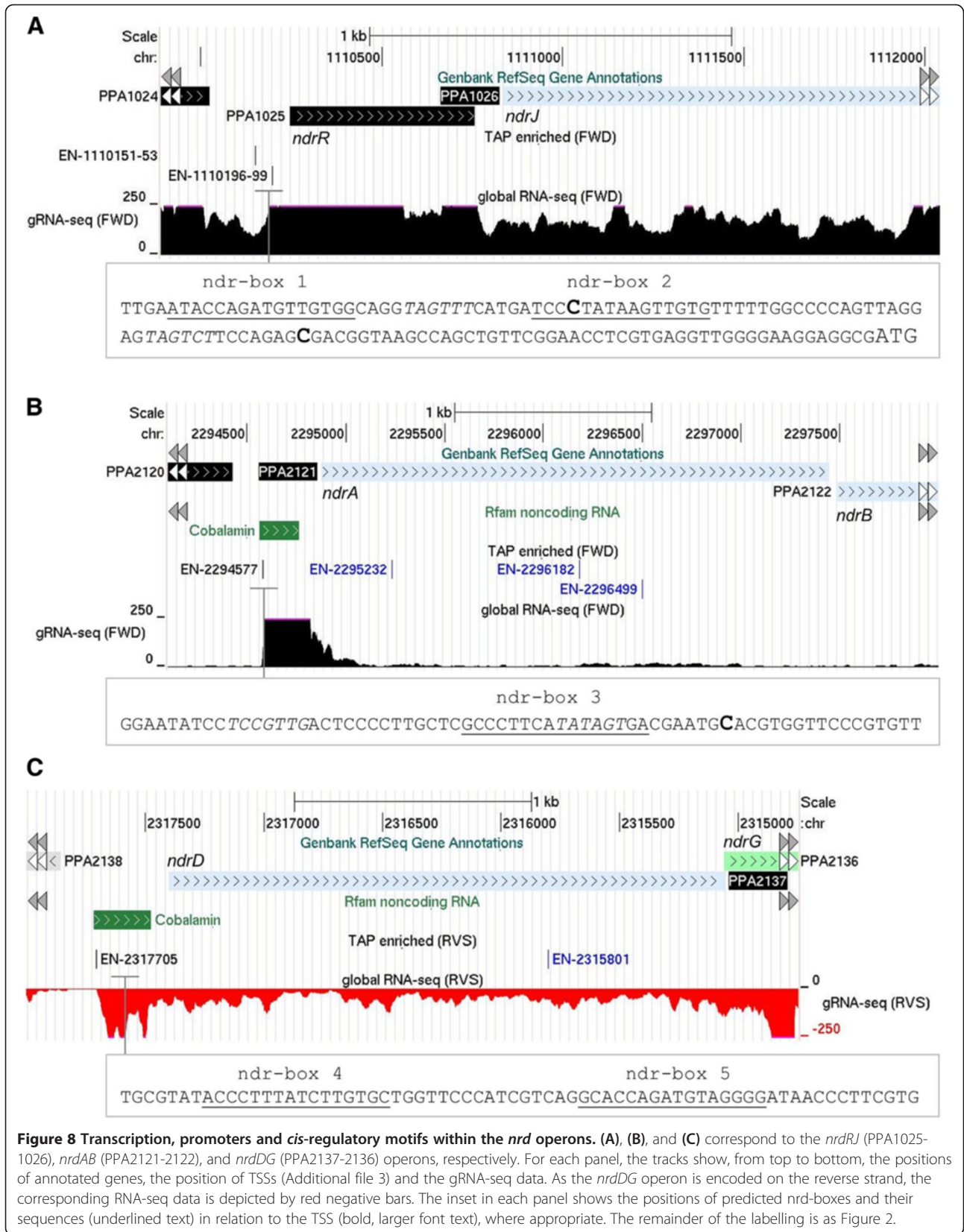


Figure 8 Transcription, promoters and cis-regulatory motifs within the ndr operons. (A), (B), and (C) correspond to the *ndrRJ* (PPA1025-1026), *ndrAB* (PPA2121-2122), and *ndrDG* (PPA2137-2136) operons, respectively. For each panel, the tracks show, from top to bottom, the positions of annotated genes, the position of TSSs (Additional file 3) and the gRNA-seq data. As the *ndrDG* operon is encoded on the reverse strand, the corresponding RNA-seq data is depicted by red negative bars. The inset in each panel shows the positions of predicted ndr-boxes and their sequences (underlined text) in relation to the TSS (bold, larger font text), where appropriate. The remainder of the labelling is as Figure 2.

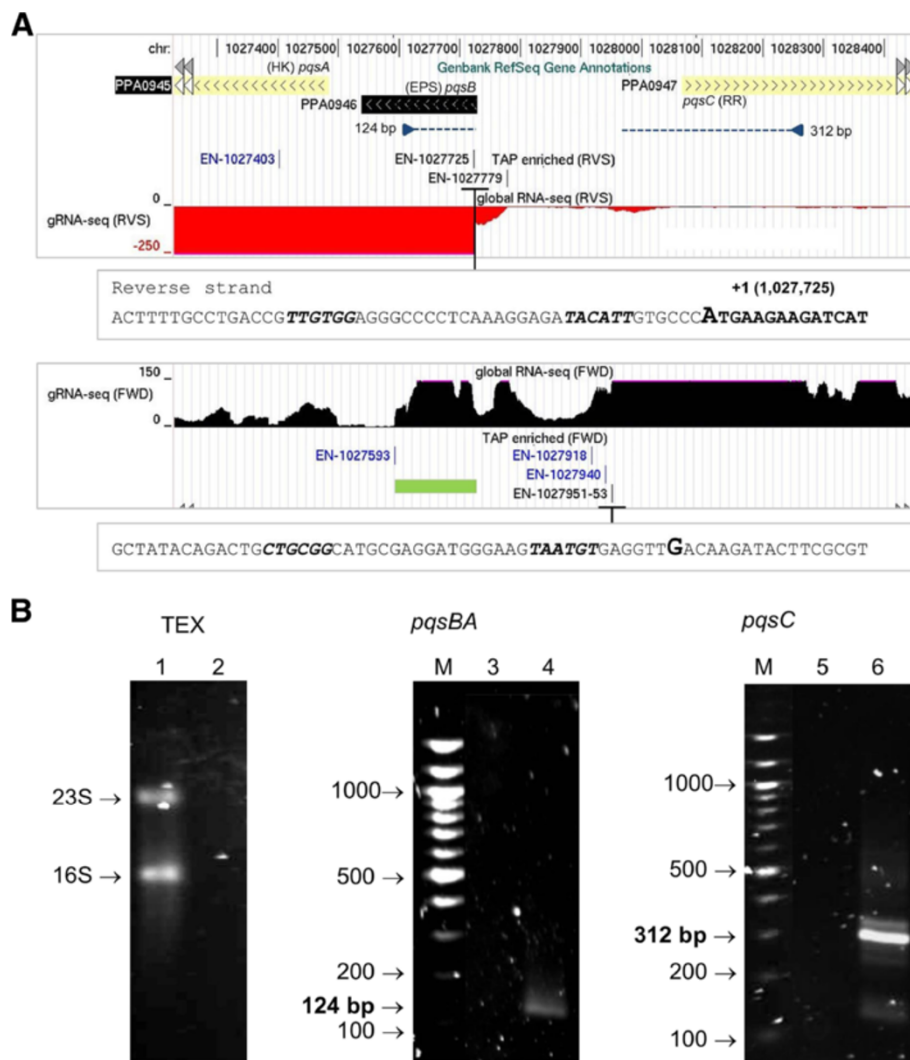


Figure 9 RNA-seq analysis of the *pqs* locus. In (A) the tracks are as detailed in Figure 8. For the insets below the gRNA-seq tracks, large bold font indicates TSSs identified by standard 5' RACE (B), and dRNA-seq (Additional file 3). The putative -10 and -35 boxes of the corresponding promoters are in italic font. The genes of the histidine kinase, response regulator extracellular and signalling peptide are PPA0945, PPA0947 and PPA0946, respectively. RNA-seq data for both strands within the *pqs* locus is shown: black positive values correspond to the forward strand, while red negative values correspond to the reverse strand. The green box marks the boundaries for a potential antisense RNA regulator of the PPA0945-0946 transcription unit. Blue arrowheads indicate the binding sites of the primers that were extended to produce cDNA for *pqsBA* and *pqsC* mRNA. The sequences of the corresponding primers were 5' CCTTGGGTTGCGTATTACAG and 5' TGTGAGACCGTCCATTTTCAG, respectively. The dashed lines indicate the sizes of the cDNAs that were produced, as determined by agarose gel electrophoresis (B) and sequencing of the product (data not shown). (B), lanes 1 and 2 show total RNA before and after treatment with TEX, respectively; lanes 3 and 4 show the reaction products of 5' RACE analysis of *pqsBA* mRNA without and with template, respectively; lanes 5 and 6, as 3 and 4, except *pqsC* mRNA was analysed. The marker (M) was a 100-bp ladder (BioLabs). For details of primers and conditions, see *Methods*.

overlaps the start codon to produce a leaderless mRNA, many more examples of which are described below.

