***High-quality genome-scale metabolic modeling of Pseudomonas putida highlights its broad metabolic capabilities***

**Juan Nogales,a,b#\*, Joshua Muellerb,c\*, Steinn Gudmundssond, Francisco J. Canalejoa, Estrella Duquee, Jonathan Monkb, Adam M. Feistb, Juan Luis Ramose,Wei Niuc, and Bernhard O. Palssonb**

aDepartment of Systems and Synthetic Biology, Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

bDepartment of Bioengineering, University of California, San Diego, La Jolla, CA, USA

cDepartment of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, USA

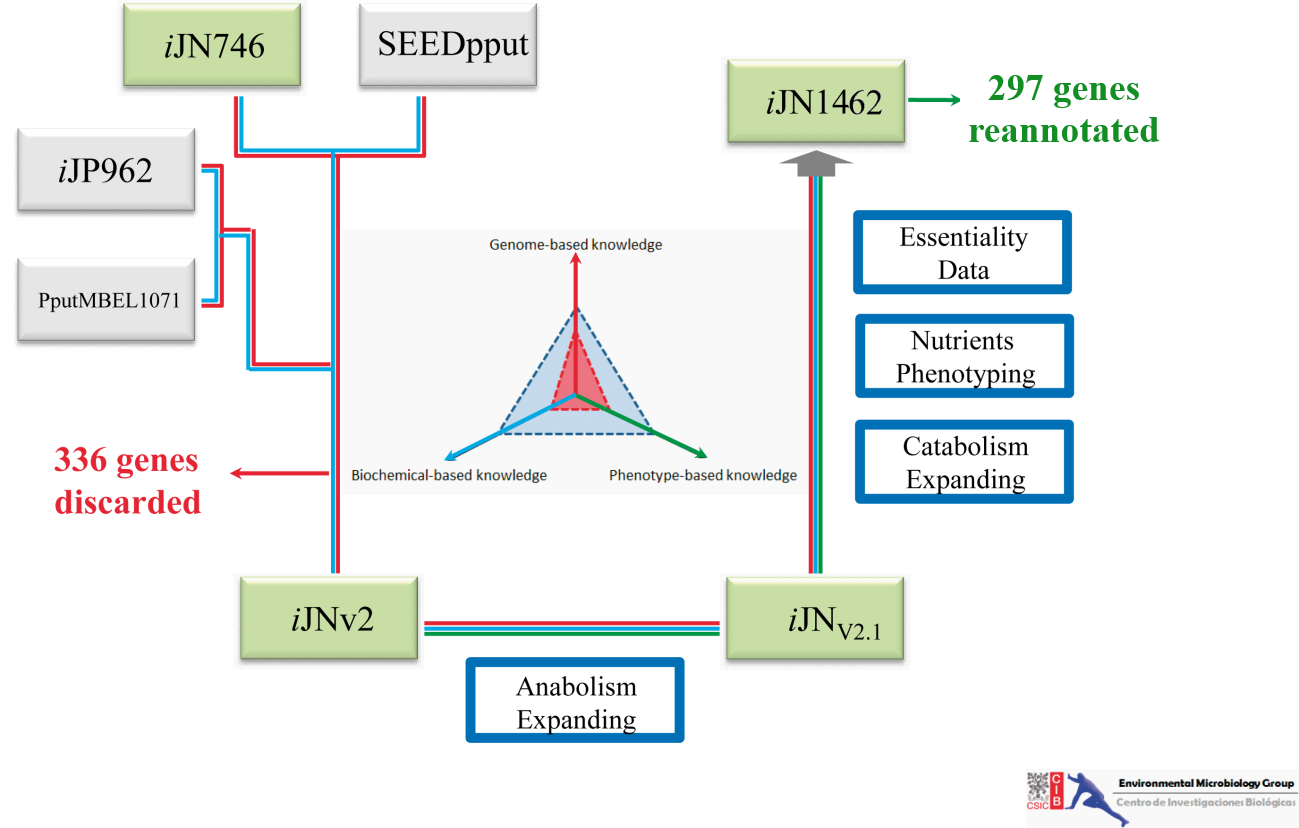
dCenter for Systems Biology, University of Iceland, Iceland

eDepartment of Environmental Protection, Estación Experimental del Zaidín (CSIC), Granada, Spain

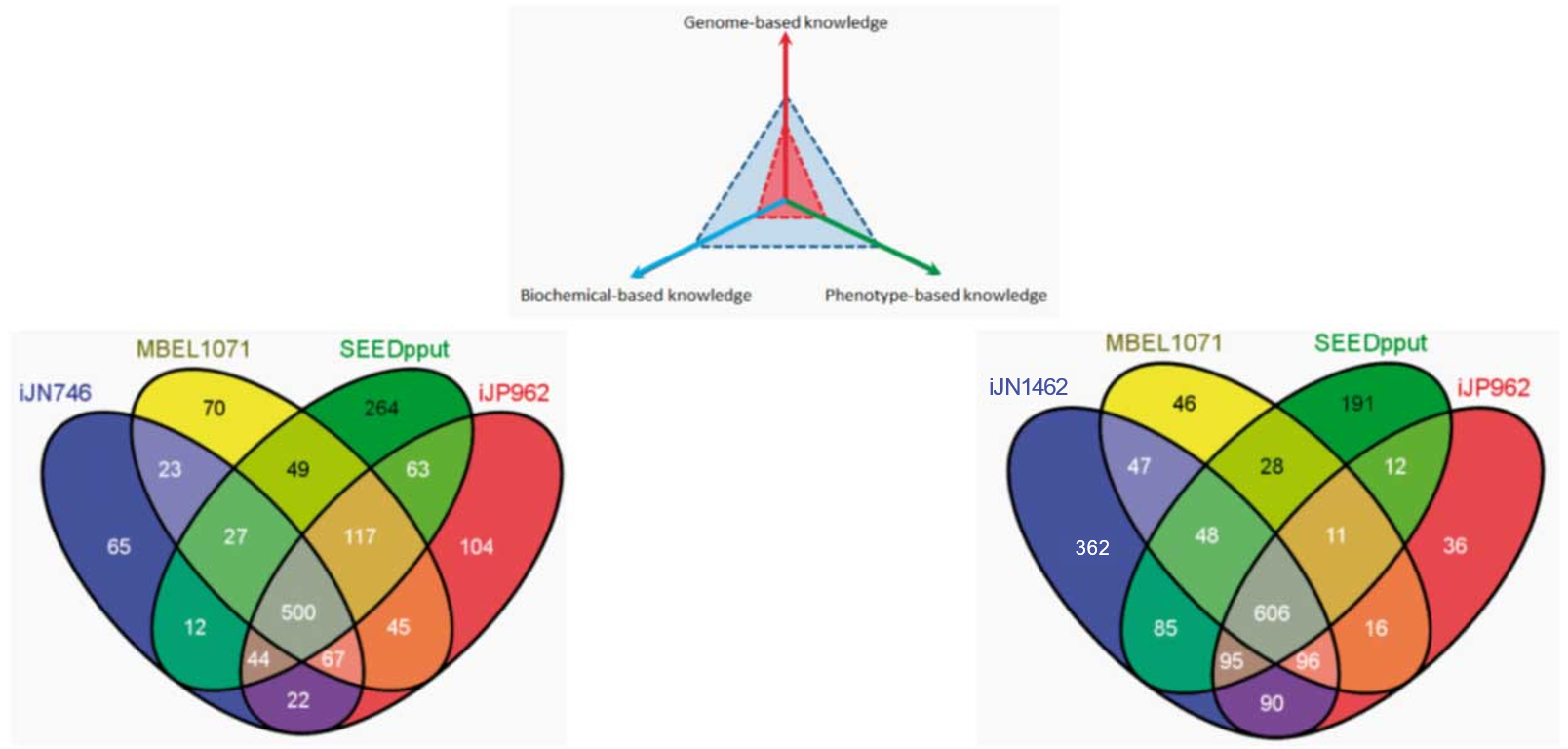
# Address correspondence to: Juan Nogales, [jnogales@cnb.csic.es](mailto:jnogales@cnb.csic.es)

