High-resolution analysis of the m-xylene/toluene biodegradation subtranscriptome of *Pseudomonas putida* mt-2

Juhyun Kim, Danilo Pérez-Pantoja,† Rafael Silva-Rocha,‡ Juan Carlos Oliveros and Víctor de Lorenzo*  
Systems Biology Program, Centro Nacional de Biotecnología-CSIC, Campus de Cantoblanco, Madrid, Spain.

**Summary**  
*Pseudomonas putida* mt-2 metabolizes m-xylene and other aromatic compounds through the enzymes encoded by the *xyl* operons of the TOL plasmid pWW0 along with other chromosomally encoded activities. Tiling arrays of densely overlapping oligonucleotides were designed to cover every gene involved in this process, allowing dissection of operon structures and exposing the interplay of plasmid and chromosomal functions. All *xyl* sequences were transcribed in response to aromatic substrates and the 3′-termini of both *upper* and *lower* mRNA operons extended beyond their coding regions, i.e. the 3′-end of the *lower* operon mRNA penetrated into the convergent *xylS* regulatory gene. Furthermore, *xylR* mRNA for the master *m*-xylene responsive regulator of the system was decreased by aromatic substrates, while the cognate *upper* operon mRNA was evenly stable throughout its full length. RNA sequencing confirmed these data at a single nucleotide level and refined the formerly misannotated *xylL* sequence. The chromosomal *ortho* route for degradation of benzoate (the *ben*, *cat* clusters and some *pca* genes) was activated by this aromatic, but not by the TOL substrates, toluene or *m*-xylene. We advocate this scenario as a testbed of natural retroactivity between a pre-existing metabolic network and a new biochemical pathway implanted through gene transfer.

**Introduction**  
*Pseudomonas putida* mt-2, host of the TOL plasmid pWW0, is able to assimilate *m*-xylene or toluene through a set of enzymes encoded by the catabolic plasmid (Worsey and Williams, 1975; Assinder and Williams, 1990; Marques et al., 1999). Two plasmid-encoded operons account for the entire metabolic process that leads to biodegradation of such aromatic substrates (Fig. 1). The *upper* operon determines enzymes that convert *m*-xylene or toluene into 3-methylbenzoate (i.e. 3MBz) or benzoate, respectively, and the *lower* operon enzymes catalysing the conversion of 3MBz into 3-methylcatechol or benzoate into catechol and its ensuing *meta* ring cleavage, which ultimately leads to intermediates of the TCA cycle (Assinder and Williams, 1990; Ramos et al., 1997). At the same time, benzoate is also degraded via the chromosomally encoded *ortho* pathway genes, i.e. *ben*, *cat* and *pca* (Harwood and Parales, 1996; Jimenez et al., 2002). The TOL plasmid-encoded transcriptional factors called XylR and XylS control the expression of the corresponding *upper* and *lower* operons. The principal regulatory interactions between metabolic and regulatory components of the TOL network are sketched in Fig. 1. In addition to these plasmid-encoded regulators and promoters, a number of host factors influence expression of the TOL functions, e.g. the integration host factor and the HU nucleoid-associated proteins (Perez-Martin and de Lorenzo, 1995) and the TurA and PprA products (Rescalli et al., 2004; Vitale et al., 2008). Besides, the TOL system is subject to various catabolite repression and physiological control checks (Velazquez et al., 2004; Aranda-Olmedo et al., 2005; Moreno et al., 2010). In this scenario, pWW0-borne proteins and host native regulators and small molecules interact to shape a complex metabolic and transcriptional circuit that has recently been reviewed (Silva-Rocha et al., 2011a) and (at least partially) modelled (Silva-Rocha et al., 2011b). Specific nodes of such a network have been studied in extraordinary molecular detail (Koutinas et al., 2010; Silva-Rocha et al., 2011a), and even more regulatory devices are likely to appear. The reason for the apparently unnecessary intricacy of the genetic circuitry found in the TOL system is still unexplained (Silva-Rocha et al., 2011b).  

Received 2 June, 2015; revised 30 August, 2015; accepted 12 September, 2015. *For correspondence. E-mail vdlorenzo@cnb.csic.es; Tel. (+34-91) 585 45 36; Fax (+34-91) 585 45 06. Present addresses: †Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Biológicas, Universidad de Concepción, Chile; ‡Department of Cell and Molecular Biology, Ribeirão Preto Medical School, Universidade de São Paulo 14049-900, Brazil.

© 2015 Society for Applied Microbiology and John Wiley & Sons Ltd
The TOL regulatory and metabolic network for m-xyylene and toluene biodegradation borne by the catabolic enzymes in the TOL plasmid. The top part of the figure shows the principal actors of the process. In the presence of m-xyylene or toluene, XyR (expressed from Pr) activates both the σ54-dependent promoters Pu (which transcribes the upper pathway operon) and Ps, which drives expression of xylS (Perez-Martin and de Lorenzo, 1995; Bertoni et al., 1998). Since the Pr promoter overlaps the XyR binding sites of the divergent Ps promoter, XyR expression is negatively auto-regulated (Bertoni et al., 1998). In turn, the XylS protein activates Pm, the promoter of the lower operon (Bertoni et al., 1998; Marques et al., 1999; Dominguez-Cuevas et al., 2008; 2010). Pr, Ps, Pu and Pm promoters are indicated in front of each cognate gene or operon. By default, XylS is expressed at a low level but its transcription is enhanced by XyR in the presence of m-xyylene. The bottom part of the figure signifies that the benzoate produced by partial metabolism of toluene can be channelled through either the meta ring-cleavage pathway encoded by the lower TOL operon or through the ortho-ring-fission route encoded in the host’s chromosome.

