Sense and nonsense from whole genome microarray data in the analysis of microbial physiology

Tim W. Overton, Rebekah Whitehead, Ying Li, Lesley Griffiths, Jeff Cole

School of Biosciences, University of Birmingham, Birmingham

Sense and nonsense from whole genome microarray data in the analysis of microbial physiology

Summary

The arrival of the post-genomic era has allowed the regulation of every gene or protein of an organism to be studied at once using microarrays for transcriptomic studies, proteomics to analyse gene products, and metabolomics to study the complete complement of products and intermediary metabolites produced by a single person or a single organism. Too often the results of such enterprises are disappointing either because many of the products cannot be identified, or because they are products of genes of unknown function. Success is far more likely to be achieved if the organism to be exploited is thoroughly understood at the levels of genome organisation, regulation, physiology and biochemistry.

Typical questions asked in biotechnology and the biopharmaceutical industries include what genes are expressed – or not expressed – when recombinant protein production is induced: can they be manipulated to provide a more productive host; and how do successful pathogens survive in the human body when exposed to oxygen starvation or chemical attack by host defence mechanisms. Examples are given of how whole genome microarray data can reveal mechanisms used by bacteria to survive when they are starved of oxygen; what genes are turned on in response to host defence mechanisms such as nitrosative attack; and how pathogens repair damage inflicted by the host defence mechanisms. Striking similarities and fascinating differences are revealed between two major groups of pathogenic bacteria: enteric bacteria that are able to adapt to life both inside and outside an animal host, and the obligate human pathogen, *Neisseria gonorrhoeae*. Proteins are identified that provide possible targets for biopharmaceutical intervention, and hence illustrate the potential value of whole-genome transcriptomic approaches to biotechnology.

Adres do korespondencji

Jeff Cole,
School of Biosciences,
University of Birmingham,
Birmingham B15 2TT, UK;
e-mail:
j.a.cole@bham.ac.uk

biotechnologia
1 (80) 15–30 2008

Key words:
microarrays, pathogenicity determinants, reactive nitrogen species, gene regulation.

1. Introduction: from targeted gene analysis to the era of systems biology

Successful biotechnology can deliver massive benefits to society, the environment, industry and the national economy – irrespective whether the ultimate goal is to cure a disease, prevent environmental pollution, provide preventative medicine at the personal level, or simply to make money by selling a product. Success is far more likely to be achieved if the organism to be exploited is thoroughly understood at the levels of genome organisation, regulation, physiology and the biochemistry.

Throughout the 20th century, massive progress was made in understanding plants, animal cells and micro-organisms by studying one gene or one protein at a time, but by the year 2000 genome-wide sequencing was well established, so the ball-game completely changed as people tried to study the regulation of every gene or protein of an organism at once using microarrays for transcriptomic studies, proteomics to analyse gene products, and metabolomics to study the complete complement of products and intermediary metabolites produced by a single person or a single organism. A significant disappointment has been the frequency with which the only results of microarray or proteomic studies have been long lists of genes or proteins that respond to the insult of the experiment. Too often the sole conclusion from such studies is that hundreds of products increase or decrease in response to drug treatment, heat or cold stress, carbon starvation, oxidative stress, diseases of genetic or microbial origin, or even during recombinant protein production using a microbial host. In many cases the results of this type of study have three things in common.

(i). Many of the products either cannot be identified, or are products of genes of unknown function.

(ii). The results reported are statistically mathematically highly significant.

(iii). The lists of genes or proteins responding to the insult are unenlightening.

In an attempt to answer the inevitable question how these global approaches can be made more productive, typical reasons for failure will first be discussed before examples of how to solve the inherent problems will be presented. Finally, examples will be given of how microarray experiments with pathogenic bacteria can reveal new insights into their physiology that suggest novel targets for biopharmaceutical intervention.

2. How not to discover changes in gene expression due to recombinant protein production

Many biotechnology projects start with a requirement to produce recombinant proteins. Sometimes the protein itself is the end product: monoclonal antibodies; subunit vaccines; factor VIII; insulin are obvious examples. More often the protein is required for structural studies that will lead to drug development for the pharmaceutical industry. While it is dangerous to generalise, these projects crudely fall into two classes. Some companies require the largest possible quantities of product at the lowest possible price, so it does not matter if the protein is recovered from inclusion bodies. Others require much small quantities of perfectly folded protein suitable for crystallography and structure determination.

Typically the production host, for example a bacterium transformed with a recombinant plasmid, is grown to the required cell density and recombinant protein production is then induced, sometimes with a shift in temperature (1). There is a rapid burst of product formation that might continue for as little as 1 hour, or in a well designed process, it might continue for 24 hours or longer. Far too often, however, the burst of product formation soon stops, and the productive bacteria stop growing (2). So this leads to the question: what genes are expressed – or not expressed – when recombinant protein production is induced, and can they be manipulated to provide a more productive host? In attempts to answer this question, RNA is prepared from bacteria before and after protein production is induced; the differences between the two samples are analysed on a microarray.