\*Contributed equally

**1. Reconstruction process**



**Figure. S1. Overall workflow followed for reconstructing the metabolic network of *P. putida* KT2440.** The reconstruction was performed through an iterative tri-dimensional expansion based on genome annotation (red lane), biochemical (blue line) and phenotypic (green lane) information. Along this approach up to 336 genes included in previous reconstruction were discarded due insufficient information available while 246 ORFs from the genome of *P. putida* were re-annotated (Table S1).



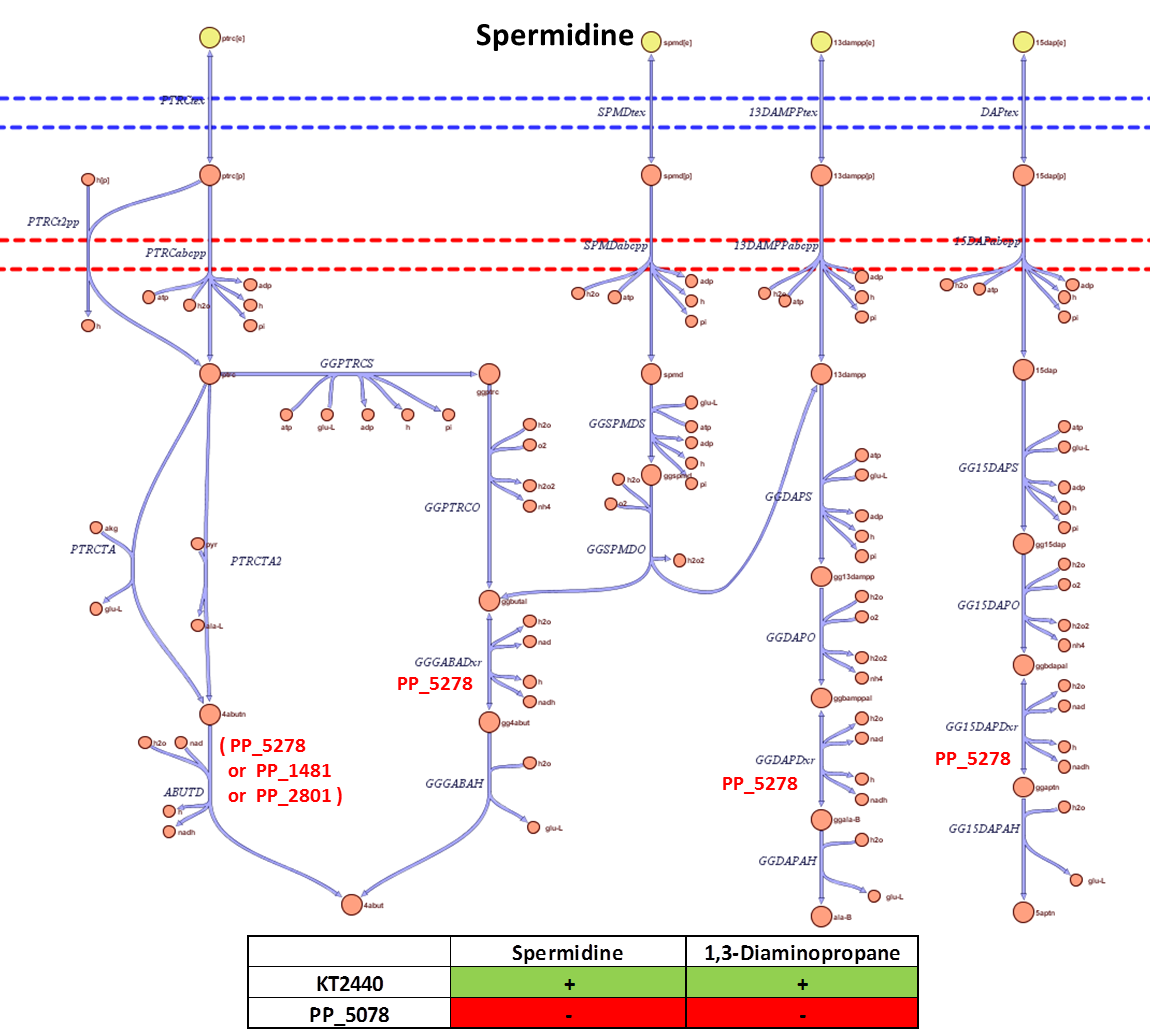
**Figure. S2. Venn diagrams showing the gene comparison between the current *P. putida* reconstructions.** Our previous reconstruction *i*JN746 (left) and *i*JN1462 (right) are shown as reference.

**2. Reconstruction Validation.**

**Figure S3. Reconstruction of the sarcosine catabolom.** Based on legacy data and manual search in the genome of *P. putida* KT2440 it was possible to reconstruct the catabolic pathway for choline sulfate, choline, glycine betaine, dimethylglycine, creatine, carnitine, carnosine and creatinine which converge in the key metabolite sarcosine. The reconstruction was validated by growth experiments (Table S2). Furthermore, knockouts available at PRCC were used to validate at a molecular level the reconstruction. Hence, while the wild type strain grew on glycine betaine and creatine, *P. putida* strains knockouts in the genes PP\_0310 and PP\_0116 were unable to growth in glycine betaine and a knockout strain in the gene PP\_3637 was unable to growth using creatine as only carbon source. The abbreviations for metabolites and reactions can be found in Table S1.