One approach to addressing the exaggerated complexity of the TOL network is to examine the regulatory behaviour of the entire system rather than its parts. Velazquez and colleagues (2005) and Dominguez-Cuevas and colleagues (2006) first documented the genome-wide transcriptional response of P. putida mt-2 when cells faced various aromatic compounds – whether they were TOL substrates or not. While these studies were instrumental in verifying the layout of the regulatory circuit proposed earlier on the basis of one-at-a-time reporter fusions to the TOL promoters (Hugouvieux-Cotte-Pattat et al., 1990; Marques et al., 1994), the technology of the time was insufficient to examine key aspects of the m-xylene subtranscriptome, e.g., the fine structure of the TOL operons, the regulatory cross-talk of plasmid and host metabolic genes, and the action of post-transcriptional RNA-based regulation devices.

To overcome this impasse and generate a high-definition expression profile of such catabolic genes, we have designed and exploited in this work a tiling array (Selinger et al., 2000; Guell et al., 2009) that spans 14 704 base pairs (bp) of DNA, including the two strands of the corresponding xyl genes of the TOL system in its entirety, as well as the chromosomal ortho-pathway genes involving ben, cat and pca sequences. Experiments with this array were complemented and validated with deep RNA sequencing (RNA-seq; Passalacqua et al., 2009; Toledo-Arana et al., 2009; Yoder-Himes et al., 2009; Kim et al., 2013) that gave further insights into the regulatory components of the TOL catabolic system - in particular the interplay between the pWW0-encoded and the chromosomally encoded pathways for metabolism of benzoate in P. putida mt-2.

Results and Discussion

Benchmarking tiling arrays and RNA-seq for visualizing the structure of xyl transcripts

As the starting point of this work, the structure and abundance of the transcripts that originate in each of the four promoters (Pu, Pm, Pr, Ps) in front of the xyl genes of the TOL plasmid (Fig. 1) were examined by combining tiling array technology and deep RNA-seq. In one case, a custom tiling microarray was designed to contain probes corresponding to the complete xyl operons of the TOL plasmid, as well as chromosome clusters covering both branches of β-ketoadiapate pathway comprising the fcs cluster, the pca genes, the ben genes and the cat genes. Note that tiling arrays are strand specific, what allows gaining details on transcript structure that may not be evident with other techniques (Huber et al., 2006; Miyakoshi et al., 2009; Mader et al., 2011). Once we had such arrays in hand, P. putida mt-2 strain was cultured to early exponential growth phase (OD600 ∼ 0.3) in succinate-
supplemented minimal medium and then exposed to saturating vapours of a 1:2 dilution of \( m \)-xylene (the canonical substrate of the TOL network) in dibutyl phthalate. Two hours later, total RNAs were isolated and the cDNAs synthesized were used for hybridization to the tiling array chip. In addition to these conditions, we also prepared a second control sample by culturing the cells in succinate as the sole carbon source and then let them run out of carbon for an extended period of time (96 h). Such prolonged carbon starvation is predicted to empty cells of the carbon for an extended period of time (96 h). Such prolonged carbon starvation is predicted to empty cells of the aromatic inducer, particularly in the case of the lower TOL operon. In addition, when using succinate only as a control, transcription intensity was not evenly distributed through the genes of the lower pathway. Under the same conditions (i.e. using succinate-cultured cells as control), transcription of the upper pathway genes was uniform throughout the entire length of the cognate DNA sequence (Fig. 2A). By contrast, the transcriptional profiles of both the upper and the lower operons were evenly distributed when the RNA from C-runout cells was used for normalization (Fig. 2B).

In sum, these data accredit that the lower pathway genes are expressed to a low but detectable level in succinate-amended minimal medium in the absence of inducers or substrates to be. In view of these results, we adopted the C-runout condition as the default reference for subsequent analyses (Fig. 2B, lower panel), as it seems to faithfully reflect the no-induction baseline of the upper and lower operons. To verify this, we revisited the net \( m \)-xylene-dependent expression profile of \( xyl \) genes with deep RNA-seq (Fig. 3). To this end, the cDNA libraries from RNA extracted from the bacteria under examination cells were sequenced using the Illumina Genome Analyzer Iix system, and the reads were mapped to the pWW0 plasmid sequence (and also the \( P. putida \) KT2440 chromosome, see below). The alignment to the upper and lower pathway genes of the TOL plasmid was visualized

![Fig. 2. High-density tiling array-based transcriptome profile of TOL xyl genes in P. putida mt-2.](image-url)
using the IGV (integrated genome viewer) software (Thorvaldsdottir et al., 2012) as shown in Fig. 3A, and the transcription activity was normalized (Fig. 3B) by trimmed mean of M values (TMM; Robinson and Oshlack, 2010). Although the expression intensity level cannot be compared quantitatively with the tiling array data, the transcription profiles from RNA-seq results were fully consistent with the array-based expression pattern (e.g. compare Figs 2 and 3). The mutually complementary tiling arrays and deep RNA-seq were thereafter adopted as a reliable approach to examine the whole \textit{P. putida} mt-2 subtranscriptome involved in biodegradation of aromatic substrates.