The resulting list will include hundreds of genes that are turned on or turned off after the inducer is added, but will tell you exactly what you already know: because growth stops sooner or later after recombinant protein production is induced, genes that respond to growth rate also respond to recombinant protein production. This includes the general stress response, the stringent response controlled by ppGpp; genes required for protein, RNA and ribosome synthesis; and usually just a few genes of unknown function which it is tempting to waste the next year or two studying in the hope they might provide a magic bullet! In short the data from the microarray studies, and any systems biology based upon it, are largely unproductive. The reasons for this will be explored in the remainder of this review using examples from medical biotechnology based upon different types of pathogenic bacteria, which will illustrate how knowledge of the microbial physiology helps uncover targets for biotechnological intervention.

3. Identification of genes required for pathogenic bacteria to survive in the human body

The availability of whole genome sequences for more than 200 strains or species of pathogenic bacteria has provided major new challenges in healthcare biotechnology. Goals that are both altruistic and potentially profitable include how to develop new vaccines to prevent disease, or the identification of new protein targets for biopharmaceutical intervention. A rational approach to answer both of these questions is to answer the question: what genes or proteins are essential for pathogens to survive in the human body?

This question can rarely be answered from laboratory experiments with aerobic cultures of bacteria growing in shake flasks because apart from the upper respiratory tract, most sites of infection in the human body are anaerobic. To be a successful pathogen, bacteria must be able to survive oxygen starvation. But human neutrophiles and macrophages respond to a bacterial attack by what is known as the oxidative burst and the nitrosative burst: our human defence mechanisms try to kill the bacteria by chemical attack. So this leads to three key questions, all of which are possible to answer using genome-wide microarray experiments.

(i) First, how do bacteria survive when starved of oxygen?

(ii) Secondly, what genes are turned on in response to host defence mechanisms, for example, nitrosative attack?

(iii) How do pathogens repair damage inflicted by the host defence mechanisms?

Although many pathogenic bacteria can grow well even when they are starved of oxygen, this involves massive changes in their physiology and biochemistry regulated by multiple transcription factors. It is therefore perhaps not surprising that some of the early microarray experiments designed to answer the first of these questions yielded disappointing results.

In *E. coli* and related bacteria, two global control circuits regulate the switch from aerobic to anaerobic growth. During anaerobic growth, many of the enzymes required for aerobic growth are down regulated, and this is achieved by a two-component regulatory system, ArcB-ArcA (3,4). The Arc system is essentially an off switch that represses synthesis of enzymes required for aerobic metabolism. If alternative electron acceptors such as fumarate, nitrate, nitrite or DMSO are available, genes required for the synthesis of anaerobic respiratory chains are turned on. This requires transcription activation provided by the FNR protein (5,6). FNR stands for regulator of fumarate and nitrate reduction. However, the above statements are misleading because there are many complications. First, FNR can function as an off switch, for example, at the promoter of the *ndh* gene that encodes the energy-dissipating NADH dehydrogenase II (7), so FNR helps ArcA do part of its work. Secondly, ArcA can also function as an on switch, helping FNR to do its work (8,9). But there is an even worse complication: during anaerobic growth, the FNR protein activates transcription of the *arcA* gene, so every gene that is directly regulated by ArcA is also indirectly regulated by FNR (10).

In the early *E. coli* microarray studies, two groups each listed genes differentially expressed in an *E. coli fnr* mutant compared with its parent. The Gunsalus group detected over 1400 genes that responded to the mutation (11); the Blattner-Kiley study found over 900 genes that were differentially transcribed (12). Although one would expect most of these 900 to be included in the Gunsalus list, only 334 genes were the same in both studies, and out of these 334, 123 were stated to be regulated in completely opposite ways by FNR. This meant that although 2073 genes were identified in one or other study, the two studies agreed about only 211 of these genes, or 10% of the total.

One of the problems that had to be resolved was that many FNR-activated genes require specific growth conditions, for example the addition of nitrate or nitrite, to be induced. Secondly, in these early studies, glucose was used as carbon source for growth, despite the fact that it represses many anaerobically induced pathways. The reason for this was that an *fnr* mutant cannot grow on a non-repressing but non-fermentable carbon source such as glycerol using nitrate as the terminal electron acceptor for anaerobic growth. To resolve this problem, our subsequent study exploited two facts.