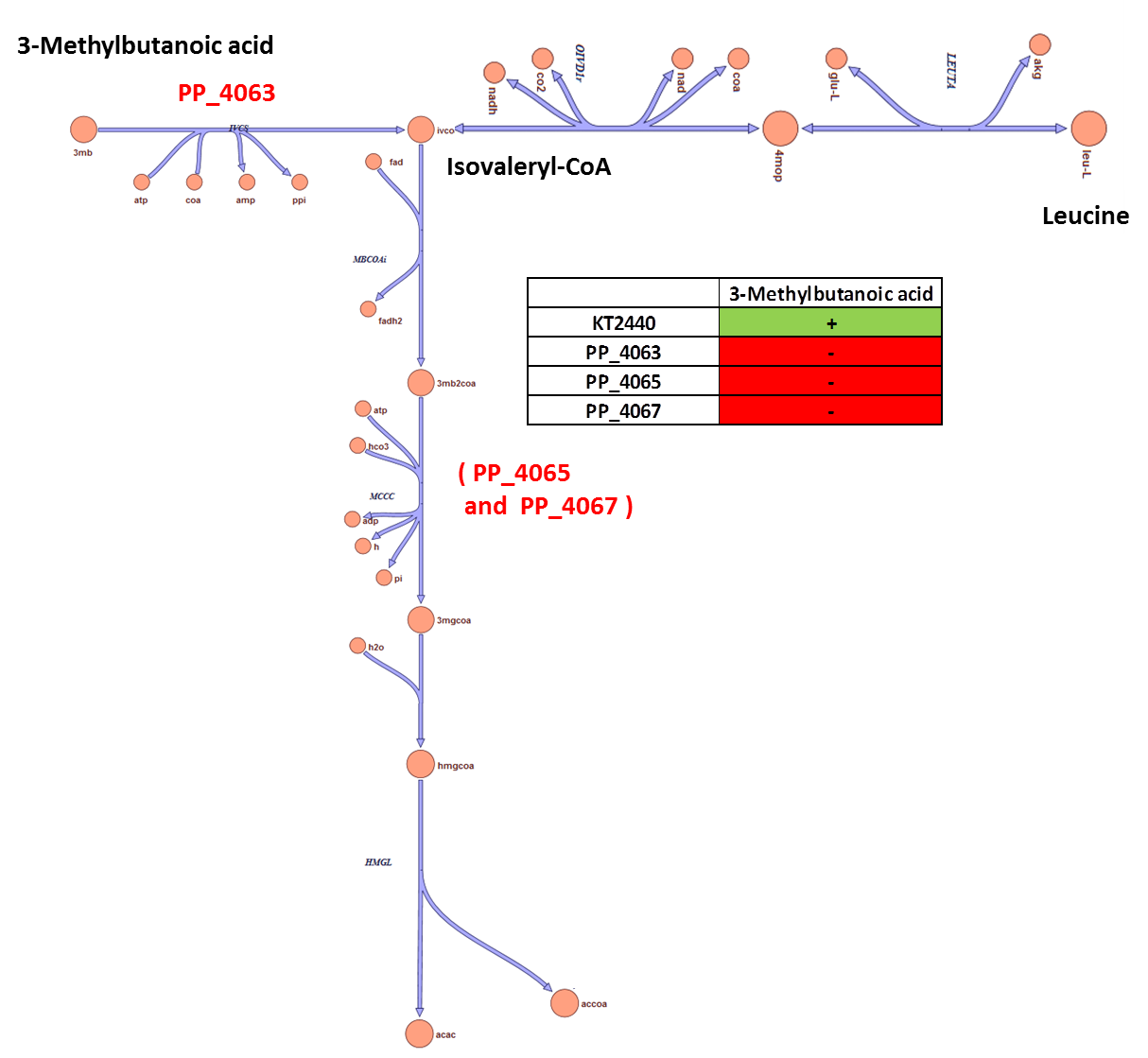


**Figure S4. Reconstruction of the 2,5 dioxopentanoate catabolom.** Based on legacy data and manual search in the genome of *P. putida* KT2440 it was possible to reconstruct the catabolic pathway for D-galacturate, D-galactarate, D-glucuronate, D-glucarate, trans-4-Hydroxy-L-proline and cis-4-Hydroxy-L-proline which converge in the key metabolite 2,5 dioxopentanoate. The reconstruction was validated by growth experiments (Table S2). Furthermore, knockouts available at PRCC were used to validate at molecular level the reconstruction. Hence, *P. putida* strains knockouts in the genes PP\_1168 and PP\_1169 were unable to growth D-Galacturonate and D-glucoronate, a knockout strain in the gene PP\_4757 failed in to growth in Glucarate. Finally, a knockout strain in the gen PP\_1259 was unable to growth using cis-4-Hydroxy-L-proline as only carbon source. The abbreviations for metabolites and reactions can be found in Table S1.



**Figure S5. Reconstruction of polyamine metabolism.** Based on legacy data and manual search in the genome of *P. putida* KT2440 it was possible to reconstruct the catabolic pathway for polyamines. The reconstruction was validated by growth experiments (Table S2). Furthermore, knockouts available at PRCC were used to validate at molecular level the reconstruction. Hence, *P. putida* strain knockout in the key gene PP\_5278 was unable to growth using Spermidine or 1,3-Diaminopropane as carbon sources. The abbreviations for metabolites and reactions can be found in Table S1.

The reviewer is right when stat that the multi-strain modeling has been done by removing metabolic content from iJN1462. Therefore, potential strain-specific content is absent in the cognate model.



**Figure S6. Reconstruction of isovaleryl-Coa metabolism.** Based on legacy data and manual search in the genome of *P. putida* KT2440 it was possible to reconstruct the catabolic pathway for isovaleryl-CoA. The reconstruction was validated by growth experiments (Table S2). Furthermore, knockouts available at PRCC were used to validate at molecular level the reconstruction. Hence, *P. putida* strain knockouts in the genes PP\_4063, PP\_4065 and PP\_4067 were unable to growth using 3-Methylbutanoic acid as carbon sources. The abbreviations for metabolites and reactions can be found in Table S1.



**Figure. S7.** **Relative production of PHA monomers by *i*JN1462.** The 223 different carbon sources supporting growth of *i*JN1462 (vertical axis) were used to maximize the production of the 24 different PHA monomers included in *i*JN1462 (horizontal axis). The relative production was computed as the fraction of carbon being transformed in the target monomer. The analysis shows that *P. putida* is well suited for producing aliphatic PHA monomers, irrespective of the carbon source, while the aromatic, thiol, and polyunsaturated PHA monomers are highly carbon source dependent. R-Hydroxyacid PHA monomers: CX:Y where the X indicates the chain length and Y the number of unsaturations. The Phenyl and Thio prefixes indicate monomers including phenyl and thiol groups in their structures, respectively.

**3. Biomass Comparison**





**Figure S8. Qualitative comparison of Biomass components present in *i*JN1462, *i*EB1050 and PpuQY1140**

**4. Nutrients source Validation.**

We validated the growth performance accuracy of *i*JN1462 taken advantage of the legacy data. For those nutrients with contradictory reports or no prior testing in *P. putida*, we performed experimental validation (72 and 46 carbon and nitrogen sources, respectively). The carbon and nitrogen sources supporting growth in *i*JN1462 together with the validation sources is provided in Table S2. Excluding those metabolites for which the validation was not possible due technical issues, the overall accuracy of the updated model was very high and it predicted correctly the 79% and 84% (two-sided p-values of Fisher‘s exact test were less than 10-12) of the phenotypes observed for carbon and nitrogen sources, respectively (Figure 2). Just a few false negatives were found and being confined to metabolites which catabolic pathways are unknown in *Pseudomonas* e.g., 3,4-dihydroxyphenylacetate, methyl pyruvate, and Tween 20 as carbon source and L-cysteine as nitrogen source. However, significant higher number of false positives (nutrients supporting growth in silico but not in vivo were detected. Many of they supported growth as nitrogen source but not as carbon source e.g., nitrogenous bases such as xanthine and hypoxanthine and polyamines like agmatine, suggesting regulatory and/or adaptive limitations. In fact, KT2440 grew on xanthine when this metabolite was used as only carbon and nitrogen source (data non-shown), suggesting that the metabolism of xanthine is repressed in presence of ammonium. In addition, when P. putida was cultured in xanthine it grew after 96 h.