**Definition of operon structure of TOL genes, novel transcriptional units and correction of gene annotation**

Improved resolution by high-density tiling microarrays and deep RNA-seq provides insights into genomic elements, e.g. discovery of new genes, definition of untranslated regions and operon structure (Sorek and Cossart, 2010). We thus proceeded to examine the structure of the TOL operons as they appear when cells are exposed to \textit{m}-xylene. Using these methods, we found some new transcriptional features in TOL \textit{xyl} genes. The first novelty was the presence of a previously unreported and apparently separate \textasciitilde 1000 bp-long transcription unit downstream of the \textit{xylN} gene (Fig. 2B; Fig. S1). Consistent with the tiling array data, this transcription unit was also observed with the RNA-seq, as reads mapped to the same region at a high rate (Fig. 3A). This mRNA extension unit overlaps in part with an open reading frame of unknown function (\textit{orf126}) that is also found in other TOL plasmids such as pD2RT, pDK1 and pWW53. \textit{orf126} is adjacent to the downstream end of the \textit{upper} operon and transcription from \textit{Pu} runs through its complete sequence (see Fig. S1 for a blow-up of the region). However, the separate transcript occurs regardless of growth conditions (Fig. 4) and might be therefore constitutively expressed – or induced by an unknown metabolic effector. 

Apart of this trait, both transcriptomic approaches confirmed the already recognized operon structures of the \textit{xyl} genes (Harayama et al., 1984; 1989; Marques et al., 1993; Williams et al., 1997). Furthermore, the relatively even signals through the whole length of both the \textit{upper} and the \textit{lower} operons suggested that the bulk of the \textit{xyl} transcripts generated under the conditions tested was long, polycistronic mRNAs, with little indication of differential stability, processing or additional internal promoters. Still, under some conditions (Fig. 4), we could detect an apparent boundary between the \textit{xylXYZL} portion of the \textit{meta} pathway transcripts and the rest of the operon, a
circumstance that has been reported before (Harayama and Rekik, 1990; Marques et al., 1993). However, this discontinuity disappeared when the system was fully induced with \( m \)-xylene (Fig. 2B). We also found that the xylL sequence present in the \( P. \ putida \) mt-2 strain under examination had a manifest divergence in respect to the corresponding data in the literature (Greated et al., 2002).

From the RNA-seq results, no sequencing read could be aligned to a 36 bp segment of the xylL gene annotated in NCBI (GenBank: AJ344068.1; Fig. S2). To test whether this was caused by a technical error in the RNA-seq or the problem stems from the gene sequence proper, we cloned the partial xylL gene of our \( P. \ putida \) mt-2 clone and analysed the DNA of that region. This revealed that the section between coordinates 52320–52355 in the plasmid was absent in our template DNA (Fig. S2). Since the product of xylL is necessary for processing cis-dihydrodiols generated during dioxygenation of 3MBz (Voss et al., 1990) and our strain grows perfectly well on \( m \)-xylene and 3MBz, it is unlikely that the loss of this 36 bp sequence could be attributed to a spontaneous deletion in the laboratory. Moreover, all other annotated xylL orthologues that are found in databases lack also the segment missing in our pWW0 variant (not shown).

Overlapping transcription between the 3’ ends of xylS and the lower operon gene

As mentioned above, tiling arrays expose details on transcript structure that may go unnoticed with bulk RNA-seq. One intriguing feature of the regulatory architecture of the TOL network is that the gene encoding the regulatory protein XylR is located in the pWW0 frame adjacent to one target promoter (\( P_s \)) but distal to the other (\( P_u \)). In contrast, the second regulatory factor (xylS) is away from the promoter it controls (\( P_m \)), but it is located adjacent and opposite to the end of the terminus of the meta operon whose transcription is driven by \( P_m \) (Fig. 1).

Inspection of the convergent strand-specific signals originated at the 3’ termini of the meta operon and the xylS gene (Figs 2B and 5A) revealed that they both overlap by at least 350 nt. Most of the overlap occurs in non-coding regions, but the shared segment reaches out short portions of either 3’-end gene sequences. Since such a confluence could have a regulatory value, we verified this state of affairs with Northern blot analysis of transcript production at this overlapping region with strand-specific RNA probes prepared with the \textit{in vitro} transcription method described in Experimental procedures. These
probes were designed for either strand of the two non-coding sequences between the xylS and the end of the lower operon region. Both probes were directly hybridized with total RNAs from m-xylene-exposed cells (Fig. 5A, red line). Several signals originated in the non-coding regions were detected by the antisense probe, indicating that they were indeed transcribed (in contrast with the more modest signals for the sense probe; Fig. 5B). When an internal probe for xylS was used (Fig. 5B, blue line), only the sense probe delivered a signal. This meant that the untranslated 3’ ends of the xylH and xylS transcripts could reach out each other well, thereby raising the question of whether such a convergent transcription of the region of confluence results in a pairing structure between sense and antisense transcripts that could have a regulatory effect (Thomason and Storz, 2010). To explore this possibility, we constructed an artificial antisense xylS expression system as sketched in Fig. 6A. The business part of plasmid pSEVA224-XS consists of a DNA segment spanning xylS and its extended 3’-end bound upstream by the native Ps promoter of xylS but added downstream with a reverse IPTG-inducible Pptrc promoter. This plasmid was introduced into the P. putida strain harbouring the Pm promoter fused to a bi-cistronic gfp-lacZ reporter system (Silva-Rocha and de Lorenzo, 2012). Should the untranslated mRNA downstream of the xylH gene (expressed from Pptrc in plasmid pSEVA224-XS) influence expression of xylS, the prediction is that we should see an effect on the activity of the Pm–GFP fusion of the host strain. To test this possibility, we inspected GFP levels in single cells with flow cytometry. As shown in Fig. 6B, Pm activity and distribution in individual bacteria was impervious to the increase of anti-sense RNA, as GFP expression followed in all cases a graded pattern with no changes in activity or expression noise (Fig. 6B). These results indicated that overlapping transcripts downstream of xylS lack any gross influence on the behaviour of the TOL regulatory network.