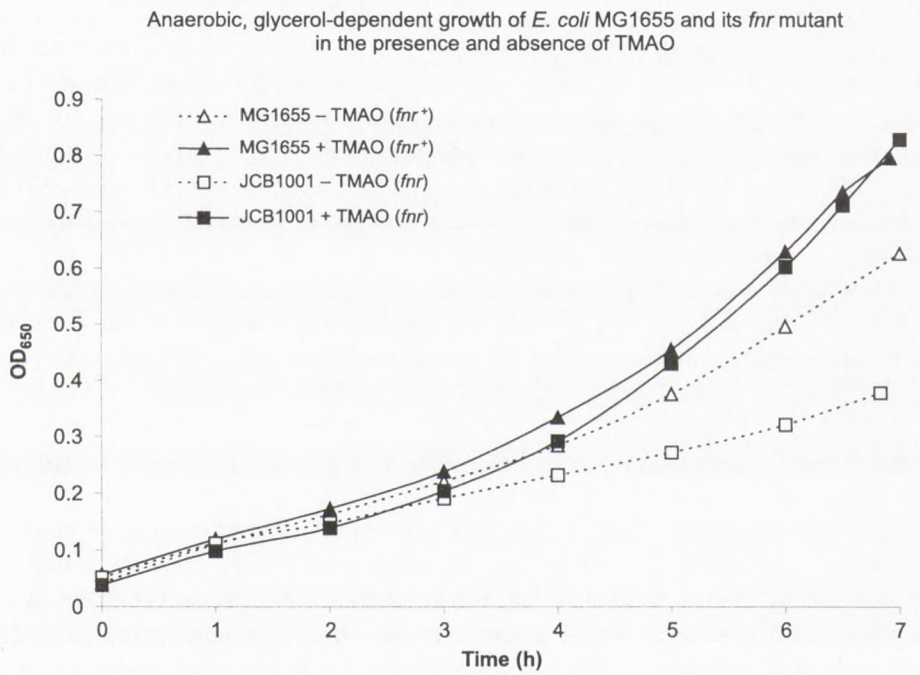

Fig. 1. Effect of the addition of trimethylamine-N-oxide (TMAO) on anaerobic growth of *Escherichia coli* K-12 strain MG1655 and an *fnr* mutant, strain JCB1001, in minimal medium supplemented with glycerol as the main carbon source and fumarate as the terminal electron acceptor. These data were originally published in reference (13).

(i) Trimethylamine-N-oxide (TMAO) is an excellent terminal electron acceptor for anaerobic growth of enteric bacteria, but TMAO reduction is completely independent of FNR.

(ii) Fumarate is also an excellent electron acceptor for anaerobic growth, but fumarate reductase synthesis is only partially dependent upon FNR.

The *fnr* mutant grew anaerobically at the same rate as the parent strain, *E. coli* MG1655, in a minimal medium slightly amended with 5% (v/v) Luria broth, glycerol, fumarate and TMAO, but grew poorly in the absence of TMAO (Fig. 1, from ref. 13). Thus the global effects of the FNR protein, nitrate and nitrite on anaerobic growth and gene regulation could be studied under conditions that avoided artefacts due to both glucose repression and differences in growth rates between strains.

A five-point strategy was then applied to identify operon previous not known to be activated or repressed by the FNR transcription factor. First, a literature search identified all operons that had previously been reported to be FNR-regulated, revealing independent biochemical or genetic evidence that 29 operons are activated and 14 operons are repressed by FNR. In some of these cases, however, apparent regulation was based upon promoter fusion data and might therefore be indirect effects. The microarray data confirmed 32 of these 43 assignments, but we failed to detect FNR-activation of 5 operons (*adhE, glpTQ, cydDC, hlyE* and *arcA*), or FNR repression of 6 operons (*hemA, narXL, tpx, yeiL, norVW* or *ubiCA*) (13). In every case there is an easy explanation for the 11 discrepancies. This provided confidence that the microarray data could be used to identify operons previously unknown to be FNR regulated. For each promoter that responded, directly or indirectly, to FNR, a bioinformatic search was used to exclude promoters that lacked a recognisable FNR-binding site (an 8 out of 10 match to the consensus sequence, TTGAT-N_4-ATCAA, where N is any nucleotide). Finally, discrepancies were resolved or new discoveries confirmed by direct experimentation.

Forty-four operons not previously known to be included in the FNR regulon were activated by FNR and a further 28 operons appeared to be repressed. The *E. coli* FNR regulon therefore includes at least 104, and possibly as many as 115, operons (13).

4. Dual two-component systems regulate the *E. coli* response to nitrate

The microarray data revealed that the FNR regulon is about twice as large as had been discovered over 30 years by a gene-by-gene approach, confirming that FNR regulates many aspects of nitrate and nitrite metabolism. Nitrate is sensed by two membrane-bound environmental sensor proteins, NarX and NarQ, respectively (14). Both are autokinases that trans-phosphorylate both of the cognate response regulator protein, NarL and NarP. Structural information about NarL is available (15), and the complex interactions between nitrate, nitrite, NarL, NarP and the DNA targets for phosphorylated NarL and NarP have been extensively documented in at least

three laboratories (16-20). As a simplification, NarQ detects and responds to micromolar concentrations of nitrate typical of those found in the gastro-intestinal tracts of warm-blooded animals, and nitrate is a far more effective ligand for activating its kinase activity than nitrite. In contrast, NarX responds to higher nitrate concentrations typical of those found in soil and waste water treatment plants – conditions rarely encountered in the human body. Again in contrast to NarQ, in the presence of nitrite, NarX promotes dephosphorylation of phosphorylated NarL, inactivating its regulatory functions (14).