It has become evident because of insufficient adaptive and/or regulatory evolution, silent metabolic pathways without immediately apparent function are encoded in bacterial genomes as ([Fong, Nanchen et al. 2006](#_ENREF_5), [Conrad, Lewis et al. 2011](#_ENREF_2), [Cornelius, Lee et al. 2011](#_ENREF_3)). Since GENREs ignore these limitations, often the *in silico* metabolic capabilities predicted surpass the experimentally observed phenotypes. This fact explains, to some extent, the presence of false positives in growth predictions. Interestingly, adaptive laboratory experiments (ALE) have shown that in silico predicted phenotypes can be achieved in vivo through short-term evolution ([Fong, Nanchen et al. 2006](#_ENREF_5), [Conrad, Lewis et al. 2011](#_ENREF_2)). In order to investigate if some of the false positives found here respond to lack of evolutionary adaptation, we grew KT2440 on ethylene glycol since it has been described that several *P. putida* strains are able to use this compound as a sole carbon and energy source but not the KT2440 strain ([Mückschel, Simon et al. 2012](#_ENREF_13)). We observed that KT2440 was unable to grow on EG after 48 h of incubation as previously has been reported, however it grew efficiently after 100 hours. This result shows that the EG catabolic pathway is indeed a latent pathway in KT2440, and it suggests that the rest of the false positives detected could be suitable targets for the catabolic expansion of this biotechnologically interesting strain.

**5. Gene-essentiality analysis validation and contextualization.**

The essentiality analysis and validation experiments provided a unique opportunity to deciphering new insights in the metabolism of KT2440 while increasing the accuracy of *i*JN1462. For instance, several *cob* genes involved in the biosynthesis of vitamin B12 were predicted as essential in the first version of *i*JN1462 contrary to the phenotype displayed by these gene knockouts in vivo ([Molina-Henares, De La Torre et al. 2010](#_ENREF_10)). Even more, based on the lack of B12 requirement of these mutants together with the presence of several *cbi* genes (responsible for synthetizing B12 under anaerobic conditions) it was suggested a putative alternative B12 biosynthetic pathway may exist. Thus, uncertainties remain about the biosynthesis of this vitamin in KT2440 ([Molina-Henares, De La Torre et al. 2010](#_ENREF_10)). In order to resolve this gap of knowledge, we performed a comprehensive search looking for B12-dependent enzymes encoded in the genome of KT2440. We found three, two of which are included in *i*JN1462 e.g., the methionine synthase MetH (PP\_2375) and the ethanolamine ammonia Lyase EutCB (PP\_0542 and PP\_0543). The ethanolamine ammonia Lyase is easier to test because the methionine synthase encoding gene (MetE) is B12-independent and exists in KT2440 making MetH non-essential even in media lacking B12, thus we decided to validate the role of the *cob* genes in B12 biosynthesis by growing the *cob* knockouts in ethanolamine as the only nitrogen source. If such mutant strains were unable to growth under these conditions, the presence of an alternative pathway for synthesizing B12 would be discarded, while the role of cob genes in this process would be unequivocally demonstrated. We found that this was the case (Fig. S9), thus we proceeded to remove the B12 from the wild-type BOF solving the initial discrepancy between *in vivo* and *in silico* essentiality data.



**Figure S9. Role of *cob* genes in B12 biosynthesis.** *P. putida* KT2440 and isogenic knockouts strains lacking *cob* genes were grown in glucose minimal medium using ammonia or ethanolamine as nitrogen source. Growth or not growth are indicated as (+) or (-), respectively.

**6. Definition of *In silico* Luria Broth (LB) medium.**

In order to define an accurate in silico LB medium (iLB) we used the basic composition based on previous reports ([Oh, Palsson et al. 2007](#_ENREF_16)) and conditional essential gene analysis from KT2440 ([Molina-Henares, De La Torre et al. 2010](#_ENREF_10)).

The default Exchange reactions were constrained as follow.

**Carbohydrates**

About ~ 29% weight; glucose was assumed as a representative carbohydrate source.

model=changeRxnBounds(model,'EX\_glc(e)',-5,'l');

**Amino acids**

model=changeRxnBounds(model,'EX\_ala\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_asp\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_glu\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_his\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_leu\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_met\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_pro\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_thr\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_tyr\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_arg\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_cys\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_gly(e)',-5,'l');

model=changeRxnBounds(model,'EX\_ile\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_lys\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_phe\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_ser\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_trp\_L(e)',-5,'l');

model=changeRxnBounds(model,'EX\_val\_L(e)',-5,'l');

**Vitamins**

**Cobalamine (B12).** Conditional essential genes found in Molina-Henares et al, 2010: PMID:20158506.

model=changeRxnBounds(model,'EX\_cbl1(e)',-0.1,'l');

**Riboflavin (Vitamin B2).** No auxotrophs were found in Molina-Henares et al, 2010: PMID:20158506. No gene encoding riboflavin transport. Assumed that *P. putida* is unable to transport this vitamin.

**Biotin (B8**). Conditional essential genes found in Molina-Henares et al, 2010: PMID:20158506.

model=changeRxnBounds(model,'EX\_btn(e)',-0.1,'l');

**Niacin (PP).** Conditional essential genes found in Molina-Henares et al, 2010: PMID:20158506. NadA knockout required nicotinic acid for growth.

model=changeRxnBounds(model,'EX\_nac(e)',-0.1,'l');

**Thiamine (B1).** No auxotrophs were found in Molina-Henares et al, 2010: PMID:20158506. No gene encoding thiamine transport. Assumed that *P. putida* is unable to transport this vitamin.

**Pyridoxine (B6).** No auxotrophs were found in Molina-Henares et al, 2010: PMID:20158506. No gene encoding pyridoxine transport. Assumed that *P. putida* is unable to transport this vitamin.

**Pantotenate.** Synthesis of pnto-R is not essential in LB (This study)

model=changeRxnBounds(model,'EX\_pnto\_R(e)',-0.1,'l');

**Folic acid (B9).** No auxotrophs were found in Molina-Henares et al, 2010: PMID:20158506. No gene encoding folic acid transport. Assumed that *P. putida* is unable to transport this vitamin.

**Chorismate.** No auxotrophs found, P. putida is able to transport aromatic aminoacid as well. Assumed present in LB medium based on Oh et al, 2007 (PMID:17573341).

model=changeRxnBounds(model,'EX\_chor(e)',-0.1,'l');

**Diaminopimelic acid.** An *dapA* (PP\_1237) knockout required diaminopimelic acid for grow in minimal media with glucose as carbon source, since *dapA* is not essential LB medium, was assumed that 26dap\_M is present in LB.

model=changeRxnBounds(model,'EX\_26dap\_M(e)',-0.1,'l');

**Siroheme.** CysG, (PP\_3999) the key enzyme in sheme biosynthesis is not essential in LB medium. It was assumed that sheme is present in LB.

model=changeRxnBounds(model,'EX\_sheme(e)',-0.1,'l');

**Other metabolites included based on Oh et al, 2007.**

model=changeRxnBounds(model,'EX\_na1(e)',-100,'l');

model=changeRxnBounds(model,'EX\_cl(e)',-100,'l');

model=changeRxnBounds(model,'EX\_so4(e)',-100,'l');

model=changeRxnBounds(model,'EX\_k(e)',-100,'l');

model=changeRxnBounds(model,'EX\_pi(e)',-100,'l');

model=changeRxnBounds(model,'EX\_ca2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_mg2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_sel(e)',-100,'l');

model=changeRxnBounds(model,'EX\_zn2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_aso4(e)',-100,'l');

model=changeRxnBounds(model,'EX\_cd2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_hg2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_pb2(e)',-100,'l');