Response of the TOL network to gratuitous induction and to 3MBz

To complete the picture of the response of the TOL network to its canonical substrate, m-xylene, we examined the expression pattern of the corresponding gene clusters under two additional conditions with tiling arrays as before. In one case, we exposed the cells to o-xylene. This is a non-metabolizable but strong effector of the master XylR regulator that should directly trigger the targets of this transcription factor (i.e. the Ps and the Pu promoters of the TOL system; Fig. 1). Indirectly, the Pm promoter is expected to be activated in the presence of o-xylene also due to the regulatory effect of overproduced XylS even in the absence of its metabolic effector (the metabolic intermediate 3MBz). Figure 4A shows the expression profile of each portion of the TOL segments of pWW0 plasmid in response to o-xylene. This gratuitous XylR effector triggered the expression of the upper operon to similar or even higher levels than the genuine pathway substrate, m-xylene (compare Fig. 4A with Fig. 2B). The same held true for the XylR-activated expression of the xylS gene. In contrast, the volume of transcripts of the meta operon originated in the Pm promoter was significantly less with o-xylene than with m-xylene. This made sense, as activation of Pm under these conditions stems only from overproduction of XylS, thereby lacking the cascade effect caused by 3MBz when it is produced out of metabolism of m-xylene (Fig. 1). Regardless of differences in levels, comparison of the mRNA profiles of the xyl genes induced with either m-xylene or o-xylene
indicated the organization and relative stability of segments throughout the mRNAs to be virtually identical. This suggested that there is little or no connection between expression of these catabolic genes and actual catabolic activity and that neither the substrate nor its metabolic intermediates influence production of full-length mRNAs through riboswitches or other regulatory devices.

The second condition tested involved addition of 3MBz to P. putida mt-2 cultures. As mentioned above, this aromatic intermediate is produced metabolically upon the action of the upper pathway gene products on m-xylene, but can also be added exogenously for separate induction of the lower TOL operon (Gonzalez-Perez et al., 2004). Figure 4B shows the result of such an induction experiment. While the upper TOL pathway remained silent and expression of the xylIS was not affected, transcription of the lower TOL pathway was upregulated but to a much lower level than the same in cells induced with m-xylene (compare Fig. 4B with Fig. 2B). These data were consistent with the current regulatory model of the TOL network, in which the Ps promoter is in reality composed of two tandem promoters (Galgles et al., 1996): one low constitutive (Ps2) and one inducible at high levels by m-xylene/XylR (Ps1). When 3MBz is added to the cells, only low levels of XylIS are available to activate the expression of the lower pathway. This thus happens to a limited extent, as the cognate regulatory node lacks in this case the amplifying effect of overproducing XylIS.

**Transcriptomic fingerprint of toluene catabolism through the TOL pathway**

The data examined above concern exclusively the transcriptomic background of degradation of m-xylene or 3MBz, the canonical substrates of the TOL system. While processing of m-xylene by upper pathway enzymes generates intermediates alien to the innate metabolic network of P. putida, the final product of the upper route 3MBz (Fig. 1) is a faulty substrate of the chromosomally encoded ortho-fission pathway, which generates dead-end metabolites (Rojo et al., 1987). This possible metabolic misrouting is solved through a regulatory device that ensures that 3MBz induces the plasmid-encoded meta-fission pathway but not the ortho-cleavage route encoded by the ben and cat genes (Perez-Pantoja et al., 2014; Silva-Rocha and de Lorenzo, 2014). When cells are grown in toluene, however, the metabolic and regulatory scenario changes, as the action of the upper TOL pathway on this aromatic produces benzoate which induces both the lower TOL pathway and the chromosomal ben genes (Fig. 7). How do these two routes manage their simultaneous action on benzoate? To answer this question, we started by examining expression of TOL plasmid-coded genes when P. putida mt-2 is exposed to toluene under the same conditions used before for m-xylene. As shown in Fig. 4D, when cells were cultured in succinate-minimal medium and treated with toluene, the gross expression pattern of the xyl genes was indistinguishable from that observed before with the m-xylene condition (compare Fig. 4D with Fig. 2B); the upper and lower operons and the xylS genes were highly expressed, whereas xylR appeared to be downregulated (Fig. 4B). Endogenously produced benzoate was thus inducing the meta pathway. When benzoate was instead added as a purified chemical to succinate-grown cells, there was induction only of the lower operon (Fig. 4E) but to a lower extent than expression observed with toluene (compare Fig. 4E with Fig. 4D). In this respect, the effects of toluene/benzoate on the regulatory network of the xyl genes are virtually indistinguishable of those of m-xylene/3MBz. Figure S3
summarizes the relative responsiveness of each of the transcriptional units that compose the TOL system when *P. putida* mt-2 cells are exposed to pathway head substrates (*m*-xylene, toluene), added intermediates (3MBz and benzoate) and the gratuitous inducer o-xylene. Expectedly, the more pronounced differences happen in the lower pathway. Yet, expression levels in response to 3MBz and benzoate (whether endogenously produced from head substrates or added to the medium as such) were almost identical — as if the ortho route was unable to divert much of the benzoate into this chromosomally encoded pathway. The experiments below allowed us to trace the destiny of benzoate — as compared with 3MBz, when the substrate bifurcates between two different catalytic pathways in this bacterium.