Identification of potential sRNAs

As indicated above, we identified a number of TSSs that were associated with relatively short transcripts with high abundance relative to transcription in the surrounding regions (Additional file 8). These include all the ubiquitous bacterial sRNAs (6S RNA, tmRNA, the

RNA component of RNase P, and Signal Recognition Particle RNA) [12] and 5' UTRs with and without structural similarity to known riboswitches (e.g. TPP, FMN and Cobalamin) [85]. To be classified as a 5' UTR, the peak of RNA abundance had to fall sharply at the boundary with the coding region of the downstream gene. Such a profile could be the result of riboswitching [85], attenuation-like mechanisms [86] or simply the higher stability (i.e. reduced susceptibility to the mRNA

turnover machinery) that is often associated with structured regions [87]. An example of a 5' UTR that has not, as far as we are aware, been linked with known riboswitches is shown in Figure 10 (panel A). Like all the potential 5' UTRs that we have identified here, this is predicted to be structured (panel B). We also identified a few potential *cis*-acting antisense regulatory RNAs that are complementary to the ribosome-binding regions of overlapping mRNA and may thereby regulate the initiation of translation [12]. A complication is that many RNAs that are antisense to ribosome-binding regions could equally be the 5' UTR of a divergent protein-coding gene. An example of a potential antisense RNA complementary to a 5' ribosome-binding site (RBS) is shown (panel C), along

with one that might interact with a RBS internal to a polycistronic mRNA (panel D). We also identified others that are complementary to sites distal to a RBS (for example, see panel E). It has recently been shown that antisense RNA binding to internal sites can initiate mRNA degradation via the stimulation of RNase E cleavage *in vitro* [88]. As mentioned previously, *P. acnes* has a homologue of this key endonuclease (Additional file 4). Small RNAs were also identified, some of which are highly abundant, that do not appear to overlap with protein-coding genes (for an example, see panel F). Our results indicate that, as is being found increasingly in other bacteria [12], sRNAs are likely to a significant role in regulating gene expression in *P. acnes*.

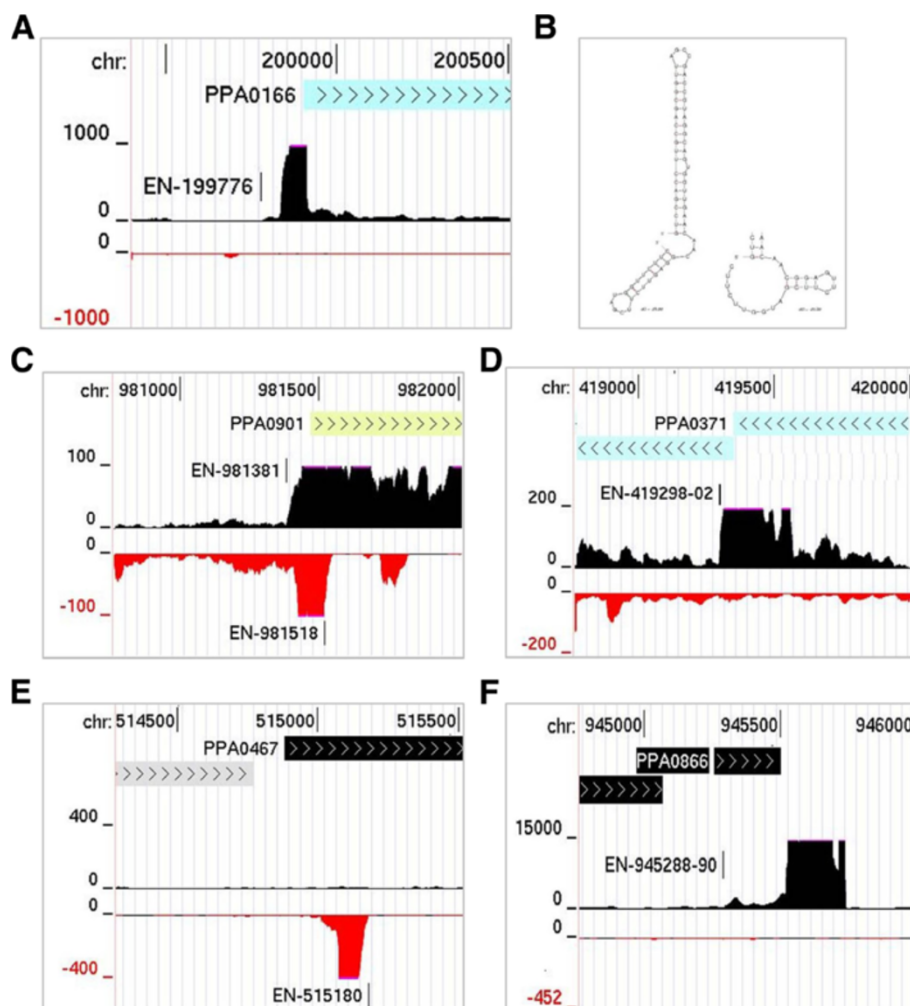


Figure 10 Examples of *P. acnes* small RNAs. (A) and (B) show the transcription and structure, respectively, of the 5' UTR of PPA0166, which encodes a putative L-lactate permease. Alternative folding is shown bottom right of (B). (C), (D), and (E) correspond to examples of potential antisense RNAs that bind 5' RBSs, internal RBSs, and sites internal to the coding region, respectively. (F) shows an example of highly abundant sRNA that does not appear to overlap mRNA. For each of these panels, the tracks show the positions of annotated genes, the position of TSSs (Additional file 3) and the gRNA-seq data for both strands (black positive values correspond to the forward strand, while red negative values correspond to the reverse strand). The remainder of the labelling is as Figure 2.

Leaderless mRNAs

While mapping TSSs, we noticed several examples that coincided with start codons for translation. This prompted us to gauge the prevalence of leaderless mRNAs in *P. acnes*. The start codons of protein-coding genes, as annotated in the NCBI database [89], were collated and their positions mapped against TSSs associated with a step increase in transcription (Additional file 3). This revealed 50 instances of annotated start codons overlapping TSSs and another 88 where start codons followed TSSs within 10 nt (Additional file 9). The latter TSSs produce a 5' leader that is generally considered too short to recruit ribosomes via the canonical Shine-Dalgarno interaction, which requires base pairing between the leader and a complementary sequence in the 3' end of 16S rRNA [90,91]. We also identified 15 instances of mRNAs with relatively short 5' leaders (<20 nt) within which we were unable to detect a Shine-Dalgarno sequence using RBSfinder [92] (Additional

file 9). From the above, we concluded that in sharp contrast to what has been found for nascent mRNAs in *E. coli* [93], translation initiation in the absence of a Shine-Dalgarno interaction appears to be prevalent in *P. acnes*. Analysis of the ontology of genes associated with 'leaderless' mRNA failed to identify enrichment of a particular function(s). This is in contrast to the situation in *E. coli* where repressors of mobile genetic elements are clearly enriched (Romero and McDowall, unpubl. results). Pyruvate kinase, a key glycolytic enzyme, and ferritin, the major store of iron, are two examples of products translated from leaderless mRNA in *P. acnes* (Figure 11, panels A & B). The proportions of AUG, GUG, UUG start codons present in *P. acnes* leaderless mRNA are 68%, 29% and 3%, respectively. This is similar to the proportion reported for another actinomycete, *Streptomyces coelicolor* [95]. To our knowledge, the mechanism by which leaderless mRNA is translated in actinomycetes has not been determined. Very recently it has been