**Other chemicals (included based on *P.putida* BOF)**

model=changeRxnBounds(model,'EX\_ni2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_cu2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_fe2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_fe3(e)',-100,'l');

model=changeRxnBounds(model,'EX\_mn2(e)',-100,'l');

model=changeRxnBounds(model,'EX\_mobd(e)',-100,'l');

model=changeRxnBounds(model,'EX\_cobalt2(e)',-100,'l');

**Nucleotides/nucleosides**

model=changeRxnBounds(model,'EX\_ins(e)',-5,'l');

model=changeRxnBounds(model,'EX\_hxan(e)',-5,'l');

model=changeRxnBounds(model,'EX\_ura(e)',-5,'l');

model=changeRxnBounds(model,'EX\_uri(e)',-5,'l');

model=changeRxnBounds(model,'EX\_adn(e)',-5,'l');

**Additional constrains.**

model=changeRxnBounds(model,'EX\_o2(e)',-21.5,'l');

model=changeRxnBounds(model,'EX\_nh4(e)',0,'l');

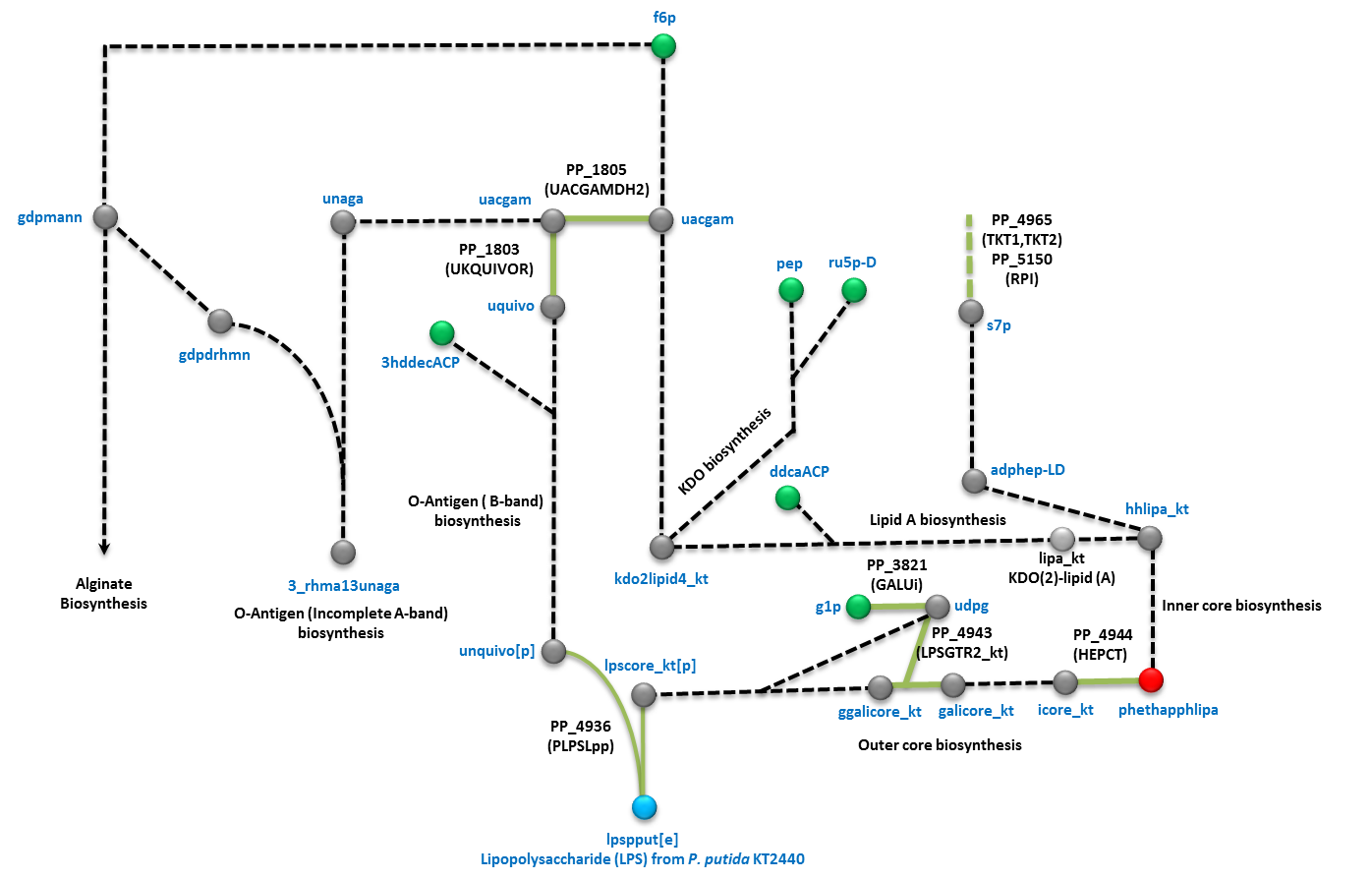
**7. Definition of in silico M9 glucose minimal medium.**

In order to define an in silico M9 (iM9) glucose minimal medium, the lower and upper bounds of several exchange reactions were constrained as follows.

|  |  |  |  |
| --- | --- | --- | --- |
| Rxn name | Rxn description | Lower Bound  (mmol.gDW-1.h-1) | Upper Bound  (mmol.gDW-1.h-1) |
| EX\_co2(e) | CO2 exchange | -100 | 1000 |
| EX\_h2o(e) | H2O exchange | -100 | 1000 |
| EX\_h(e) | H exchange | -100 | 1000 |
| EX\_o2(e) | O2 exchange | -30 | 1000 |
| EX\_ca2(e) | Calcium exchange | -10 | 1000 |
| EX\_cl(e) | Chloride exchange | -10 | 1000 |
| EX\_cobalt2(e) | Co2 exchange | -10 | 1000 |
| EX\_cu2(e) | Cu2 exchange | -10 | 1000 |
| EX\_fe2(e) | Fe2 exchange | -10 | 1000 |
| EX\_hco3(e) | Bicarbonate exchange | -10 | 1000 |
| EX\_k(e) | K exchange | -10 | 1000 |
| EX\_mg2(e) | Mg exchange | -10 | 1000 |
| EX\_mn2(e) | Mn2 exchange | -10 | 1000 |
| EX\_mobd(e) | Molybdate exchange | -10 | 1000 |
| EX\_na1(e) | Sodium exchange | -10 | 1000 |
| EX\_nh4(e) | Ammonia exchange | -10 | 1000 |
| EX\_ni2(e) | Ni2 exchange | -10 | 1000 |
| EX\_pi(e) | Phosphate exchange | -10 | 1000 |
| EX\_sel(e) | Selenate exchange | -10 | 1000 |
| EX\_so4(e) | Sulfate exchange | -10 | 1000 |
| EX\_tungs(e) | tungstate exchange | -10 | 1000 |
| EX\_zn2(e) | Zinc exchange | -10 | 1000 |
| EX\_glc(e) | D Glucose exchange | -6,3 | 1000 |