*The benzoate biodegradation subtranscriptome of *P. putida* mt-2*

Figure 7 sketches the whole of genes known to participate in degradation of *m*-xylene, toluene and their intermediates in *P. putida* mt-2 (Jimenez et al., 2002). On one hand, catechol can be generated by benzoate dioxygenase encoded by the *benABCD* genes of pWW0 and the benzoate dioxygenase chromosomally encoded by *benABCD*. On the other hand, once catechol is formed, it can proceed either through the rest of the meta-cleavage pathway determined by the single, polycistronic lower TOL operon or by the alternative route formed by biochemical activities determined by the independent gene cluster *catABC* and additional enzymes recruited from a separate protocatechuate degradation pathway (*pcaDUF*; Jimenez et al., 2002). Since these chromosomally encoded routes are independently regulated by pathway substrates and cognate metabolic intermediates, their transcription is a plausible proxy of their biochemical activities.

On this background, we set out to examine with tiling array technology the expression of the *ben*, *cat* and *pca* gene clusters in response to the same panel of inducers than those used for inspection of the *xyl* genes above. Figure 8 shows the results, which embody a considerable amount of information. The control conditions were those of Fig. 8A, in which succinate-grown bacteria were supplied with exogenously added benzoate. The tiling array signals show a considerable induction of the *benABCD* segment (determining the key benzoate dioxygenase activity), *catB*, and *catA* genes and to a lesser level, the rest of the *ben* (*benKA*), *cat* (*catABC*), and *pca* (*pcaDUF*) clusters. Since the signal intensities along the cognate gene sequences are calculated on the basis of the C-runout conditions, comparison of Fig. 8A with Fig. 8B (succinate-grown cells, no other additions) allowed us to discriminate genuine induction by benzoate from default constitutive expression. Under such a criterion, expression of the whole *ben* and *cat* operons seems to be genuinely triggered by benzoate, while that of *pca* genes was generally constitutive or semi-constitutive. No changes were detected either in other connected pathways that share *pca* genes, e.g. coumaric acid, caffeic acid and ferulic acid (Jimenez et al., 2002; Fig. 7). These compounds are channelled towards protocatechuate by enzymes encoded by the *fcs* cluster (*fcs, vdh, ech* genes) and the *pobA* gene and further taken towards β-ketoadipate by *pca*-encoded products. As shown in Fig. S4, these gene clusters were expressed roughly constitutively regardless of culture conditions. This made us focus our attention on the *ben* and *cat* operons.

As shown in Fig. 8A, *ben* genes appear as a polycistronic operon with at least two differentially stable segments, the most conspicuous being the one encompassing the *benABCD* genes. That the *cat* genes are also induced signifies that at least part of the catechol resulting from the activity of benzoate dioxygenase (or the toluate dioxygenase, see above) has indeed been channelled through the ortho route. This is because CatR, the activator of the *cat* cluster, is turned on by *cis-cis* muconate, the product of the ortho-ring cleavage (Rothmel et al., 2002).
Fig. 8. Expression of ortho pathway genes

The metabolic fate of endogenously produced benzoate

The next question was the status of the ben/cat/pca ortho pathway transcriptome when benzoate is generated endogenously out of toluene metabolism by the upper TOL pathway enzymes (not supplied from outside as before). The results of such a situation are shown in Fig. 8C. Unlike exogenous addition of benzoate (Fig. 8A), we could detect no induction of benKAxF, catCBA and pca genes. Still, there was a clear signal in the tiling array spanning the benABCD sequences. However, unlike the results with benzoate in the same region (Fig. 8A), the level and continuity of the signals through that region were not even, thereby raising the suspicion of artefactual cross-hybridization with the highly similar xylXYZL genes of the lower TOL pathway (Harayama and Rekik, 1993), which become expressed at high levels in the presence of toluene (Fig. 4D). To clarify this uncertainty, we constructed plasmids bearing transcriptional (PbenA→′lacZ) and translational (PbenA→benA′′lacZ) reporter fusions, passed them to P. putida mt-2 and measured accumulation of β-galactosidase in the same conditions adopted for RNA extraction. The data shown in Fig. S5 clearly indicated that endogenous benzoate altogether fails to bring about activity of the PbenA promoter. Therefore, the tiling array signals coming from benABCD in Fig. 8C can be traced to the aforementioned similarity between benABCD and xylXYZL sequences. The same holds true for the signals in the same region caused by m-xylene (Fig. 8D) and o-xylene (Fig. 8E) as none of these inducers had significant effect on the activity of the PbenA′′lacZ and PbenA′′lacZ fusions (Fig. S5). Finally, 3MBz (Fig. 8F) is known not to be an effector of BenR and also that the complex XylS/3MBz is unable to activate PbenK (Perez-Pantoja et al., 2014). This reiterates that there is little or no activity of benABCD pathway and the other ortho pathway clusters in conditions other than external addition of benzoate.