Table 1

Examples of NarXL and nitrate regulated operons

Transcript in each subgroup[a]	Ratio[b]	
	wt + NO_3^- / wt – NO_3^-	*narXL* + NO_3^- / wt – NO_3^-
1	2	3
a) NarXL activated in response to nitrate (44 operons)		
*narK**	14.3	1.04
*narG**	33.9	0.91
Ogt	12.8	1.51
cyoA	13.1	1.47
soda	6.7	1.05
b) Induced by nitrate even in the *narXL* mutant (7 operons)		
*hcp**	30.0	13.6
*fdnG**	15.2	2.0
yeaR	89.9	5.5
yedF	6.94	2.66
hmpA	20.4	9.07
yibI	5.05	2.31
ytfE	50.5	10.2
c) Induced by nitrate but not significantly NarXL-activated (11 operons)		
*nirB**	9.87	11.02
ydjY	3.36	3.10
d) Induced by nitrate only in the *narXL* mutant (3 operons)		
*napA**	1.01	14.8
*nrfA**	0.32	8.81
yjiM	1.62	3.4
e) Possibly repressed by both NarL and NarP (4 operons)		
*hyaA**	0.09	0.48
appC	0.12	0.39
ynfE	0.03	0.13
*fumB**	0.03	0.27

1	2	3
f) Repressed by nitrate and NarXL (34 operons)		
*dmsA**	0.10	0.61
ansB	0.25	1.18
*frdA**	0.03	0.87
fdoG	0.24	0.84
*hybA**	0.17	0.85

[a] Genes marked with asterisks have been previously reported to be regulated by NarL and / or NarP.

[b] The reported ratios represent the ratio of transcript abundance in the RNA extracted from either the wild type or *narXL* mutant grown in the presence of nitrate compared to transcript abundance in the pool RNA (anaerobically grown parental strain in the absence of nitrate). This table is reproduced from reference (13).

FNR is an activator of dimethylsulfoxide reductase, fumarate reductase, the periplasmic nitrate and nitrite reductases Nap and Nrf, and the cytoplasmic nitrate and nitrite reductases, NarG and NirB. Although both FNR and NarL work positively to activate the *narG* promoter, they work against each other and neutralise each other's action at the *nap* promoter, which regulates the periplasmic nitrate reductase operon. As many promoters are co-regulated by both FNR and one or both of NarL and NarP, the next objective was to extend the microarray studies to determine (a) how many operons are regulated positively or negatively by the NarX-NarL two-component regulatory system; (b) whether nitrate-activated NarL can function as a transcription activator in the absence of FNR; (c) the extent of the sparsely-studied *narP* regulon; and (d) whether nitrate-activated NarP can function as a repressor, a function not previously documented directly, but for which preliminary indication were available (21). Two mutants, one with a *narXL* deletion (these genes are contiguous and are co-transcribed), the other with deletions in both *narXL* and *narP*, were used in further microarray experiments: in each experiment, the same reference RNA was included on every microarray, including those used for the FNR study, so that data from independent experiments over a long time period could be compared. The results summarised in Table 1 and reference (13) revealed that transcription of 51 operons is activated, directly or indirectly, by NarL in response to nitrate, and a further 41 are repressed. However, unlike the FNR regulon, there was no obvious way to distinguish between direct and indirect effects of NarL mediated by metabolic events. Closer inspection revealed that the results were remarkable for at least four reasons.

The first point is that although the *narL* gene was discovered by Valley Stewart 25 years ago (22), most of the operons revealed in our experiments to be regulated by NarL have remained undiscovered. This illustrates the power of genome-wide transcription studies in leading to new discoveries.

Perhaps more significantly, however, is the second point, that only five operons were strongly activated by NarL, and all five of them were discovered long ago.

In contrast, the third point is that our micro-array data showed us for the first time that ***many*** new operons, including genes of unknown function, are repressed by NarL, so this confirms that nitrate really is a major regulator of anaerobic metabolism in *E. coli*.

Greater insight into how pathogens survive in the human body was provided by the transcriptomic studies of the NarP regulon (recall that unlike NarL, NarP is activated by nitrate concentrations similar to those found in the human body). Although only 14 promoters are activated by NarP, at least 37 operons were strongly repressed. This answered the fourth key question: NarP CAN act as an effective transcription off switch, especially when only very low concentrations of nitrate are present in the environment. These are exactly the conditions that are found at the sites of infection in the human body, so they are extremely relevant to mechanisms of pathogenicity, but what does this reveal about how pathogens survive in the human body?

One of the longest standing controversies about anaerobic bacterial metabolism is why bacteria with moderate sized genomes have retained multiple operons for apparently identical functions. Some believe them to be mutually redundant. A more appealing hypothesis is that the encoded functions are required under different environmental conditions, and hence that they fulfil different physiological roles (23,24). Analysis of the *E. coli* NarP regulon provided strong evidence that the