**8. Definition of *a CORE biomass objective function.***

We proceeded to formulate the core biomass objective function based on the minimal biosynthetic precursors supporting growth according to the knockouts present at PRCC ([Duque, Molina-Henares et al. 2007](#_ENREF_4)). For instance, we included in the core biomass function the metabolite phosphor-6-heptosyl-1,3-ethanolaminephosphate-2-phospho-4-heptosyl-1,5-kdo2\_lipidA (phethapphlipa) instead of the complete lipopolysaccharide (lpspput) since knockouts strains upstream of this metabolite in the biosynthetic pathway were not present at PRCC. Thereby we assume that phethapphlipa is the minimal lipopolysaccharide supporting growth in *P. putida.*



**Figure S10. Reconstruction of the lipopolysaccharide biosynthetic pathway in *P. putida* KT2440.** Precursor and intermediate metabolites in the biosynthesis of the lipopolysaccharide of KT2440 are shown as green and grey balls, respectively. The complete lipopolysaccharide included in the wild type biomass is shown in blue while the metabolite phethapphlipa present in the core biomass is shown in red. Dashed lines indicate lumped reactions in the biosynthetic pathway. *P. putida* gene knockouts strains present at PRCC are indicated in green lines. Note that no knockouts strains upstream of phethapphlipa are present at PRCC. The abbreviations for metabolites and reactions can be found in Table S1.

**9. BarSeq Sensitivity Analysis**

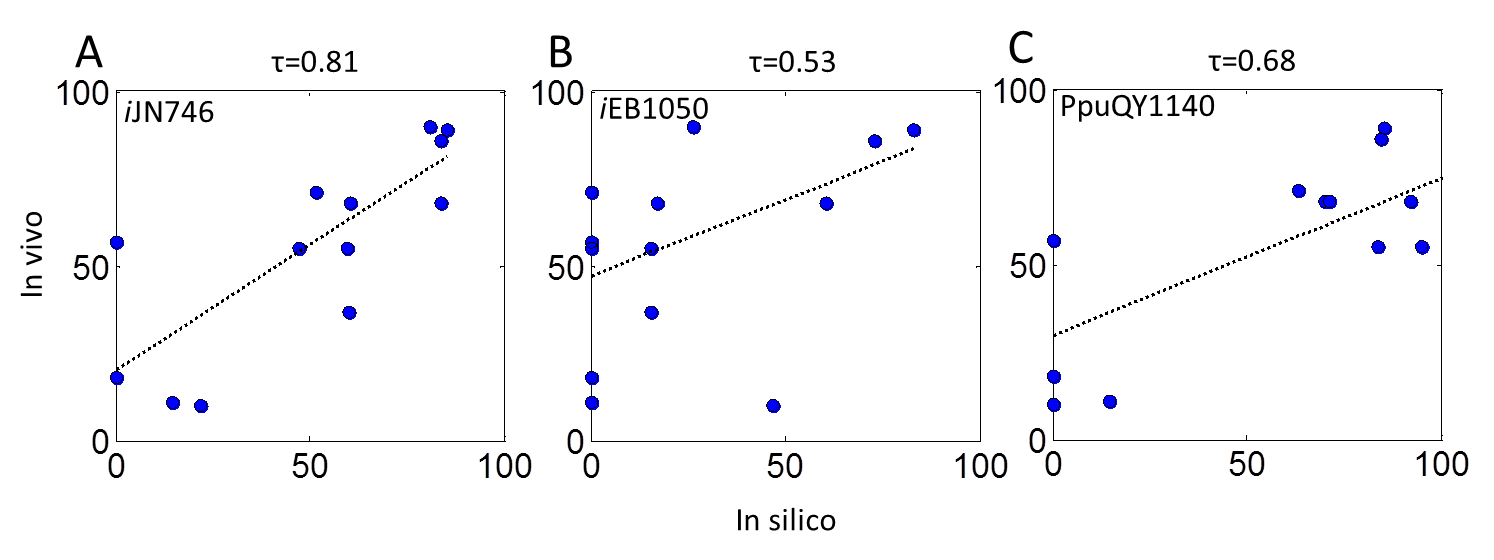
BarSeq experiments generate fitness values that express how well a deletion strain grew compared to all other deletion strains. The fitness ratio is calculating by generating a bar-coded library of all possible knock-outs. Samples are taken at the beginning and the end of the experiment and fitness ratios are calculated based off of how well a given knockout was able to replicate in comparison to the entire library. Positive numbers indicate beneficial knockouts while negative numbers indicate detrimental knockouts. Because of this methodology it is difficult to exactly identify essential genes, barring universally essential genes that don’t appear in the BarSeq library due to the inability to culture them in any media. Fitness values are generally just looked at relative to one another to try and identify genes that are more or less beneficial in certain growth conditions.

In order to compare the BarSeq data with model predictions, we varied a threshold for BarSeq fitness values. Genes with values below the threshold were considered essential for the growth condition. For model predictions, the *single\_gene\_deletion* function from Cobrapy was used to identify the essential genes for each growth condition. Results between the model predictions and BarSeq experiments were compared and the BarSeq threshold that resulted in the highest Matthews correlation coefficient was used. This resulted in a BarSeq fitness threshold of -2.7 for essential genes.

**10. Multi-strain *P. putida* modeling.**

For the modeling of the sequenced *P. putida* strains we followed a similar approach used previously by Monk and colleagues ([Monk, Charusanti et al. 2013](#_ENREF_12)). Strains with greater than 300 genes missing compared to iJN1462 were discarded. These were considered low quality or significantly diverged and possibly misclassified from KT2440. The list of orthologous genes between the different strains was determined using Bi-directional Best Blast hits (BBH) with a cutoff of 80% percentage identity across greater than 80% of the length of the reference gene in *P*. *putida* KT2440 (Table S5). After that, we removed from *i*JN1462 those genes missing and their reactions associated in each *P. putida* strain being reconstructed by using the function “*deleteModelGenes*” implemented in the Cobrapy. Finally, we tested the ability of the new models to produce biomass in *i*LB using the core biomass objective function as BOF. We noted that most of the models were unable to produce biomass under these conditions. Gapfilling was performed in order to identify reactions from the KT2440 model that were necessary for growth on glucose. These reactions were automatically added back into the strain-specific models.

**11. Carbon flux predictions.**

****

**Figure S12. Carbon flux predictions from previous models of *P. putida* KT2440.**

Comparisons between experimentally reported flux values in the central metabolism of *P. putida* growing on glucose ([Blank, Ionidis et al. 2008](#_ENREF_1)) and predicted flux values obtained with *i*JN746, *i*EB1050 and PpuQY1140.