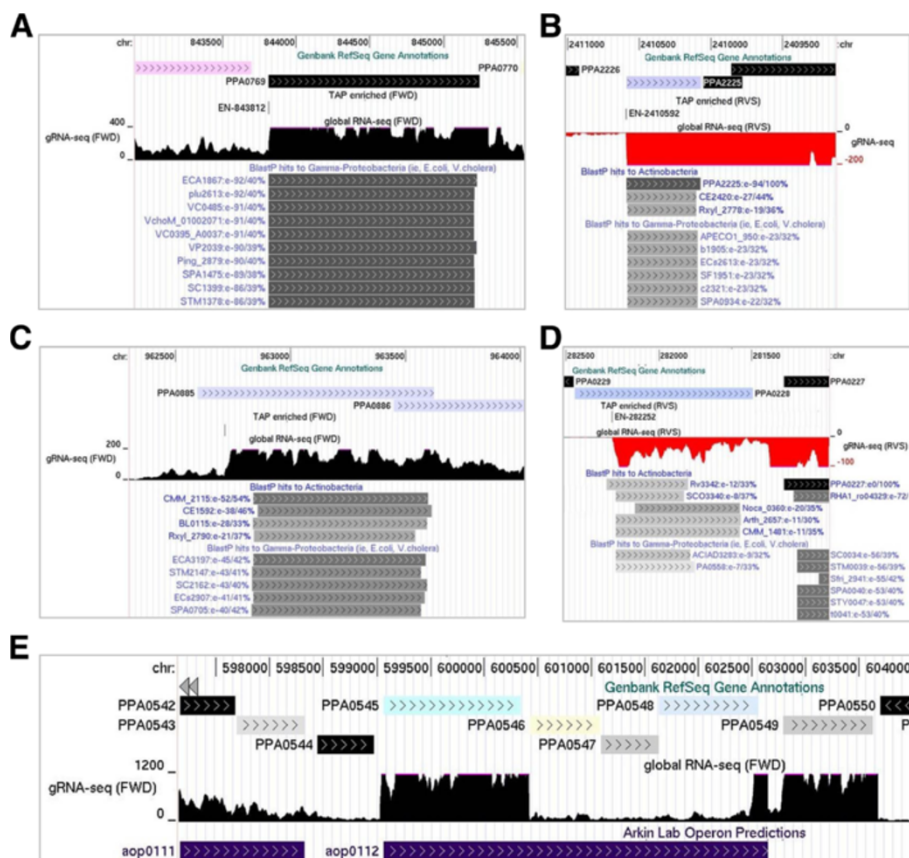


Figure 11 Leaderless mRNAs and transcripts of genes requiring reannotation. (A) and (B) correspond to examples of leaderless mRNA. PPA0769 encodes pyruvate kinase, and PPA2225 ferritin. (C) and (D) correspond to examples of genes requiring reannotation. PPA0885 encodes hydroxyethylthiazole kinase, and PPA0228 a putative methyltransferase. For each of these panels, the tracks show, from top to bottom, the positions of annotated genes, the position of TSSs (Additional file 3), gRNA-seq data, and BlastP hits or operon predictions. The remainder of labelling is as Figure 2. (E) shows the transcription data for *eno* (PPA0545), which encodes phosphopyruvate hydratase. Purple boxes indicate computational predictions of the operon structure in this region [94]. As in previous figures, all the genes are drawn left to right with their corresponding RNA-seq data (black positive values for forward strand, red negative values for the reverse strand).

shown that leaderless mRNAs in *E. coli* can be generated post-transcriptionally by a stress-induced mRNAse that is the toxic component of a toxin-antitoxin system [96]. Should such processing also exist in *P. acnes*, the detection of the corresponding sites would require us to add a phosphorylation step to facilitate the cloning and sequencing of the 5'-hydroxylated downstream products of cleavage by toxin mRNAases.

Re-annotation of protein-coding genes and operon structures

We also noticed that a significant proportion of TSSs associated with significant step increases in transcription were internal to the 5' half of annotated genes (for examples, see Figure 11, panels C & D) suggesting that the actual gene might be shorter. Consistent with this notion, we have been able to find RBS along with appropriately spaced start codons downstream of many of these TSSs (Additional file 10) and homologous genes that lack sequences matching the 5' end of the original gene annotation (panels C and D). Our combined RNA-seq approach also revealed many examples of operon structures that differ significantly from bioinformatics predictions (for example, see panel E). This was not particularly surprising; it is known that even the best bioinformatic approaches are not completely accurate [97]. Nevertheless, achieving accurate information on gene and operon structures is essential for gene expression and regulation to be modelled [24] at the level of the whole cell [98]. Our transcriptome approach and data should hasten the achievement of this goal for *P. acnes*.

Discussion

Here we describe a number of mechanistic insights gained from an improved dRNA-seq approach that distinguished sites of transcription initiation without erasing the secondary transcriptome using TEX, an enzyme that in our hands can degrade a substantial proportion of 5'-triphosphorylated RNA under the conditions recommended by the vendor (Additional file 11). With the advent of sequencing techniques that can provide in excess of 100 million reads (e.g. Illumina Solexa) there is now no need to erase the secondary transcriptome in order to detect transcription start sites. We simply used TAP [15] to distinguish tri- from mono-phosphorylated 5' ends. This enzyme was used in earlier dRNA-seq approaches, but to facilitate the cloning of 5'-fragments remaining after TEX treatment [14,95]. Other improvements were to fragment the RNA after the addition of the 5' adaptor to improve the efficient cloning of 5' ends from large transcripts, and to combine with a gRNA-seq approach that does not require an amplification step [18]. The latter – used here for the first time to study bacterial transcription – allowed us to identify the

3', as well as 5', boundaries of transcripts. It should be particularly beneficial for the study of organisms with a high GC content in their genomes, as the amplification of segments of such genomes (or transcriptomes) is prone to PCR bias and artefacts [99]. Moreover, as we included biological replicates in our dRNA-seq approach and incorporated a statistical method that is ideally suited to the analysis of dRNA-seq data, we are confident that our analyses are not dominated by false positives. This is particularly relevant with regard to our finding that the majority of TSSs in *P. acnes* are not associated with step increases in transcription that continued into annotated genes or produced discrete RNAs of high abundance relative to flanking regions (Additional file 3).

Evidence for pervasive transcription on a genome scale has previously been obtained for several bacteria [14,32-46,100]. However, this has largely been in the form of the identification of transcripts antisense to annotated genes [35]. By including biological replicates and mapping TSSs in addition to transcripts, our study shows that pervasive transcription in *P. acnes* stems as much from transcription of the coding strand as the non-coding strand of annotated genes. Of the thousands of TSSs identified within annotated genes (and not associated with obvious transcription of a flanking gene), approximately half produced transcripts sense to the coding strand. The only strong bias we have identified so far is for the leading strand of replication (data not shown). Our interpretation is that clashes with oncoming DNA replication hinder pervasive transcription initiation on the lagging strand.

The initiation of pervasive transcription is not random. MEME analysis of sequences immediately upstream of TSSs associated with pervasive transcription revealed that a high proportion (61%) matched the single most common sequence variant (TAnnnT) of the -10 box of *P. acnes* vegetative promoters (data not shown). It is our view that much of the pervasive transcription observed in bacteria is a consequence of the ability of RNA polymerases to recognise a range of sequences, which means that these enzymes can initiate transcription from sub-optimal sites albeit at reduced frequency. Just because a region is transcribed (i.e. active) does not mean that it has a function [101]. Nevertheless, pervasive transcription may be of evolutionary significance by facilitating the transcription of genes acquired horizontally and in turn the production of products that could confer a selective advantage. Pervasive transcription might also have been the source of *bona fide* sRNA regulation. Through spontaneous mutation of promoter regions upstream of TSSs or the acquisition of mobile elements such as REP sequences, which are known to stabilise transcripts [87], small RNAs could increase in abundance and their likelihood of being subject to selection.

Very recent work suggests that the folding architecture of the chromosome influences the location of pervasive transcription [102]. This is consistent with a model in which pervasive transcription is largely driven by the ability of RNA polymerases, given access, to initiate transcription from sub-optimal sites albeit at reduced frequency. Nevertheless, the growing evidence for pervasive transcription in bacteria challenges the dogma that transcription starts and ends just before and after a gene. Moreover, it seems likely that mechanisms exist to minimise the accumulation of pervasive transcripts. One means of controlling pervasive transcripts is likely to be RNA degradation. It is well established that RNA turnover is particularly rapid in the absence of protection by RNA-binding proteins or translating ribosomes and appropriately located structures [9,17,103,104]. In other words, the transcriptional landscape may be defined by the stability of bacterial transcripts more than is currently appreciated.

In addition to providing additional evidence for pervasive transcription, we have identified over a thousand TSSs and associated transcripts encompassing all classes of functional RNA, mono- and poly-cistronic mRNA, ribosomal RNA, transfer RNA, and ubiquitous small RNAs (Figure 2). This alone is an important milestone along the route to understanding the cellular workings of *P. acnes*. In addition, we have identified changes in gene expression resulting from potassium downshift (Figure 3), post-transcriptional steps (Figures 5 and 6) including unusual 3' processing of tRNA (Figure 4), features of vegetative promoters (Figure 7), potential transcription factor-binding sites (Figures 8 and 9), functioning riboswitches as well as a number of sRNAs (Figure 10), and an abundance of leaderless mRNAs (Figure 11). We have also shown how knowledge of the above can be used to build models of gene regulation that inform experimental investigation (Figures 8 and 9). Simply knowing the position of a TSS narrows the search area for transcription factor-binding sites. For example, inspection of the promoter of the *kdp* operon identified a direct repeat overlapping the -35 region (Figure 3, inset), which is characteristic of the binding sites of the response regulators of two-component systems [105].