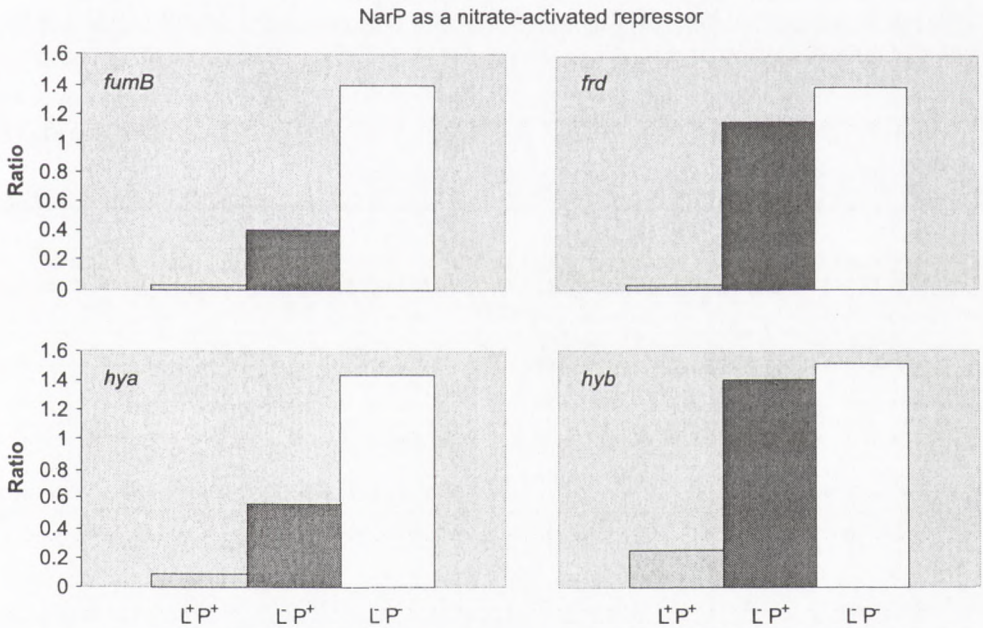

Fig. 2. Repression of *E. coli* genes for dicarboxylate and hydrogen metabolism by either NarL alone, or by both NarL and NarP, during anaerobic growth in the presence of nitrate.

latter explanation is correct. Essentially the role of NarP is to repress expression of genes whose products are required for fermentative growth when low concentrations of nitrate are available to support the more energy-efficient nitrate or nitrite respiratory growth. Thus genes for one of the hydrogenases, HydA, was repressed not only by NarL, but also by NarP, while a second hydrogenase operon, *hydB*, was repressed only by NarL (Fig. 2). This implies that HydA is a key enzyme for life in the human body, while HydB supports fermentative growth in the outside environment. Similarly duplicate pathways for dicarboxylate metabolism during anaerobic growth were differentially repressed in response to nitrate availability either by NarL alone (for growth outside the human body), or by both NarL and NarP.

5. Genes strongly up-regulated in response to nitrite

Throughout the microarray project, many new regulatory phenomena were revealed, not least how *E. coli* might protect itself against toxic nitrogen compounds such as nitric oxide and other reactive nitrogen species generated as part of the host defence mechanisms (13). Genes involved in protection against nitrosative stress were first noticed because some of them were the most highly regulated genes discovered in the project. Noting that they responded differently to FNR, nitrate and nitrite, attempts were made to group them according to their transcriptional control. Some were activated by FNR; others appeared to be repressed by FNR – and in some cases, they were repressed by FNR only in the presence of nitrite (Table 2).

Table 2

Genes required for, or implicated in, reactive nitrogen metabolism

Gene	Ratios[a]			
	Nitrate cultures		Nitrite cultures	
	FNR^+	FNR^-	FNR^+	FNR^-
1	2	3	4	5
(a)				
napA	1.0	0.3	4.6	0.4
nrfA	0.3	0.03	4.7	0.03
(b)				
nirB	10.0	0.2	8.8	0.4
hcp	30.3	1.6	47.8	2.5
(c)				
hmpA	20.4	75.9	29.9	149.7
ytfE	51.4	128.8	38.5	174.7

1	2	3	4	5
ygbA	10.1	17.4	6.7	13.4
(d)				
yeaR	87.5	146.2	3.8	60.4
yibI	5.1	9.5	1.7	7.7
yibH	3.7	6.0	1.4	5.0
cyoA	13.1	61.3	1.1	43.0
ogt	12.8	9.8	1.3	7.9

[a] The reported ratios represent the ratio of transcript abundance in the RNA extracted from the stated growth condition (anaerobic plus nitrate or nitrite) and strain (wild type or *fnr* mutant) compared to transcript abundance in the pool of RNA from the anaerobically grown parental strain. This table is reproduced from reference (13).

The *hcp* gene codes for the prismane or hybrid cluster protein. Some people claim that HCP is a hydroxylamine reductase (25,26); others believe it to be a peroxidase (27), but the evidence is not totally convincing. The microarray data revealed that the *hcp* gene is regulated by FNR activation, but is even more strongly activated by nitrate and nitrite than the *nirB* promoter, suggesting that HCP provides defence against a reactive nitrogen compound generated when the nitrate and nitrite reductases, NarG and NirB, are active.