**12. Material and Methods**

***Metabolic reconstruction process of P. putida KT2440***

The workflow of the reconstruction process started with the GEMs of *P. putida* available at that time (November 2011): *i*JN746 (36), *i*JP850 (37), PpuMBEL1071 (38), and *i*JP962 (39). As often happens for bacteria being reconstructed, the available models for *P. putida* KT2440 are significantly different and surprisingly they only share 523 genes from the 1213 unique genes included in the four models. This data highlights the large bias and lack of manual curation inherent in many metabolic reconstruction processes ([Monk, Nogales et al. 2014](#_ENREF_11), [Heavner and Price 2015](#_ENREF_6)). The overall workflow for the reconstruction process used here is shown in the supplemental material (SI1, Fig. S1). Briefly, the reconstruction was performed manually following an iterative tri-dimensional expansion based on genome annotation, biochemical, and phenotypic legacy knowledge. For genome-based expansion, the *P. putida* KT2440 model, Seed160488.1, (PputSEED) was downloaded from the Model SEED database ([Henry, DeJongh et al. 2010](#_ENREF_7)) and its content mapped on *i*JN746. The subsequent comparison shows that PputSEED included 497, 881 and 655 exclusive genes, reactions and metabolites, respectively (SI1, Fig. S2, Table S1). These sets of genes, reactions and metabolites absent in *i*JN746 were further manually investigated one by one in order to justify their inclusion in the updated model based on legacy knowledge and/or computational evidence. When appropriate, the new content was included in the reconstruction and the new reactions and metabolites were named following the BIGG nomenclature ([King, Lu et al. 2015](#_ENREF_9)). This approach significantly increased the content of *i*JN746, however we found several inconsistences in PputSEED in four categories; i) lack of specie-specific reaction formulations, including inappropriate substrate and/or cofactors and/or reversibility; ii) inaccurate GPR associations; iii) inaccurate compartmentalization of reactions; and iv) incorrect modeling of most of the biosynthetic pathways including the cell envelope, phospholipids. Multiple reactions were therefore excluded or reformulated based on *Pseudomonas* legacy knowledge.

In the second step, the content from two additional metabolic reconstructions of *P. putida*, *i*JP962 ([Oberhardt, Puchałka et al. 2011](#_ENREF_15)) and PpuMBEL1071 ([Sohn, Kim et al. 2010](#_ENREF_23)) were investigated following the above workflow and when appropriate, new content was added to the model. This second genome-based expansion step provided minor additions compared to PputSEED and similarly, a large number of genes were discarded and/or included in different GPRs. Following the recommendations transparence guidelines for metabolic reconstructions ([Heavner and Price 2015](#_ENREF_6)), the list of discarded genes from previous GENREs of *P. putida* (up to 336) and the reason of their exclusion are provided in Table S1.

The biochemical and phenotypic expansion was performed simultaneously by modeling new anabolic and catabolic pathways (SI1, Fig. S1). During this step, legacy knowledge from *Pseudomonas* found in databases such as the Pseudomonas Database, BRENDA etc., as well as in the primary literature was widely used. As a result, up to 410 unique citations are included in the final reconstruction and 2048 of the reactions have, at least, one citation supporting its inclusion in the reconstruction. The list of citations is provided in SI (Table S1). This detailed search for biological knowledge in *Pseudomonas* beyond the genome-annotation allowed the accurate modeling of multiple new biosynthetic and catabolic pathways, many of which were previously unknown in *P. putida* KT2440 and have been modeled here for the first time. Finally, the model built on the reconstruction was thoroughly evaluated in order to detect inconsistencies by experimental nutrient phenotyping and gene essentially data ([Duque, Molina-Henares et al. 2007](#_ENREF_4), [Molina-Henares, De La Torre et al. 2010](#_ENREF_10)) (SI, Table S2). This approach allowed the re-annotation and/or more accurate assignment of function to 297 genes encoded in the *P. putida* KT2440 genome. The complete list of re-annotated genes is provided in Table S1.

The SimPheny™ (Genomatica Inc., San Diego, CA) software platform was used to build the reconstruction. All the metabolites in the reconstruction were introduced according to their chemical formula and charge using their pKa value at pH 7.2. All reactions were subsequently mass and charge balanced. The reversibility for each reaction in the reconstruction was determined from the primary literature, when possible, or taken for phylogenetically related organisms. In addition, for each reaction included in the model a confidence score (CS) ranging from 1 to 4 was assigned (Table S1). A value of 1 indicates *in silico* evidence supporting the inclusion of a given reaction, e.g., the reaction is solely required for the functionality of the model. A value of 2 indicates genomic or physiological evidence. Reactions with a confidence score of 3 are supported by genetic evidence such as knockout characterization and a value of 4 indicates that the target GPR has been completely characterized. The average CS was 2.59 (Table S1). The model in SBML and json format is provided in SI3 and it will be made available in the BIGG database after publication.

***Constraints-based analysis.***

The *i*JN1462 model was exported from SimPheny as an SBML file and analyzed with the COBRA Toolbox v2.0 ([Schellenberger, Que et al. 2011](#_ENREF_21)) within the MATLAB environment (The MathWorks Inc.) or using the Cobrapy package within Jupyter Notebooks running Python 3.6. Tomlab CPLEX, Gurobi, and the GNU Linear Programming Kit (<http://www.gnu.org/software/glpk>) were used for solving the linear programing problems. The constrain-based model consists of a 2155 x 2929 matrix containing all the stoichiometric coefficients in the model of 2155 metabolites and 2929 reactions (**S**). Flux balance analysis (FBA) was used to predict growth and flux distributions ([Orth, Thiele et al. 2010](#_ENREF_18)). FBA is based on solving a linear optimization problem by maximizing or minimizing a given objective function **Z** subject to a set of constraints. The constraints **S·v = 0** correspond to a situation of steady-state mass conservation where the change in concentration of the metabolites as a function of time is zero. The vector **v** represents the individual flux values for each reaction. These fluxes are further constrained by defining lower and upper limits for flux values. For reversible reactions an upper and lower bound of -1000 mmol.gDW-1.h-1 and 1000 mmol.gDW-1.h-1 were used, respectively. A lower bound of 0 mmol.gDW-1.h-1 was used in case of irreversible reactions. For simulating condition-specific growth conditions, lower bounds of the corresponding exchange reactions were modified accordingly (See SI1). By default, the maximum growth rate was used as the cellular objective. Additional model constraints sink and demand reactions required for the functionality of the model can be found in SI1.

For modelling and analysis some additional constraints were applied. The bounds of the Pit7pp (Na-dependent phosphate transport) reaction were constrained to 0 mmol.gDW-1.h-1 to avoid unrealistic ATP production. Sink and demand reactions are modeling reactions required for the functionability of the model. Sink reactions are included in order to provide key metabolites of unknown origin while demand reactions are required for the removal dead end metabolites. *i*JN1462 includes two sink reactions, sink\_PHAg and sink\_pqqA which provide the PHA granule required for PHA biosynthesis and the initial peptide required for PQQ biosynthesis, respectively. There are 31 demand reactions of which six are required to allow dead metabolites to leave the system e.g., DM\_acmum6p, DM\_5DRIB, DM\_acgam, DM\_AMOB, DM\_doxopa, and DM\_tripeptide, and 25 are needed for allow the accumulation of cytoplasmic polymers, including the 24 monomers of PHA and polyphosphate

**Expansion and validation of nutrient sources supporting growth.**

To model the metabolic versatility of *P. putida* KT2440 primary literature and high-throughput nutrient phenotyping analyses of *Pseudomonas* spp were extensively scrutinized. The identification of a nutrient supporting growth in any *Pseudomonas* spp was the starting point for searching for potential genes encoding this ability in the genome of KT2440. If enough computational evidences supported the inclusion of the target catabolic pathway based on sequence identity, the corresponding reactions were added to the reconstruction. This iterative process resulted in the inclusion of hundreds of new reactions, many of them modeled in *i*JN1462 for the first time. This process concluded by adding the corresponding transport and exchange reactions. The transport reactions databases TCDB ([Saier, Reddy et al. 2014](#_ENREF_20)) and TransportDT ([Ren, Chen et al. 2007](#_ENREF_19)) were used for this purpose.