The prevalence of leaderless mRNAs in *P. acnes* is in stark contrast to the situation in *E. coli*, the main bacterial system in which the translation of leaderless mRNA has been studied [106,107]. Indeed, only two examples of *E. coli* leaderless mRNA have been widely reported, the *cI* repressor of bacteriophage lambda [108] and the *tetR* repressor of transposon Tn1721 [109]. The association between leaderless mRNA and repressors within mobile genetic elements in *E. coli* has been extended to the repressors of the Rac, e14 and Qin prophages by our

own deep RNA-seq analysis of the *E. coli* transcriptome (unpubl. result). We speculate that some aspect of the translation of these leaderless mRNAs may be important in linking the mobilisation of the corresponding genetic elements and the physiological of their host. Recently, it has been shown that stress induces the production of specialised ribosomes that selectively translate a group of mRNAs made leaderless by MazF [96], an endoribonuclease of a toxin-antitoxin (TA) module. Like their mRNA targets, the specialised ribosomes are produced by MazF cleavage, which removes 43 nt from the 3' end of *E. coli* 16S rRNA [96]. Intriguingly, we have mapped a processing site 53 nt from the 3' end of *P. acnes* 16S rRNA (data not shown). This raises the possibility that specialised ribosomes, similar to those generated by MazF in *E. coli*, could mediate much of the translation in *P. acnes*. However, unlike the situation described for *E. coli* [110-112], the translation of leaderless mRNA in *P. acnes* does not appear to require that the start codon is AUG. As described above, a significant proportion of the leaderless mRNA in *P. acnes* have GUG (29%) or UUG (3%) start codons in place of AUG (68%) (Additional file 9).

Regardless of the actual mechanism by which leaderless mRNAs are translated, our study adds to a growing body of evidence that leaderless mRNAs are prevalent outside *E. coli* and its closest relatives and the notion that the mechanism of their translation may represent an ancient milestone in the evolution of gene expression [106,107]. A gene ontology analysis of leaderless mRNA in *P. acnes* revealed a wide distribution of cellular roles (data not shown). Thus, since the emergence of translation mediated by a Shine-Dalgarno interaction, there does not seem to have been divergence in terms of cellular functions that are dependent on leaderless translation in *P. acnes*. It will be interesting to establish for *P. acnes* whether there is a correlation between leaderless translation and the level of gene expression, as measured by protein levels, or the response of genes under conditions of stress or both.

We estimate that around 250 reads (gRNA-seq) correspond to 1 transcript per cell, assuming that the levels of RNase P and tmRNA are similar to those in *E. coli* [113-116]. This appears to be reasonable; our estimate yields 60,000 ribosomes per cell, which is within the range reported for *E. coli* [117] and other bacteria, when applied to the 15 million reads for *P. acnes* 5S rRNA. Many of the mRNAs we identified were associated with less than 250 reads. This does not necessarily mean that not every cell expresses the corresponding gene, only that it was expressed for a proportion of the *P. acnes* cell cycle. Moreover, many mRNAs will not need to be continuously present to ensure translation produces sufficient protein for the daughter cell at division, given that the doubling time is relatively long (~6 hours). We estimate

that transcription associated with pervasive initiation corresponds on average to 0.06 transcripts per cell (15 reads). This relatively low level is perhaps consistent with a background of sporadic transcription.

Conclusions

In summary, by combining the differential and global RNA-sequencing approaches used here to analyse *P. acnes*, it is clear that major step increases in both knowledge and understanding of gene organisation and regulation can be obtained for any bacterial species, which in turn would inform experimental investigation, as illustrated using the *nrdr* operon and *pqs* locus, as well as further computational approaches. With regard to the latter, our data can be mined further, for example, to identify *cis*-regulatory signals that control transcription termination, initiation from the promoters of genes with shared cellular function, RNA processing and degradation, and to predict the potential structures and targets of small RNAs. Our RNA-seq approach can also be used to identify and compare genes expressed under particular conditions. There is excellent agreement between the microarray and gRNA-seq data [52] and we are currently comparing the expression profiles obtained here for cells grown in liquid culture with those of cells grown as a biofilm (Lin and McDowall, unpublished data) as well as cells grown by others in a rich complex medium [13]. With continuing annotation, it should be possible eventually to computationally model the *P. acnes* cell and its interaction with the environment and animal hosts [98].

Methods

P. acnes and its cultivation in defined media

Propionibacterium acnes strain KPA171202 was obtained from Ulm University, Göttingen, Germany [8] and cultivated in an anaerobic workstation (MACS-MG-1000, Don Whately Scientific) at 34°C under 80% [v/v] N₂, 10% [v/v] CO₂, and 10% [v/v] H₂. All analyses were done using cells cultivated without shaking in 100 mL of modified Holland Synthetic Medium (HSM) [84,118] in a 250-mL Erlenmeyer flask. Inocula were prepared in two stages. First a single colony isolated from the surface of reinforced clostridial agar [119] was used to inoculate 10 mL of tryptone-yeast extract-glucose (TYG) broth [120] in a 30 mL plastic universal bottle. After growing to stationary phase, an aliquot was used to inoculate 100 mL of TYG broth to an OD₆₀₀ of 0.2. The culture was then incubated to an OD₆₀₀ of 1.0, after which cells were harvested by centrifugation (3,000 × *g* for 20 min) and washed by first resuspending in 10 mL of HSM (pre-warmed to 34°C) and then repeating the harvesting step. Finally, the cells were resuspended in 10 mL of pre-warmed HSM and an appropriate aliquot was used to inoculate 100 mL of pre-warmed HSM to an OD₆₀₀

of 0.2. To study the effects of a potassium downshift (depletion from growth media), a 100 mL culture of *P. acnes* was prepared as above and grown to an OD₆₀₀ of 1.0, after which the culture was separated into two equal halves and cells harvested as described above. One half was washed using standard HSM, then used to inoculate 100 mL of fresh HSM and reincubated. The other half was processed in the same way, except the HSM lacked potassium (KH₂PO₄ and K₂HPO₄). After 1 h of reincubation, 12.5 mL of stop solution (95% [v/v] ethanol; 5% [v/v] phenol) was added to inhibit cell metabolism [121], and the cells were harvested by centrifugation. When necessary, cell pellets were stored frozen at -80°C.

Isolation of bacterial RNA

Cell pellets of *P. acnes* were resuspended in Kirby mix [122], 1 OD₆₀₀ unit per 100 μL, and transferred to Lysing Matrix B tubes containing fine silica beads (MP Biomedical). Tubes were then placed in a high-speed benchtop homogenizer (Fastprep-24, MP Biomedical; set at 6.5 M/s). Cells were lysed by three cycles of homogenisation for 1 min with cooling between each cycle in an ice-water bath for 1 min. Lysates were extracted using an equal volume of acidic phenol: chloroform: isoamyl alcohol (50: 50: 1) and then chloroform: isoamyl alcohol (49: 1). Nucleic acid in the aqueous phase was precipitated by adding NaCl to 150 mM and 2.5 × volumes of 100% [v/v] ethanol, chilling at -20°C for 1 h, and then harvested by centrifugation (13,000 × *g* for 30 min at 4°C). Nucleic acid pellets were washed twice with 700 μL of 70% [v/v] ethanol, air dried for 5 min and resuspended in RNase-free water. To remove contaminating DNA, samples were treated with DNase I using conditions described by the vendor (Ambion) and extracted with phenol: chloroform as described above. The concentration and integrity of RNA samples were determined using a NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific) and agarose gel electrophoresis [121], respectively.

Transcriptome analyses

The differential RNA-seq data was generated by vertis Biotechnologie AG (Germany) as a service that included the construction of cDNA libraries before and after treatment with TAP (Additional file 1), sequencing of libraries using an Illumina HiSeq platform (single end, 50-bp read length), and the alignment of sequences to the genome (NCBI, accession number AE017283). It should be noted that the 5'-sequencing adaptor was ligated to transcripts prior to fragmentation, thereby allowing the 5' ends of both long and short transcripts to be detected. RNA was fragmented using a Bioruptor® Next Gen UCD-300™ sonication system (Diagenode),

then tailed at the 3' end using poly(A) polymerase (New England BioLabs), copied into cDNA using M-MLV reverse transcriptase (RNase H minus, AffinityScript, Agilent) and an oligo-dT primer, amplified by PCR and fractionated using gel electrophoresis. Fragments of 250–500 bp were selected for Illumina sequencing. Reads were trimmed of 5' adapter and poly(A) sequences and mapped using the CLC Genomics Workbench and standard settings. Prior to differential RNA-seq, samples were enriched for mRNA using *MICROBExpress*[™]-Bacteria beads, as described by the manufacturer (Ambion).