At the other extreme, it was known that the *hmpA* gene, which codes for a bacterial flavohemoglobin, is repressed by FNR and is more strongly induced by nitrite than by nitrate (13,28). It is known that Hmp provides a nitric oxide reductase activity during anaerobic growth and protects bacteria when exposure to NO is severe (29,30). Microarray data revealed that the *ytfE* gene is regulated almost co-ordinately with *hmpA*, suggesting that YtfE protein also provides protection against reactive nitrogen species under the same conditions that HmpA is required. Ligia Saraiva's team have already shown that this conclusion is correct (31,32).

The *yoaG-yeaR* operon is also strongly induced by nitrite, especially in the *fnr* mutant, so this operon is also likely to be involved in protection against reactive nitrogen species. Other operons also encode genes of unknown function that, on the basis of their regulation, are likely to play a role in protecting *E. coli* against reactive nitrogen species.

6. The nitrosative stress response repressor, NsrR

Some operons are repressed by FNR, but this repression becomes most obvious only during anaerobic growth in the presence of nitrite. Nitrite still induced expression of some of these operons even in double mutants deleted for *narXL* and for *narP*. This indicated that there is yet another control circuit interacting with FNR

and toxic nitrogen compounds derived from nitrite. The missing regulator is the repressor, NsrR, which is inactivated by nitric oxide (33,34). This repressor protein, a member of the Rrf2 family of transcription factors, represses expression of about 20 genes in 9 *E. coli* operons, and possibly also activates one or more transcripts (35).

7. How pathogenic *Neisseria* survive oxygen starvation and nitrosative stress

Other successful pathogens like the pathogenic neisseria that cause meningitis and the sexually transmitted disease, gonorrhoea, must also be able to survive in different sites in the human body: some sites are aerobic; others are anaerobic. But they must also avoid being killed by products of their own metabolism, for example nitric oxide generated as an intermediate of denitrification, as well as by reactive oxygen species and reactive nitrogen compounds fired at them by host defence mechanisms. *Neisseria gonorrhoeae* (the gonococcus) is essentially a microaerophile: it grows best with a low concentration of oxygen, and uses essentially a single cytochrome oxidase, cytochrome cbb_3, to reduce oxygen to water. It can also survive anaerobically using a partial denitrification pathway in which nitrite is reduced first to nitric oxide, and nitric oxide is then reduced to nitrous oxide (36-39).

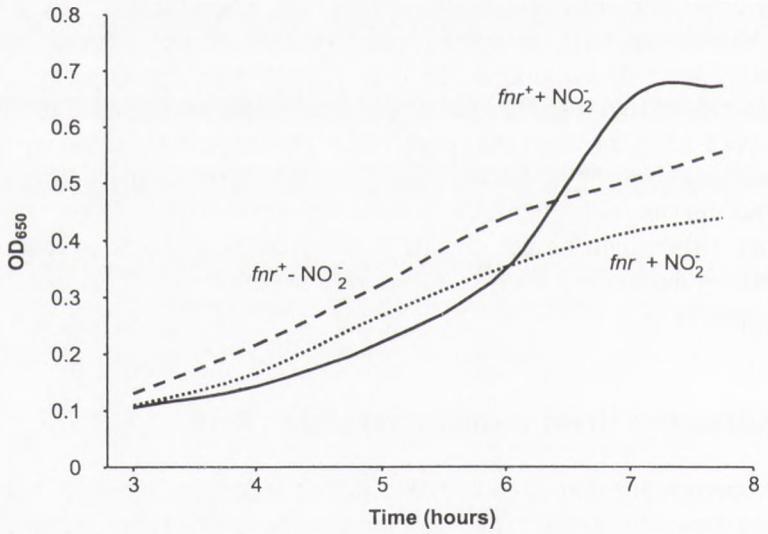

Fig. 3. Oxygen limited growth of the *Neisseria gonorrhoeae fnr* mutant in the presence of nitrite, and of its *fnr*$^+$ parent, strain F62, in the presence or absence of nitrite. Modified from reference (41).

As in *E. coli*, there a gonococcal FNR protein, which contains an iron-sulphur centre that acts as an oxygen sensor (40). There is also a two-component regulatory system that we designated NarQ-NarP (38). The pathogenic Neisseria have small genomes and are relatively fastidious in their growth requirements, presenting challenges for experiments to reveal their physiology and biochemistry. Their only mode of oxygen-limited growth required FNR-dependent expression of the *aniA* (*nirK*) gene that encodes a copper-containing nitrite reductase, so gonococcal *fnr* mutants cannot grow anaerobically. Once more whole-genome transcriptomic experiments could be invalidated by the growth rate problem: in oxygen limited media supplemented with nitrite, the *fnr* mutant will grow more slowly that the parent strain (Fig. 3). To solve this problem, bacteria were grown with poor aeration, and as a control the parent strain was grown both with and without nitrite in the medium. In this way, it was possible to differentiate between transcripts that respond to changes in growth rate, and genes that are regulated positively or negatively by the FNR transcription factor.