These results tell us that while benzoate is indeed produced out of toluene in P. putida mt-2, it is instantly directed to the pWW0-encoded meta pathway and not to the chromosomal ortho route. This is reflected in the lack of induction of ben/cat/pca genes with toluene, as if the chromosomal pathway (unlike the TOL system) could not
see the intracellular benzoate. This may be due to: (i) differences in kinetic and thermodynamic parameters of the enzymes involved that prevent benzoate to accumulate to levels sufficient to activate BenR/Pcat/PcatA, (ii) an active inhibition of ben/cat/pca expression by unknown TOL-encoded functions or (iii) a physical channelling (Castellana et al., 2014; de Lorenzo et al., 2014) of the benzoate coming from the upper TOL pathway into the lower enzymatic complex. While these mechanistic questions are beyond the scope of this work, they expose that *P. putida* mt-2 had to evolve some sort of biochemical and/or genetic patch to solve the problem created by the coexistence of two potentially conflicting pathways (Jimenez et al., 2014).

### Conclusion

The tiling array described in this work for examining the TOL sub-transcriptome has been invaluable to clarify a number of standing questions that could not be addressed before with one-at-a-time approaches that have largely dominated the study of this archetypical metabolic and regulatory network. One key technical choice to this end has been the adoption of the C-runoff samples as the reference condition (Guelle et al., 2009) for establishing a baseline that allowed us to faithfully compare all other cases. This stratagem exposed: (i) a significant level of basal expression of the lower TOL pathway in the absence of inducers (in contrast with virtually no expression of the upper route), (ii) the presence of a distinct transcript downstream of the *xylN* gene and (iii) the overlapping transcription between the 3’ ends of the convergent *xylS* and *xylH* mRNAs, a somewhat frequent occurrence in genomes (Hovik et al., 2012). Yet, the most surprising finding was that benzoate generated from toluene catabolism by the upper pathway failed to induce expression of the ortho-cleavage genes. These results are compatible with a hypothetical scenario in which the enzymes of the upper pathway could be localized close to the enzymes of the lower pathway. In this way, a metabolically produced benzoate could be directly channelled towards the *meta* route. When benzoate is added directly to culture medium, such a channelling effect may disappear and both the *meta* and *ortho* pathway genes can be simultaneously expressed. This hypothetical setting – or any of its alternatives mentioned above, puts some focus on the retroactivity between a pre-existing metabolic chassis and the acquisition of a new biochemical module (Del Vecchio, 2015) whether by naturally occurring horizontal gene transfer or by genetic engineering. Specifically, the data above suggest a range of enzymatic, regulatory (and perhaps physical) devices that work as the *interlocks* of engineered items, i.e. they help prevent the system from damaging itself by interrupting its operation when it meets trouble. Enzymatic misrouting between coexisting pathways is one of the most common of such biochemical conflicts. In previous reports (Jimenez et al., 2014; Perez-Pantoja et al., 2014), we have identified different metabolic and regulatory devices for avoiding such clashes at various checkpoints of the aromatic biodegradation network of *P. putida* mt-2, and many more are likely to exist (Danchin, 2009). Understanding the mutual adaptation between chassis (e.g. environmental bacteria) and implants (e.g. metabolic pathways found in transmissible plasmids) will be of great value for engineering microbial production of biofuels and added-value chemicals (Kung et al., 2012; Nikel and de Lorenzo, 2014; Nikel et al., 2014).

### Experimental procedures

#### Strains and growth condition

Unless otherwise indicated, *Escherichia coli* and *P. putida* were routinely grown at 37 and 30°C, respectively, in Luria–Bertani or M9 minimal medium including 6 g l⁻¹ NaHPO₄, 3 g l⁻¹ KH₂PO₄, 1.4 g l⁻¹ (NH₄)₂SO₄, 0.5 g l⁻¹ NaCl, 0.2 g l⁻¹ MgSO₄·7H₂O and 2.5 ml l⁻¹ of a trace element solution as noted (Abril et al., 1989; Nikel and de Lorenzo, 2013), with a carbon source as described below. For induction of TOL catabolic genes during transcripctomic experiments, the *P. putida* mt-2 strain was precultured overnight in M9 medium with succinate, and bacterial cultures were then diluted 100-fold in the same medium and grown to the exponential phase (OD₆₀₀ = 0.3–0.5). Samples were then either cultured further without additional substrate or exposed to vaporeous m-xylene, toluene or o-xylene (one-half dilution in dibutyl phthalate, which is a non-effector for TOL genes) in a flask (2 h). For soluble substrates, benzoate (5 mM) or 3MBz (5 mM) was added to the culture medium once growth had reached OD₆₀₀ ∼0.3. For the C-runout condition, the cells were cultured for 96 h in succinate-supplemented minimal medium. The antibiotics kanamycin (Km) 50 µg ml⁻¹, ampicillin (Amp) 150 µg ml⁻¹, streptomycin (Sm) 100 µg ml⁻¹, chloramphenicol (Cm) 30 µg ml⁻¹ and gentamycin (Gm) 10 µg ml⁻¹ were added to bacterial cultures when necessary.