Global transcriptome sequencing was performed by Dr Lira Mamanova (Wellcome Trust, Sanger Centre, Cambridge, UK) using a published methodology [18] on samples that were enriched for mRNA. RNA sequences from the global analysis were processed in-house using Galaxy [123]. Adapter sequences were removed and reads trimmed for quality before being mapped to the genome using Bowtie 1.0 [124] with custom parameter: -l 28 for read1, -l 20 for read2, and -y -a --best --strata. Pairs of datasets were compared using M-A (ratio-intensity) scatterplots, where M is \log_2 (reads plus/minus TAP treatment), and A is $(\log_2 \text{ plus} + \log_2 \text{ minus})/2$. For each value of A, we calculated the average (μ) and standard deviation (σ) of M in a moving window of 5,000 pairs that were sorted in ascending order of A. Upper and lower envelopes were defined by the equation: $\mu \pm X\sigma$, and positions outside the envelope recorded, as described previously [22,23]. The value of X was user defined (see text below for details). Microarray data was collected by Roche NimbleGen (Iceland) as a service using 4 × 72 k format arrays. RNA samples were analysed from duplicate cultures of *P. acnes* following subculture with and without potassium downshift.

5' RACE and RT-PCR analysis

The mapping of the 5' ends of specific transcripts was done using a 2nd generation 5'/3' RACE kit as described by the vendor (Roche Applied Science), except an aliquot of each RNA sample was treated with Terminator[™] 5'-Phosphate-Dependent Exonuclease (TEX), as described by the vendor (Epicentre[®] Biotechnologies), prior to starting the mapping protocol. The sequences of specific primers are indicated in the legends to relevant figures. Primers were designed with the assistance of Primer 3 software [125] and purchased from Eurogentec. For the Reverse-transcription polymerase chain reaction (RT-PCR), cDNA was synthesised using SuperScript[®] RT III (Invitrogen) with random hexamers (100 nM) and 200 ng of RNA template, the rest of the protocol was carried out as stated by manufacturer with no modifications. The PCR reaction was carried out using GoTaq[®] DNA polymerase (Promega) according to the vendor's

instruction using cDNA diluted with RNase-free water as the template.

Availability of supporting data

Both raw and processed RNA-seq data has been submitted to NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>). Accession: GSE46883 and GSE46810.

Additional files

Additional file 1: Schematic illustration of combined differential and global RNA-seq approach. RNA samples were enriched for mRNA by depleting 16S and 23S rRNA. To differentiate 5'-triphosphorylated ends (three yellow, filled circles) generated by transcription from 5'-monophosphorylated ends (single yellow, filled circles) produced by RNA processing or degradation, an aliquot was incubated with tobacco acid pyrophosphatase (TAP; branch on left), which leaves a monophosphate on 5' ends that were originally triphosphorylated. As a control, another aliquot was incubated under the same conditions, but without TAP (branch in centre). Both aliquots of this pair were then incubated separately with an adapter (red bar) that is only able to ligate to 5'-monophosphorylated ends. After terminating the 5'-end ligation reaction, the RNA was fragmented (not shown) to generate 3' ends close to the native 5' ends tagged with adaptor. The 3' ends of the resulting fragments were then tailed using poly(A) polymerase (broken black bar). Fragmentation following the attachment of the 5' adaptor allowed the efficient cloning of 5' ends associated with long as well as short transcripts. To identify transcription units onto which 5' ends could be mapped, a separate aliquot of the enriched mRNA was analysed using FRT-seq, an amplification-free form of strand-specific global RNA-seq (branch on right). To allow the mapping of all segments of transcripts, the RNA was fragmented and dephosphorylated prior to the ligation of adapter to 3' ends (solid black bar). The RNA was then 5' monophosphorylated to allow the ligation of a second adapter (red bar). The individual fragments were then reverse transcribed on the flow-cell without amplification, thus avoiding PCR biases and duplicates. The fragments were sequenced using an Illumina Genome Analyzer.

Additional file 2: Batch culture of *P. acnes* in Holland Synthetic Medium. (A), the growth profiles were constructed from the results of replicate cultures ($n = 13$). The black and red plots correspond to growth following subculture without and with a potassium downshift, respectively. The error bars indicate the standard deviation of the OD_{600} readings. Times given are from the point of subculture. RNA was isolated from cultures 1 h after subculturing. There was no distinct exponential phase following the potassium downshift: cells grew without a discernible lag phase, but their growth rate appeared to decrease steadily with time. This may reflect the utilisation of phosphate reserves that were accumulated during growth on TYG. The doubling-time and specific growth rate during exponential growth in the absence of potassium downshift were 6.2 h and 0.111 h^{-1} , respectively. (B), RT-PCR analysis of RNA isolated with and without potassium downshift: cDNA was synthesised from equal amount of *P. acnes* RNA, and used as template for PCR amplification of segments of the target genes PPA0010 (*gyrA*) and PPA0116 (*kdpB*) and analysed by electrophoresis using a 2.0% [w/v] agarose gel. Lane M shows a 100-bp DNA ladder (Fermentas), lanes 1 and 2 show PCR product using genomic DNA and no template, respectively. Lanes 3 and 4 show the PCR products of using cDNA synthesised from RNA isolated from cells subcultured without and with downshift, respectively. Lanes 5 and 6 are as lanes 3 and 4, but for a biological replicate. The sequences of the PCR primers were PPA0010F, 5'-CCCGTACTGGTCAGCGTTTA; PPA0010R, 5'-GCCGTCGCTGTACAGGTT; PPA0116F, 5'-CGGCAAGCAACTACTCATCA; and PPA0116R, 5'-TAAAGATGATCGCCGAGAGC. The *gyrA* transcript served as an internal control. Following potassium downshift, the *kdpB* gene is clearly induced.

Additional file 3: Transcriptional start sites identified for *P. acnes*. All of the positions listed in this table were either enriched (EN) following treatment with TAP in all 4 of 4 experiments (see Figure 1) or associated

with an obvious leading edge of transcription (LE) as judged by manual inspection of the global RNA-seq data or both. Nucleotide positions within 8 nt of each other were classified as belonging to the same TSS. The p-values are the probability that the number of reads corresponding to a 5' end increase following treatment with TAP, i.e. are associated with a TSS, according to Rank product analysis [25,26].

Additional file 4: *P. acnes* homologues of major factors involved in RNA processing and degradation. Further details of factors involved in eubacterial RNA degradation and processing can be found in several recent reviews [9,10,16,104].

Additional file 5: Genes with altered expression as a consequence of potassium downshift. The microarray data obtained for each of the two duplicate cultures was analysed using M-A (ratio-intensity) scatterplots (data not shown). The vast majority of the points in each comparison (with vs. without downshift) were contained within boundaries described by the equation $\mu \pm 3\sigma$, where μ and σ are the average and standard deviation, respectively, of M in a moving window of 5,000 data-points sorted in ascending order of A [22,23]. The genes listed in this table were outside the boundaries in both of the two comparisons. The microarray data was also analysed using an online version of Rank Product algorithm [26], which detects differentially regulated genes in replicated microarray experiments. Overall the analysis of the M-A scatterplots appears to have been more sensitive. In two cases, it identified all of the genes in a cluster with related function, while Rank Product did not (see PPA1287-90 and PPA1758-60). It should be noted that genes linked to iron homeostasis were detected, e.g. genes encoding the ferrous iron transport proteins A and B (PPA1676 and PPA1677, respectively), and the production of a peptide-based iron chelators (PPA1287-1291). The potassium used to culture *P. acnes* was contaminated with trace amounts of iron. Thus, removing the potassium also removed a source of iron. P-values are for change in expression. †Transcription starts within PPA0114.

Additional file 6: Processing of *P. acnes* tRNA. n.d. = not detected (scale of 0–1000 dRNA-seq reads).

Additional file 7: Sequence alignment of *P. acnes* promoters associated with the translational machinery. (A) shows ungapped sequences (+5 to –60) aligned to the '–10 box' (consensus sequence of T.A.n.n.T), which was identified using MEME [69] and an initial search window of –1 to –15. (B) as (A), except gaps have been introduced 5 nt upstream of the –10 boxes to maximise alignment to the '–35 box' (consensus sequence of G/A.n.T/G.T/G.n.G). Highlighting indicates nucleotides that match the consensus sequences (Figure 7).

Additional file 8: List of annotated and possible sRNAs in *P. acnes*.

Additional file 9: List of possible leaderless mRNAs.

Additional file 10: List of genes requiring reannotation.

Additional file 11: Degradation of 5'-triphosphorylated RNA by TEX. Total *P. acnes* RNA was isolated as described in *Methods*. A 5'-triphosphorylated form of *E. coli* *cspA* mRNA was synthesised by *in vitro* transcription using T7 RNA polymerase (Invitrogen) using conditions stated by the manufacturer [121]. 0.5 µg of *cspA* was added to 1.0 µg of total RNA and treated with TEX using Reaction Buffer B and 1 U of enzyme for 1 h as specified by the vendor (Epicentre® Biotechnologies). The reaction products were purified by phenol: chloroform extraction and analysed by gel electrophoresis (1.2% [w/v] agarose). Lane 1 shows the control sample before treatment. Lanes 2 and 3 shows the products following incubation without (reaction buffer only) and with TEX, respectively.

Competing interests

The authors declare that they have no competing interests of a financial or non-financial nature.