At most 14 transcripts were found to be activated by FNR, and a further 6 were repressed (41), but again these microarray data included genes that are regulated indirectly rather than directly by the transcription factor being studied. We therefore completed a bioinformatic analysis of the 200 bases upstream of each of these genes searching for FNR-binding sites, and we also constructed a gonococcal strain with the *fnr* gene fused in-frame to a FLAG tag for which a commercial antibody is available. The FLAG tag does not interfere with the binding of FNR to its DNA target sites, so it was possible to develop the first gonococcal chromatin immunoprecipitation (ChIP) experiments to locate FNR on the gonococcal chromosome (42). The conclusion was that, apart from genes involved or implicated in the truncated denitrification pathway, very few other transcripts are regulated as part of the gonococcal FNR regulon.

Similar experiments with a gonococcal *narQP* mutant revealed that the gonococcal NarP regulon is even smaller than the FNR regulon, including at most five transcripts (43). Two transcripts, *aniA* and *norB*, appeared to be activated by NarQ-NarP, and three, including the *narQP* operon, were repressed. However, it was reported seven years ago that the *norB* gene is not regulated by FNR during anaerobic growth, but transcription is activated by another transcription factor in response to the availability of nitric oxide (44). This effect of NarP on *norB* transcription is therefore almost certainly an indirect effect: NarP activates *aniA*, AniA catalyses the reduction of nitrite to NO; and it is the NO produced as a result of nitrite reduction that activates *norB* transcription.

In *E. coli*, NarQ is an environmental sensor that detects and responds to very low concentrations of nitrate, but it does not respond to nitrite (45). In contrast, the gonococcus cannot detect nitrate, but responds to nitrite. So how does the gonococcus detect and respond to nitrite? We initially assumed that it was the gonococcal NarQ that responds to nitrite, but realised that this is incorrect when it was

shown that a gonococcal *narQP* deletion mutant still responds to nitrite. This raised the possibility, which is now known to be correct, that *aniA* transcription responds not to nitrite, but to nitric oxide generated as the product of nitrite reduction. If so, a small molecule like NO cannot regulate transcription on its own, so there must be another protein that enables gonococci to sense and respond to nitrite. This protein was again identified in the bioinformatic analysis of all bacteria for which complete genome sequences were available (34) and that were known to respond to nitrosative stress; it was the neisserial NsrR protein. Rodionov et al. (34) predicted that, as in *E. coli*, there is a gonococcal transcription factor, NsrR, that is a repressor of genes required for defence against nitric oxide, and that NO binds to NsrR to inactivate it. The prediction that NsrR regulates expression of both of the denitrification genes, *aniA* and *norB*, as well as a gene of unknown function, *dnrN*, was confirmed (41). However, a fourth prediction, that NsrR regulates expression of *narQ-narP*, has been shown to be incorrect (43).

8. Conclusions

In summary, the combination of microarray experiments and targeted gene analysis has revealed both striking similarities and fascinating differences between two major groups of pathogenic bacteria: enteric bacteria that are able to adapt to life both inside and outside an animal host, and the obligate human pathogen, *Neisseria gonorrhoeae*. It has been reported that *E. coli* YtfE, which is a homologue of DnrN, repairs damage to iron-sulphur centres by removing the NO groups (32). DnrN should also be able to remove NO from transcription factors like FNR and NsrR, providing a possible explanation of how bacteria exploit their NsrR regulons to repair damage introduced during nitrosative stress. If so, proteins like NsrR, YtfE, DnrN, AniA and NorB, and even the transcription factors FNR and NarP, all provide possible targets for biopharmaceutical intervention, and hence illustrate the potential value of whole-genome transcriptomic approaches to biotechnology.

Acknowledgements

The authors gratefully acknowledge the technical support of Antony Jones at the Birmingham Functional Genomics Laboratory for the facilities used in this study, and Dr David Grainger for help with the chromatin immunoprecipitation experiments. The research was funded by the UK Biotechnology and Biological Sciences Research Council through grants JIF13209, P20180 to JAC and EGA16107.

Literature

1. Gill R. T., Valdes J. J., Bentley W. E., (2000), Metanol. Eng., 2, 178-189.
2. Kurland C. G., Dong H. J., (1996), Mol. Microbiol., 21, 1-4.
3. Iuchi S., Lin E. C. C., (1988), Proc. Natl. Acad. Sci. USA, 85, 1888-1892.
4. Iuchi S., Lin E. C. C., (1993), Mol. Microbiol., 9, 9-15.