#### RNA preparation

Cell cultures were transferred to 1/10 sample volume of ice-cold ethanol/phenol solution (5% water-saturated phenol in ethanol) to protect RNA from degradation and harvested by centrifugation (3800 r.p.m., 15 min, 4°C). After supernatant aspiration, pellets were frozen in liquid nitrogen and stored at −80°C until required. Total RNA was extracted by using the miRNeasy kit (Qiagen) with some modifications. The collected pellets were resuspended into 0.3 ml Tris—HCl (pH 7.5) containing 2 mg ml⁻¹ lysozyme and incubated (10 min, 37°C). Lysate (0.1 ml) was used according to manufacturer’s instructions. RNase-free DNase (Qiagen) treatment was performed during the isolation procedure to eliminate residual DNA and quality was evaluated on a 2100 Bioanalyzer System (Agilent).
High-density tiling array design and analysis

A custom tiling microarray was designed to contain probes corresponding to the complete xyl operon of the TOL plasmid, as well as chromosomal genes encoding both branches of β-ketoadipate pathway and peripheral routes such as fcs cluster, pca genes, ben genes and cat genes. Oligonucleotide probes were 60 nucleotides long and were designed to be spaced every 10 nt and overlapping 50 nt. Both sense and antisense probes were designed to integrate the same regions and to identify natural antisense RNAs. Microarrays were synthesized by Agilent in the 8 × 15K format using the online tool eArray (https://earray.chem.agilent.com/earray/). Four biological replicates were hybridized independently for each transcriptomic comparison. Total RNA (20 μg each) was retrotranscribed and aminoallyl labelled using the SuperScript Indirect cDNA Labeling System (Invitrogen) and was then retrotranscribed and aminoallyl labelled using the 5′-(3-aminoallyl)-2′-deoxyuridine-5′-triphosphate (aa-dUTP, Ambion) following manufacturer’s protocols. To avoid antisense artefacts from second-strand cDNA during reverse transcription, actinomycin D was added after the denaturing step at 70°C to a final concentration of 6 nmol/l. To avoid antisense artefacts from second-strand cDNA during reverse transcription, actinomycin D was added after the denaturing step at 70°C to a final concentration of 6 nmol/l, as described (Perocchi et al., 2007). For each sample, aminoallyl-labelled cDNA was resuspended in 0.1 M Na₂CO₃ (pH 9.0) and conjugated with Cy3 or Hyper 5 Mono NHS Ester (CyDye™ Post-labelling Reactive Dye Pack, Amersham), following a dye-swap strategy. Samples were purified with MEGAcear (Ambion) following the manufacturer’s instructions. Cy3 and Hyper 5 incorporation was measured in a Nanodrop spectrophotometer (Nanodrop Technologies). Probe preparation and hybridizations were performed as described (Two-Color Microarray-Based Prokaryote Analysis, Agilent Technologies). Samples were placed on ice and rapidly loaded onto arrays, hybridized (65°C, 17 h), then washed once in GE Wash Buffer 1 (room temperature, 1 min) and once in GE Wash Buffer 2 (37°C, 1 min). Arrays were drained by centrifugation (2000 r.p.m., 2 min).

Images from Cy3 and Hyper 5 channels were equilibrated and captured with a GenePix 4000B Microarray Scanner (Axon) and spots were quantified using GenePix software (Axon) for tiling array data sets (m-xylene versus succinate). For other tiling array data sets that are reference condition-based comparisons, we used the Agilent DNA Microarray Scanner and quantified images using Agilent Feature Extraction software. After scanning and quantification, data from both data sets were analysed using the same methods. Raw intensities were background corrected by the normexp method with an offset of 50 (Smyth and Speed, 2003). Background-corrected intensities were converted to log₂ scale and normalized by adjusting the quantiles of all replicates as described (Bolstad et al., 2003). After normalization, differential expression for each probe was calculated as log₂Ratio = log₂Intensity (experimental condition) – log₂Intensity (control condition). Differential expression at the gene level was calculated by averaging all log₂Ratios of the probes that overlap with the gene coordinates. All statistical calculations were carried out with R software (The Development Core Team, 2011) and the Bioconductor package limma (Smyth, 2004). Integrated Genome Browser (IGB) (Nicol et al., 2009) was used to represent the log₂Ratios of all probes in the genomic regions. The Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo) accession number for both raw and processed data of the tiling and RNA-seq is GSE71853.

cDNA library construction for RNA deep sequencing

Total RNAs from m-xylene-treated cells or reference conditions were prepared. rRNA was depleted using the RiboZero rRNA removal kit (Epicerent Biotechnologies) for Gram-negative bacteria with a total RNA input of 4 μg. To confirm rRNA removal, we analysed the samples with the bioanayser system as described above. RNA libraries were prepared using the TruSeq RNA sample prep kit (Illumina) following the protocol described in the TruSeq RNA Sample Preparation Guide, starting at the fragmentation step. The library was validated by running a High Sensitivity DNA chip (Agilent Technologies) on the bioanayser system and quantified by qPCR according to the Illumina Sequencing Library qPCR Quantification Guide using KAPA SYBR FAST qPCR kit for LC480 (KAPA Biosystems). The libraries were denatured and adjusted to 14 P.M prior to generating the clusters as indicated in the Illumina Cluster Generation Guide for a Cluster Station instrument. Clusters were generated in one lane of the flow cell, which was sequenced in a single read 1 × 75 bp run by the Genome Analyzer Ix system.

Sequence read alignment and normalization for RNA deep sequencing

Short reads were aligned to the P. putida genome and pWW0 plasmid genes with the BWA algorithm with default settings (Li and Durbin, 2009). Mapped counts were normalized by TMM (Robinson and Oshlack, 2010) using the sample whose 75%-ile (of library-scale-scaled counts) is closest to the mean of 75%-iles as the reference. This is the default normalization method when applying the edgeR method (Robinson and Smyth, 2008) for analysing gene expression differences by RNA-seq. After normalization, differential expression was estimated by the method implemented in edgeR software, based on the negative binomial distribution of the reads. For estimating data dispersion, edgeR uses the quantile-adjusted conditional maximum likelihood (qCML) method. In this case, a common dispersion was considered for all samples. The alignment was visualized using the IGV (Thorvaldsdottir et al., 2012).