Authors' contributions

KJM, Y-Fl and DRA designed the transcriptome study. KJM and Y-Fl performed the bulk of the RNA-seq analysis, DRA provided valuable input. SG undertook the 5' RACE analysis. LM performed the global RNA-sequencing. KJM and Y-Fl wrote the paper. All authors read and approved the final manuscript.

Acknowledgments

This work was supported in part by a grant from the Leeds Foundation for Dermatological Research to R.A. Bojar. Y-Fl gratefully acknowledges a White Rose PhD Studentship awarded to K. Stephenson as part of the MICROSYST network. Y-Fl and SG thank K. Stephenson, K.T. Holland and R.A. Bojar for guidance during their PhD studies. The global RNA-sequencing was done at The Wellcome Trust Sanger Institute. The McDowall laboratory is supported by funding from the BBSRC.

Author details

¹Astbury Centre for Structural Molecular Biology, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK. ²The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

Received: 29 May 2013 Accepted: 11 September 2013

Published: 14 September 2013

References

1. Kurokawa I, Danby FW, Ju Q, Wang XL, Xiang LF, Xia LQ, Chen WC, Nagy I, Picardo M, Suh DH, et al: **New developments in our understanding of acne pathogenesis and treatment.** *Exp Dermatol* 2009, **18**(10):821–832.
2. Perry A, Lambert P: **Propionibacterium acnes: infection beyond the skin.** *Exp Rev Anti-infe* 2011, **9**(12):1149–1156.
3. Brook I: **Meningitis and shunt infection caused by anaerobic bacteria in children.** *Pediatr Neurol* 2002, **26**(2):99–105.
4. Clayton JJ, Baig W, Reynolds GW, Sandoe JA: **Endocarditis caused by Propionibacterium species: a report of three cases and a review of clinical features and diagnostic difficulties.** *J Med Microbiol* 2006, **55**(Pt 8):981–987.
5. Albert HB, Sorensen JS, Christensen BS, Manniche C: **Antibiotic treatment in patients with chronic low back pain and vertebral bone edema (Modic type 1 changes): a double-blind randomized clinical controlled trial of efficacy.** *Eur Spine J* 2013, **22**(4):697–707.
6. Cogen AL, Nizet V, Gallo RL: **Skin microbiota: a source of disease or defence?** *Br J Dermatol* 2008, **158**(3):442–455.
7. Bojar RA, Holland KT: **Review: the human cutaneous microflora and factors controlling colonisation.** *World J Microb Biot* 2002, **18**(9):889–903.
8. Bruggemann H, Henne A, Hoster F, Liesegang H, Wiezer A, Strittmatter A, Hujer S, Durre P, Gottschalk G: **The complete genome sequence of Propionibacterium acnes, a commensal of human skin.** *Science* 2004, **305**(5684):671–673.
9. Carpousis AJ, Luisi BF, McDowall KJ: **Endonucleolytic Initiation of mRNA Decay in Escherichia coli.** In *Molecular Biology of RNA Processing and Decay in Prokaryotes*. Edited by Condon C. London, UK: Academic Press; 2009:91–135.
10. Deutscher MP: **Maturation and degradation of ribosomal RNA in bacteria.** In *Molecular Biology of RNA Processing and Decay in Prokaryotes*. Edited by Condon C. London, UK: Academic press; 2009:369–391.
11. Hartmann RK, Gossringer M, Spath B, Fischer S, Marchfelder A: **The Making of tRNAs and More - RNase P and tRNase Z.** In *Molecular Biology of RNA Processing and Decay in Prokaryotes*. Edited by Condon C. London, UK: Academic Press; 2009:319–368.
12. Storz G, Vogel J, Wassarman KM: **Regulation by small RNAs in bacteria: expanding frontiers.** *Mol Cell* 2011, **43**(6):880–891.
13. Brzuszkiewicz E, Weiner J, Wollherr A, Thurner A, Hupeden J, Lomholt HB, Kilian M, Gottschalk G, Daniel R, Mollenkopf HJ, et al: **Comparative genomics and transcriptomics of Propionibacterium acnes.** *Plos One* 2011, **6**(6):e21581.
14. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermuller J, Reinhardt R, et al: **The primary transcriptome of the major human pathogen Helicobacter pylori.** *Nature* 2010, **464**(7286):250–255.
15. Breter HJ, Rhoads RE: **Analysis of NaIO₄-oxidized/NaBH₄-reduced mRNA cap analogs by high-performance liquid anion-exchange chromatography and tobacco acid pyrophosphatase (EC 3.6.1.9).** *H-S Z Physiol Chem* 1979, **360**(3):240–240.
16. Bechhofer DH: **Messenger RNA Decay and Maturation in Bacillus subtilis.** In *Molecular Biology of RNA Processing and Decay in Prokaryotes*. Edited by Condon C. London, UK: Academic Press; 2009:231–273.
17. Belasco JG: **All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay.** *Nat Rev Mol Cell Biol* 2010, **11**(7):467–478.