5. Spiro S., Guest J. R., (1990), FEMS Microbiol. Rev., 6, 399-428.
6. Spiro S., Guest J. R., (1991), Trends Biochem. Sci., 16, 310-314.
7. Green J., Guest J. R., (1994), Mol. Microbiol., 12, 433-444.
8. Sawers G., Suppmann B., (1992), J. Bacteriol., 174, 3474-3478.
9. Tseng C. P., Albrecht J., Gunsalus R. P., (1996), J. Bacteriol., 178, 1094-1098.
10. Compan I., Touati D., (1994), Mol. Microbiol., 11, 955-964.
11. Salmon K., Hung S.-P., Mekjian K., Baldi P., Hatfield G. W., Gunsalus R. P., (2003), J. Biol. Chem., 278, 29837-29855.
12. Kang Y., Weber J. D., Qiu Y., Kiley P. J., Blattner F. R., (2005), J. Bacteriol., 187, 1135-1160.
13. Constantinidou C. C., Hobman J. L., Patel M. D., Penn C. W., Cole J. A., Overton T. W., (2006), J. Biol. Chem., 281, 4802-4808.
14. Rabin R. S., Stewart V., (1993), J. Bacteriol., 178, 3259-3268.
15. Baikalov I., Schroder I., Kaczor-Grzeskowiak M., Grzeskowiak K., Gunsalus R. P., Dickerson R. E., (1996), Biochemistry, 35, 11053-11061.
16. Tyson K. L., Bell A. I., Cole J. A., Busby S. J. W., (1993), Mol. Microbiol., 7, 151-157.
17. Tyson K. L., Cole J. A., Busby S. J. W., (1994), Mol. Microbiol., 13, 1045-1055.
18. Cavicchioli R., Schroder I., Constanti M., Gunsalus R. P., (1995), J. Bacteriol., 177, 2416-2424.
19. Darwin A. J., Tyson K. L., Busby S. J. W., Stewart V., (1995), Mol. Microbiol., 25, 583-595.
20. Kalman L. V., Gunsalus R. P., (1989), J. Bacteriol., 171, 3810-3816.
21. Goh E. B., Bledsoe P. J., Chen L. L., Gyaneshwar P., Stewart V., Igo M. M., (2005), J. Bacteriol., 187, 4890-4899.
22. Stewart V., (1982), J. Bacteriol., 151, 1320-1325.
23. Cole J., (1996), FEMS Microbiol. Lett., 136, 1-11.
24. Potter L., Millington P., Thomas G., Cole J., (1999), Biochem. J., 344, 77-84.
25. Wolfe M. T., Heo J., Garavelli J. S., Ludden P. W., (2002), J. Bacteriol., 184, 5898-5902.
26. Cabello P., Pino C., Olmo-Mira M. F., Castillo F., Roldan M. D., Moreno-Vivian C., (2004), J. Biol. Chem., 279, 45485-45494.
27. Almeida C. C., Romao C. V., Lindley P. F., Teixeira M., Saraiva L. M., (2006), J. Biol. Chem., 281, 32445--32450.
28. Cruz-Ramos H., Crack J., Wu G., Hughes M. N., Scott C., Thompson A. J., Green J., Poole R. K., (2002), EMBO J., 21, 235-3244.
29. Hausladen A., Gow A. J., Stamler J. S., (1998), Proc. Natl Acad. Sci. USA, 14100-14105.
30. Kim S. O., Orii Y., Lloyd D., Hughes M. N., Poole R. K., (1999), FEBS Lett., 445, 389-394.
31. Justino M. C., Almeida C. C, Goncalves V. L., Teixeira M., Saraiva L. M., (2006), FEMS Microbiol. Lett., 257, 278-284.
32. Justino M. C., Almeida C. C, Teixeira M., Saraiva L. M., (2007), J. Biol. Chem., 282, 10352-10359.
33. Bodenmiller D. M., Spiro S., (2006), J. Bacteriol., 188, 874-881.
34. Rodionov D. A., Dubchak I. L., Arkin A. P., Alm E. J., Gelfand M. S., (2005), PLoS Comput Biol., 1, e55.
35. Filenko N., Spiro S., Browning D., Squire D., Overton T., Cole J., Constantinidou C., (2007), J. Bacteriol., 189, 4410-4417.
36. Knapp J. S., Clark V. L., (1984), Infect. Immun., 46, 176-181.
37. Clark V. L., Knapp J. S., Thompson S., Klimpel K. W., (1988), Microb. Pathog., 5, 381-390.
38. Lissenden S., Mohan S., Overton T., Regan T., Crooke H., Cardinale J. A., Householder T. C., Adams P., O'Connor C. D., Clark V. L., Smith H., Cole J. A., (2000), Mol. Microbiol., 37, 839-855.
39. Householder T. C., Belli W. A., Lissenden S., Cole J. A., Clark V. L., (1999), J. Bacteriol., 181, 541-551.
40. Overton T., Reid E. G. F., Foxall R., Smith H., Busby S. J. W., Cole J. A., (2003), J. Bacteriol., 185, 4734-4747.
41. Whitehead R. N., Overton T. W., Snyder L. A. S., McGowan S. J., Smith H., Cole J. A., Saunders N. J., (2007), BMC Genomics 8, 35.
42. Grainger D. C., Overton T. W., Reppas N., Wade J. T., Tamai E., Hobman J. L., Constantinidou C. C., Struhl K., Church G., Busby S. J., (2004), J. Bacteriol., 186, 6938-6943.

43. Overton T., Whitehead R., Li Y., Snyder L. A. S., Saunders N. J., Smith H., Cole J. A., (2006), J. Biol. Chem., 281, 33115-33126.
44. Householder T. C., Fozo E. M., Cardinale J. A., Clark V. L., (2000), Infect. Immunol., 68, 5241-5246.
45. Wang H., Tseng C.-P., Gunsalus R. P., (1999), J. Bacteriol., 181, 5303-5308.