Preparation of sense strand and antisense strand RNA probes

The xylS intergenic region (260 bp) and the non-coding region (170 bp) between xylS and xylH were amplified by PCR using DNA template from P. putida mt-2 with the primer pairs xyl/Str-F (CCGCGATGCACAACTTCGCCATCGATCG) /xyl/Str-R (CGCGAGCTTC ATCATGCCAGCTCCGTAA) and xyl/Shr-F (CCGCGATGCACAACTTCGCCATCG) /xyl/Shr-R (CCGAGGCTCTGAGTGGATCGGACCGGTC). The PCR products were cloned into the pGEM-T-easy vector to generating pGEM-T-xyl/Sr and pGEM-T-xyl/Str respectively. To prepare sense RNA
probes, the recombination vectors were SacI digested. After the enzyme reaction, Klenow (NEB) was used to avoid non-specific transcription. For in vitro transcription, we used T7 RNA polymerase with 32P-radiolabelled CTP (Promega, Riboblock combination system). To synthesize antisense RNA probes, SphI enzyme and SP6 RNA polymerase were used with the same procedure as sense RNA probe synthesis. Labelled probes were purified using NICK columns (GE Healthcare).

Northern blot analysis

RNA samples (5 μg) were fractionated on 1% agarose-formaldehyde gels and transferred to membranes (Zeta-Probe, Bio-Rad) by capillary methods (Chomczynski, 1992), and fixed by ultraviolet cross-linking. Membranes were hybridized with 32P-CTP radiolabelled strand-specific sense- and antisense RNA probes for the xylS region and the non-coding region between xylS and xylH in ULTRAhyb ultrase sensitive hybridization buffer (Ambion) (68°C, overnight). Membranes were washed twice with 2× SSC (0.1% SDS, 68°C, 15 min) and twice with 0.1× SSC (0.1% SDS) (68°C, 5 min), then exposed to X-ray film.

Cloning procedures and construction of reporter fusions

For both sense and artificial antisense transcriptions of the xylS, the 1458 bp fragment from the xylS region covering its own promoter (Pα) and its extended 3′ end was amplified using primers xylS-F (5′CGGCTAGGCTAAAAGAAGCTCTTGGTT3′) and xylS-R (5′CGGCAAGTCCGATGCAGCCGCGCTGACCACTGCT3′). The amplicon was cloned into the XbaI/SalI cloning sites of the pSEVA224 vector (Silva-Rocha et al., 2013) including the LacIq Pm cargo to transcribe xylS in the antisense direction, to generate pSEVA224-XS. The plasmid construct was then introduced into E. coli CC118 by transformation. The pSEVA224-XS plasmid was then introduced by triparental conjugation into P. putida ME33D harbouring the Pm promoter fused to the gfp-lacZ reporter gene on chromosome (Silva-Rocha and de Lorenzo, 2012), thus resulting in P. putida ME3G3 (pSEVA224-XS). To confirm the integration of signals through Ccr and PtsN in catabolite repression of PbenA to lacZ, reporter strains were cultured in M9-succinate medium until the exponential phase, inducers were added where appropriate (2 h), and cells were prepared to measure β-galactosidase activity. Activity was measured with the Galactone-Light Plus system (Applied Biosystems), which is based on a chemiluminescent substrate of the enzyme (Galacto-Light Plus, Tropix). A sample of cells (500 μL; OD600 = 0.3–0.6) was centrifuged (13 000 r.p.m., 1 min, room temperature) and the pellet resuspended in 200 μL of lysis buffer (100 mM potassium phosphate, pH 7.8, 0.2% Triton X-100). The lysis mixture was subjected to two freeze-thaw cycles and centrifuged to eliminate cell debris. Lysed supernatant (20 μL) was incubated (30 min) with 80 μL of reaction buffer (100 mM sodium phosphate, pH 8.0, 1 mM MgCl2, 1X Galacto-Plus). These reactions were performed in 96-well plates and recorded in a luminometer for 30 s immediately after addition of 125 μL of a light emission accelerator (Accelerator-II Sapphire-II) following manufacturer’s instructions.

Acknowledgements

The authors are grateful to J.M. Franco (Genomics Unit, CNB) for his help with the design of the tiling array. This work was supported by the CAMBIOS Project of the Spanish Ministry of Economy and Competitiveness, the ST-FLOW, EVOPROG, ARISYS and EMPOWERPUTIDA contracts of the EU and the PROMPT Project of the Autonomous Community of Madrid. The authors declare no conflict of interest.

References


© 2015 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*


Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

**Fig. S1.** Blow-up of tiling array signals through the 3′ end of the upper TOL operon.

**Fig. S2.** Correction of sequence annotation on the *xylL* gene in the pWW0 plasmid of *P. putida* mt-2.

**Fig. S3.** Average intensity of the probes for the upper and lower operons, and the *xylS* and *xylR* genes in response to different TOL inducers.

**Fig. S4.** Average intensity of the probes for phenylpropenoid and *p*-hydroxybenzoate catabolic genes in response to different TOL inducers.

**Fig. S5.** *PbenA* promoter activity in *P. putida* mt-2 in response to TOL inducers.