18. Mamanova L, Andrews RM, James KD, Sheridan EM, Ellis PD, Langford CF, Ost TWB, Collins JE, Turner DJ: **FRT-seq: amplification-free, strand-specific transcriptome sequencing.** *Nat Methods* 2010, **7**(2):130–U163.
19. Berenson CS, Burch GE: **A study of the Na, K, Cl content of thermal sweat of man collected from small isolated areas of the skin.** *J Lab Clin Med* 1953, **42**:58–77.
20. Borelli F: **Research on quantitative variations of potassium of the blood and skin in normal and pathological conditions.** *Derm Sifilograf* 1932, **7**:353.
21. Ballal A, Basu B, Apte SK: **The Kdp-ATPase system and its regulation.** *J Biosci* 2007, **32**(3):559–568.
22. Hovatta I, Kimppa K, Lehmuusola A, Pasanen T, Saarela J, Saarikko I, Saharinen J, Tiikkainen P, Toivanen T, Tolvanen M, et al: **DNA Microarray Data Analysis.** 2nd edition. Helsinki: CSC - Scientific Computing Ltd; 2005.
23. Marincs F, Manfield IW, Stead JA, McDowall KJ, Stockley PG: **Transcript analysis reveals an extended regulon and the importance of protein-protein co-operativity for the *Escherichia coli* methionine repressor.** *Biochem J* 2006, **396**(2):227–234.
24. Salgado H, Gama-Castro S, Peralta-Gil M, Diaz-Peredo E, Sanchez-Solano F, Santos-Zavaleta A, Martinez-Flores I, Jimenez-Jacinto V, Bonavides-Martinez C, Segura-Salazar J, et al: **RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions.** *Nucleic Acids Res* 2006, **34**:D394–D397.
25. Breitling R, Armengaud P, Amtmann A, Herzyk P: **Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments.** *FEBS Lett* 2004, **573**(1–3):83–92.
26. Laing E, Smith CP: **RankProdIT: A web-interactive Rank Products analysis tool.** *BMC Res Notes* 2010, **3**:221.
27. Chan PP, Holmes AD, Smith AM, Tran D, Lowe TM: **The UCSC Archaeal Genome Browser: 2012 update.** *Nucleic Acids Res* 2012, **40**:D646–652.
28. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW: **GenBank.** *Nucleic Acids Res* 2013, **41**:D36–42.
29. Schneider KL, Pollard KS, Baertsch R, Pohl A, Lowe TM: **The UCSC Archaeal Genome Browser.** *Nucleic Acids Res* 2005, **34**:D407–410.
30. Jacquier A: **The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs.** *Nat Rev Genet* 2009, **10**(12):833–844.
31. Marguerat S, Bahler J: **RNA-seq: from technology to biology.** *Cell Mol Life Sci* 2010, **67**(4):569–579.
32. Albrecht M, Sharma CM, Reinhardt R, Vogel J, Rudel T: **Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome.** *Nucleic Acids Res* 2010, **38**(3):868–877.
33. Beaume M, Hernandez D, Farinelli L, Deluen C, Linder P, Gaspin C, Romby P, Schrenzel J, Francois P: **Cartography of methicillin-resistant *S. aureus* transcripts: detection, orientation and temporal expression during growth phase and stress conditions.** *Plos One* 2010, **5**(5):e10725.
34. Cho BK, Zengler K, Qiu Y, Park YS, Knight EM, Barrett CL, Gao Y, Palsson BO: **The transcription unit architecture of the *Escherichia coli* genome.** *Nat Biotechnol* 2009, **27**(11):1043–1049.
35. Dornenburg JE, DeVita AM, Palumbo MJ, Wade JT: **Widespread antisense transcription in *Escherichia coli*.** *Mbio* 2010, **1**(1):e00024–10.
36. Filiatrault MJ, Stodghill PV, Bronstein PA, Moll S, Lindeberg M, Grills G, Schweitzer P, Wang W, Schroth GP, Luo SJ, et al: **Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity.** *J Bacteriol* 2010, **192**(9):2359–2372.
37. Georg J, Voss B, Scholz I, Mitschke J, Wilde A, Hess WR: **Evidence for a major role of antisense RNAs in cyanobacterial gene regulation.** *Mol Syst Biol* 2009, **5**. doi:10.1038/msb.2009.63.
38. Guell M, van Noort V, Yus E, Chen WH, Leigh-Bell J, Michalodimitrakis K, Yamada T, Arumugam M, Doerks T, Kuhner S, et al: **Transcriptome complexity in a genome-reduced bacterium.** *Science* 2009, **326**(5957):1268–1271.
39. Jager D, Sharma CM, Thomsen J, Ehlers C, Vogel J, Schmitz RA: **Deep sequencing analysis of the *Methanosarcina mazei* Go1 transcriptome in response to nitrogen availability.** *Proc Natl Acad Sci USA* 2009, **106**(51):21878–21882.
40. Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penades JR, Valle J, et al: **Genome-wide antisense transcription drives mRNA processing in bacteria.** *Proc Natl Acad Sci USA* 2011, **108**(50):20172–20177.
41. Martin J, Zhu WH, Passalacqua KD, Bergman N, Borodovsky M: ***Bacillus anthracis* genome organization in light of whole transcriptome sequencing.** *Bmc Bioinformatics* 2010, **11**. doi:10.1186/1471-2105-11-S3-S10.
42. Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, Jimenez-Jacinto V, Salgado H, Juarez K, Contreras-Moreira B, et al: **Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*.** *Plos One* 2009, **4**(10):e7526.
43. Mitschke J, Georg J, Scholz I, Sharma CM, Dienst D, Bantscheff J, Voss B, Steglich C, Wilde A, Vogel J, et al: **An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp PCC6803.** *Proc Natl Acad Sci USA* 2011, **108**(5):2124–2129.
44. Rasmussen S, Nielsen HB, Jarmer H: **The transcriptionally active regions in the genome of *Bacillus subtilis*.** *Mol Microbiol* 2009, **73**(6):1043–1057.
45. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, Loh E, Gripenland J, Tiensuu T, Vaitkevicius K, et al: **The *Listeria* transcriptional landscape from saprophytism to virulence.** *Nature* 2009, **459**(7249):950–956.
46. Wurtzel O, Sapra R, Chen F, Zhu YW, Simmons BA, Sorek R: **A single-base resolution map of an archaeal transcriptome.** *Genome Res* 2010, **20**(1):133–141.
47. Ghora BK, Apirion D: **Structural analysis and *in vitro* processing to ρ 5 rRNA of a 9S RNA molecule isolated from an *rne* mutant of *E. coli*.** *Cell* 1978, **15**(3):1055–1066.
48. Shahbabian K, Jamali A, Zig L, Putzer H: **RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*.** *EMBO J* 2009, **28**(22):3523–3533.
49. Robertson HD, Webster RE, Zinder ND: **Purification and properties of ribonuclease III from *Escherichia coli*.** *J Biol Chem* 1968, **243**(1):82–91.
50. Even S, Pellegrini O, Zig L, Labas V, Vinh J, Brechemmier-Baey D, Putzer H: **Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E. coli* RNase E.** *Nucleic Acids Res* 2005, **33**(7):2141–2152.
51. Mathy N, Benard L, Pellegrini O, Daou R, Wen TY, Condon C: **5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA.** *Cell* 2007, **129**(4):681–692.
52. Lin Y: **Genome-wide analysis of *Propionibacterium acnes* gene regulation.** Leeds, UK: University of Leeds; 2013.
53. Dittmar KA, Mobley EM, Radek AJ, Pan T: **Exploring the regulation of tRNA distribution on the genomic scale.** *J Mol Biol* 2004, **337**(1):31–47.
54. Betat H, Rammelt C, Morl M: **tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization.** *Cell Mol Life Sci* 2010, **67**(9):1447–1463.
55. McDowall KJ, Lin-Chao S, Cohen SN: **A + U content rather than a particular nucleotide order determines the specificity of RNase E cleavage.** *J Biol Chem* 1994, **269**(14):10790–10796.
56. de Boer HA, Gilbert SF, Nomura M: **DNA sequences of promoter regions for rRNA operons *rne* and *rna* in *Escherichia coli*.** *Cell* 1979, **17**(1):201–209.
57. Gilbert SF, de Boer HA, Nomura M: **Identification of initiation sites for the *in vitro* transcription of rRNA operons *rne* and *rna* in *Escherichia coli*.** *Cell* 1979, **17**(1):211–224.
58. Young RA, Steitz JA: **Tandem promoters direct *Escherichia coli* rRNA synthesis.** *Cell* 1979, **17**(1):225–234.
59. Stewart GC, Bott KF: **DNA sequence of the tandem rRNA promoter for *B. subtilis* operon *rrnB*.** *Nucleic Acids Res* 1983, **11**(18):6289–6300.
60. Nicholson AW: **The ribonuclease III superfamily: forms and functions in RNA maturation, decay, and gene silencing.** In *RNAi: A Guide to Gene Silencing*. Edited by Hannon GJ. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2003.
61. Jacob AJ, Kohrer C, Davies BW, Rajbhandary UL, Walker GC: **Conserved Bacterial RNase YbeY Plays Key Roles in 70S Ribosome Quality Control and 16S rRNA Maturation.** *Mol Cell* 2013, **49**(3):427–438.
62. Gausing K: **Regulation of ribosome production in *Escherichia coli*: synthesis and stability of ribosomal RNA and of ribosomal protein messenger RNA at different growth rates.** *J Mol Biol* 1977, **115**(3):335–354.
63. Norris TE, Koch AL: **Effect of growth rate on the relative rates of synthesis of messenger, ribosomal and transfer RNA in *Escherichia coli*.** *J Mol Biol* 1972, **64**(3):633–649.
64. Zuker M: **Mfold web server for nucleic acid folding and hybridization prediction.** *Nucleic Acids Res* 2003, **31**(13):3406–3415.
65. Jarrige AC, Mathy N, Portier C: **PNPase autocontrols its expression by degrading a double-stranded structure in the *pnp* mRNA leader.** *EMBO J* 2001, **20**(23):6845–6855.
66. Robert-Le meur M, Portier C: ***Escherichia coli* polynucleotide phosphorylase expression is autoregulated through an RNase III-dependent mechanism.** *EMBO J* 1992, **11**(7):2633–2641.

67. Robert-Le meur M, Portier C: Polynucleotide phosphorylase of *Escherichia coli* induces the degradation of its RNase III-processed messenger by preventing its translation. *Nucleic Acids Res* 1994, **22**(3):397–403.
68. Gatewood ML, Bralley P, Jones GH: RNase III-dependent expression of the *rpsO-pnp* operon of *Streptomyces coelicolor*. *J Bacteriol* 2011, **193**(17):4371–4379.
69. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 2009, **37**:W202–208.
70. Crooks GE, Hon G, Chandonia JM, Brenner SE: WebLogo: A sequence logo generator. *Genome Res* 2004, **14**(6):1188–1190.
71. Harley CB, Reynolds RP: analysis of *Escherichia coli* promoter sequences. *Nucleic Acids Res* 1987, **15**(5):2343–2361.
72. Lissner S, Margalit H: Compilation of *Escherichia coli* mRNA promoter sequences. *Nucleic Acids Res* 1993, **21**(7):1507–1516.
73. Mulligan ME, Hawley DK, Entriken R, McClure WR: *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. *Nucleic Acids Res* 1984, **12**(1 Pt 2):789–800.
74. Borovok I, Gorovitz B, Yanku M, Schreiber R, Gust B, Chater K, Aharonowitz Y, Cohen G: Alternative oxygen-dependent and oxygen-independent ribonucleotide reductases in *Streptomyces*: cross-regulation and physiological role in response to oxygen limitation. *Mol Microbiol* 2004, **54**(4):1022–1035.
75. Torrents E, Grinberg I, Gorovitz-Harris B, Lundstrom H, Borovok I, Aharonowitz Y, Sjoberg BM, Cohen G: NrdR controls differential expression of the *Escherichia coli* ribonucleotide reductase genes. *J Bacteriol* 2007, **189**(14):5012–5021.
76. Herrick J, Sclavi B: Ribonucleotide reductase and the regulation of DNA replication: an old story and an ancient heritage. *Mol Microbiol* 2007, **63**(1):22–34.
77. Hiard S, Maree R, Colson S, Hoskisson PA, Titgemeyer F, van Wezel GP, Joris B, Wehenkel L, Rigali S: PREDetector: a new tool to identify regulatory elements in bacterial genomes. *Biochem Biophys Res Commun* 2007, **357**(4):861–864.
78. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A: Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res* 2005, **33**:D121–D124.
79. Barrick JE, Breaker RR: The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* 2007, **8**(11):R239.
80. Franklund CV, Kadner RJ: Multiple transcribed elements control expression of the *Escherichia coli* *btuB* gene. *J Bacteriol* 1997, **179**(12):4039–4042.
81. Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR: Genetic control by a metabolite binding mRNA. *Chem Biol* 2002, **9**(9):1043–1049.
82. Vitreschak AG, Rodionov DA, Mironov AA, Gelfand MS: Regulation of the vitamin B₁₂ metabolism and transport in bacteria by a conserved RNA structural element. *RNA* 2003, **9**(9):1084–1097.
83. Nordlund N, Reichard P: Ribonucleotide reductases. *Annu Rev Biochem* 2006, **75**:681–706.
84. Guan S: A novel two-component signal transduction system in *Propionibacterium acnes* and its association with a putative extracellular signalling peptide. Leeds: University of Leeds; 2011.
85. Breaker RR: Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* 2012, **4**(2):a003566.
86. Henkin TM, Yanofsky C: Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays* 2002, **24**(8):700–707.
87. Higgins CF, McLaren RS, Newbury SF: Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion? A review. *Gene* 1988, **72**(1–2):3–14.
88. Bandyra KJ, Said N, Pfeiffer V, Gorna MW, Vogel J, Luisi BF: The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Mol Cell* 2012, **47**(6):943–953.
89. Pruitt KD, Tatusova T, Maglott DR: NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 2007, **35**:D61–D65.
90. Shine J, Dalgarno L: 3'-terminal sequence of *Escherichia coli* 16S rRNA: possible role in initiation and termination of protein synthesis. *P Aust Biochem Soc* 1974, **7**:72–72.
91. Shine J, Dalgarno L: Terminal sequence analysis of bacterial rRNA: correlation between 3'-terminal polypyrimidine sequence of 16S RNA and translational specificity of ribosome. *Eur J Biochem* 1975, **57**(1):221–230.
92. Suzek BE, Ermolaeva MD, Schreiber M, Salzberg SL: A probabilistic, method for identifying start codons in bacterial genomes. *Bioinformatics* 2001, **17**(12):1123–1130.
93. Janssen GR: Eubacterial, archaeobacterial, and eukaryotic genes that encode leaderless mRNA. In *Industrial Microorganisms: Basic and Applied Molecular Genetics*. Edited by Baltz RH, Hegeman G. Washington, D.C, United States: ASM Press; 1993:59–67.
94. Price MN, Huang KH, Alm EJ, Arkin AP: A novel method for accurate operon predictions in all sequenced prokaryotes. *Nucleic Acids Res* 2005, **33**(3):880–892.
95. Vockenhuber MP, Sharma CM, Statt MG, Schmidt D, Xu ZJ, Dietrich S, Liesegang H, Mathews DH, Sues B: Deep sequencing-based identification of small non-coding RNAs in *Streptomyces coelicolor*. *RNA Biol* 2011, **8**(3):468–477.
96. Vesper O, Amitai S, Belitsky M, Byrgazov K, Kaberdina AC, Engelberg-Kulka H, Moll I: Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell* 2011, **147**(1):147–157.
97. Chuang LY, Chang HW, Tsai JH, Yang CH: Features for computational operon prediction in prokaryotes. *Brief Funct Genomics* 2012, **11**(4):291–299.
98. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B Jr, Assad-Garcia N, Glass JI, Covert MW: A whole-cell computational model predicts phenotype from genotype. *Cell* 2012, **150**(2):389–401.
99. McDowell DG, Burns NA, Parkes HC: Localised sequence regions possessing high melting temperatures prevent the amplification of a DNA mimic in competitive PCR. *Nucleic Acids Res* 1998, **26**(14):3340–3347.
100. Liu JM, Livny J, Lawrence MS, Kimball MD, Waldor MK, Camilli A: Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. *Nucleic Acids Res* 2009, **37**(6):e46.
101. Graur D, Zheng Y, Price N, Azevedo RB, Zufall RA, Elhaik E: On the immortality of television sets: “function” in the human genome according to the evolution-free gospel of ENCODE. *Genome Biol Evol* 2013, **5**(3):578–590.
102. Chintakayala K, Singh SS, Rossiter AE, Shahapure R, Dame RT, Grainger DC: *E. coli* Fis protein insulates the *cbpA* gene from uncontrolled transcription. *Plos Genetics* 2013, **9**(1):e1003152.
103. Dreyfus M: Killer and protective ribosomes. In *Molecular Biology of Rna Processing and Decay in Prokaryotes*, Volume 85. San Diego: Elsevier Academic Press Inc; 2009:423–466.
104. Andrade JM, Pobre V, Silva U, Domingues S, Arraiano CM: The role of 3'-5' exoribonucleases in RNA degradation. In *Molecular Biology of RNA Processing and Decay in Prokaryotes*. Edited by Condon C. London, UK: Academic Press; 2009:187–229.
105. Jung K, Fried L, Behr S, Heermann R: Histidine kinases and response regulators in networks. *Curr Opin Microbiol* 2012, **15**(2):118–124.
106. Moll I, Grill S, Gualerzi CO, Blasi U: Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Mol Microbiol* 2002, **43**(1):239–246.
107. Malys N, McCarthy JEG: Translation initiation: variations in the mechanism can be anticipated. *Cell Mol Life Sci* 2011, **68**(6):991–1003.
108. Walz A, Pirrotta V, Ineichen K: Lambda repressor regulates switch between P_R and P_{RM} promoters. *Nature* 1976, **262**(5570):665–669.
109. Baumeister R, Flache P, Meleforts O, Vongabain A, Hillen W: Lack of a 5' noncoding region in Tn1721-encoded *tetR* mRNA is associated with a low efficiency of translation and a short half-life in *Escherichia coli*. *Nucleic Acids Res* 1991, **19**(17):4595–4600.
110. Van Etten WJ, Janssen GR: An AUG initiation codon, not codon-anticodon complementarity, is required for the translation of unleadered mRNA in *Escherichia coli*. *Mol Microbiol* 1998, **27**(5):987–1001.
111. O'Donnell SA, Janssen GR: Leaderless mRNAs bind 70S ribosomes more strongly than 30S ribosomal subunits in *Escherichia coli*. *J Bacteriol* 2002, **184**(23):6730–6733.
112. Brock JE, Pourshahian S, Giliberti J, Limbach PA, Janssen GR: Ribosomes bind leaderless mRNA in *Escherichia coli* through recognition of their 5' terminal AUG. *RNA* 2008, **14**(10):2159–2169.
113. Vioque A, Arnez J, Altman S: Protein-RNA interactions in the RNase P holoenzyme from *Escherichia coli*. *J Mol Biol* 1988, **202**(4):835–848.
114. Chauhan AK, Apirion D: The gene for a small stable RNA (10Sa RNA) of *Escherichia coli*. *Mol Microbiol* 1989, **3**(11):1481–1485.
115. Lee SY, Bailey SC, Apirion D: Small stable RNAs from *Escherichia coli*: evidence for existence of new molecules and for a new ribonucleoprotein particle containing 6S RNA. *J Bacteriol* 1978, **133**(2):1015–1023.

116. Glynn B, Lacey K, Palta P, Kaplinski L, Remm M, Barry T, Smith T, Maher M: **Demonstration of the application of the tmRNA transcript of the bacterial *ssrA* gene as a molecular diagnostic target using a combination of NASBA and BiaCore technologies.** *Int J Antimicrob Ag* 2007, **29**:S392–S392.
117. Bremer H, Dennis PP: **Modulation of chemical composition and other parameters of the cell by growth rate.** In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. 2nd edition. Edited by Neidhardt FC. Washington, DC: ASM Press; 1996:1553–1569.
118. Holland KT, Greenman J, Cunliffe WJ: **Growth of cutaneous *Propionibacteria* on synthetic medium - growth yields and exoenzyme production.** *J Appl Bacteriol* 1979, **47**(3):383–394.
119. Hirsch A, Grinstead E: **Methods for the growth and enumeration of anaerobic spore-formers from cheese, with observations on the effect of nisin.** *J Dairy Res* 1954, **21**(1):101–110.
120. Kim J, Naylor HB: **Spore production by *Bacillus stearothermophilus*.** *Appl Microbiol* 1966, **14**(4):690–691.
121. Kime L, Jourdan SS, McDowall KJ: **Identifying and characterizing substrates of the RNase E/G family of enzyme.** In *RNA Turnover in Bacteria, Archaea and Organelles*. Edited by Maquat LE, Arraiano CM. San Diego, California, USA: Academic Press; 2008:215–241.
122. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA: *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation; 2000.
123. Goecks J, Nekrutenko A, Taylor J, Team G: **Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences.** *Genome Biol* 2010, **11**(8):R86.
124. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**(3):R25.
125. Rozen S, Skaletsky HJ: **Primer3 on the WWW for general users and for biologist programmers.** In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Edited by Krawetz S, Misener S. Totowa, NJ: Humana Press; 2000:365–386.

doi:10.1186/1471-2164-14-620

Cite this article as: Lin et al.: A combination of improved differential and global RNA-seq reveals pervasive transcription initiation and events in all stages of the life-cycle of functional RNAs in *Propionibacterium acnes*, a major contributor to wide-spread human disease. *BMC Genomics* 2013 **14**:620.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