The potential nutrients sources supporting growth on *in silico* M9 medium (*i*M9, See SI1) including glucose, ammonium, inorganic phosphate, sulfate and Fe2+ as default carbon, nitrogen, phosphate, sulfur and iron sources, respectively were identified *in silico* by maximizing the BOF. Carbon sources were identified constraining the glucose uptake rate to zero and testing sequentially all the metabolites for growth which an exchange reaction was present in the reconstruction. Nitrogen, sulfur, phosphate and iron sources were predicted similarly by constraining the uptake of the corresponding default nutrient to zero. Any metabolite providing a non-zero growth rate was considered as true nutrient.

The predicted carbon and nitrogen carbon sources were subject to bibliomic and/or experimental validation. All disagreements between predicted and experimental values were further carefully analyzed. Several false negatives (growth in vivo, but not in silico) were resolved by manual gap-filling resulting in the inclusion of new reactions and genes in the reconstruction. This process contributed to the reannotation of many metabolic genes in *P. putida* (the annotation update can be found in SI1, Table S1). If the gene encoding the target enzymatic activity was unknown, we decided to fit the experimental data by including orphan reactions only if enough bibliomic support was available. For instance, while the gene encoding the coniferyl alcohol dehydrogenase (COALCDH) appears to be missing from the genome of the KT2440, coniferyl alcohol can be utilized as the sole carbon and energy source by this strain ([Jiménez, Nogales et al. 2010](#_ENREF_8)). For false positives (growth in silico, but not in vivo), the criteria that we followed was to keep the corresponding catabolic pathway in the model if strong computational evidence (sequence identity) was available. For instance, although *P. putida* KT2440 is unable to use ethylene glycol as a sole carbon source, the genes encoding its degradation are present in KT2440, and multiple *P. putida* strains grow on this compound ([Mückschel, Simon et al. 2012](#_ENREF_13)). Therefore, the incongruences still remaining in the model pave the way towards targeted identification of new genes responsible for orphan reactions, the deciphering of unknown regulatory mechanisms as well as a guide for future adaptive laboratory evolution (ALE) experiments.

**Growth experiments on carbon and nitrogen sources.**

Individual colonies of *P. putida* KT2440 and mutant strains were picked from the surface of cultures freshly grown on LB medium plates supplemented with 30 μg/ml of chloramphenicol, streaked onto M8 pre-growth medium plates (0.1% [wt/vol] glucose, 0.1 g/liter NH4Cl, 1 mM MgSO4, 0.6 mg/L Fe-citrate, and micronutrients), and grown overnight at 30°C. Pre-growth of cells on M8 pre-growth medium was sufficient to deplete nutrient reserves such that the subsequent growth assays with different carbon, nitrogen, and sulfur sources were dependent on the nutritional sources provided. The biomass of the overnight plates described above was recovered from the plate surface and suspended in 15 ml of M9 or M8 liquid medium (Daniels et al, 2010) to a turbidity at 660 nm (OD660) of 0.1. The wells of the microplates were filled with 180 μl of the cellular suspension, and 20 μl of each carbon, nitrogen, or sulfur source was added to reach a final concentration of 5 mM. For sulfur source assays, the MgSO4 in the medium was replaced with MgCl2. Positive-control wells consisted of full minimal medium containing glucose, NH4Cl, and MgSO4 as carbon, nitrogen, and sulfur sources, respectively; negative-control wells contained medium without cell inoculate.

All data recordings were performed using a type FP-1100-C Bioscreen C MBR analyzer system (OY Growth Curves Ab Ltd., Raisio, Finland) at 30°C, with continuous agitation. The turbidity was measured using a wideband filter at 420 to 580 nm every 60 min over a 24-h period. Each strain was assayed at least three times for each of the compounds tested, and plates were visually examined following each assay in order to verify the results.

**Gene essentiality predictions on *i*LB and glucose.**

*In silico* LB medium (*i*LB) was formulate based on the composition of commercial LB medium and the conditional essential gene analysis in *P. putida* ([Molina-Henares, De La Torre et al. 2010](#_ENREF_10)) (SI1). For predicting gene essentiality in glucose, a glucose minimal medium was simulated as described in SI1. The *singleGeneDeletion* function in the Cobra Toolbox ([Schellenberger, Que et al. 2011](#_ENREF_21)) with minimization of metabolic adjustment (MOMA) algorithm ([Segrè, Vitkup et al. 2002](#_ENREF_22)) were used to simulate knockouts. A gene was considered to be essential if its removal reduced the growth rate below 5% of the growth rate in the original model. Glucose, ammonium, inorganic phosphate, sulfate and Fe2+ were used as default carbon, nitrogen, phosphate, sulfur and iron sources, respectively. Growth in the *i*LB medium and nutrients used by *i*JN1462 (Figure 2) was simulated and the *singleGeneDeletion* function was used to identify the essential genes in each condition.

**12. Supplementary references.**

Blank, L. M., et al. (2008). "Metabolic response of *Pseudomonas putida* during redox biocatalysis in the presence of a second octanol phase." FEBS Journal **275**(20): 5173-5190.

Conrad, T. M., et al. (2011). Microbial laboratory evolution in the era of genome‐scale science.

Cornelius, S. P., et al. (2011). "Dispensability of Escherichia coli’s latent pathways." Proceedings of the National Academy of Sciences **108**(8): 3124-3129.

Duque, E., et al. (2007). Towards a Genome-Wide Mutant Library of *Pseudomonas putida* Strain KT2440. Pseudomonas. J.-L. Ramos and A. Filloux, Springer Netherlands**:** 227-251.

Fong, S. S., et al. (2006). "Latent Pathway Activation and Increased Pathway Capacity Enable Escherichia coli Adaptation to Loss of Key Metabolic Enzymes." Journal of Biological Chemistry **281**(12): 8024-8033.

Heavner, B. D. and N. D. Price (2015). "Transparency in metabolic network reconstruction enables scalable biological discovery." Current Opinion in Biotechnology **34**(0): 105-109.

Henry, C. S., et al. (2010). "High-throughput generation, optimization and analysis of genome-scale metabolic models." Nat Biotech **28**(9): 977-982.

Jiménez, J. I., et al. (2010). A Genomic View of the Catabolism of Aromatic Compounds in *Pseudomonas*. Handbook of Hydrocarbon and Lipid Microbiology. K. Timmis, Springer Berlin Heidelberg**:** 1297-1325.

King, Z. A., et al. (2015). "BiGG Models: A platform for integrating, standardizing and sharing genome-scale models." Nucleic Acids Research.

Molina-Henares, M. A., et al. (2010). "Identification of conditionally essential genes for growth of *Pseudomonas putida* KT2440 on minimal medium through the screening of a genome-wide mutant library." Environmental Microbiology **12**(6): 1468-1485.

Monk, J., et al. (2014). "Optimizing genome-scale network reconstructions." Nat Biotech **32**(5): 447-452.

Monk, J. M., et al. (2013). "Genome-scale metabolic reconstructions of multiple Escherichia coli strains highlight strain-specific adaptations to nutritional environments." Proc Natl Acad Sci U S A **110**(50): 20338-20343.

Mückschel, B., et al. (2012). "Ethylene Glycol Metabolism by *Pseudomonas putida*." Applied and Environmental Microbiology **78**(24): 8531-8539.

Nogales, J. (2014). A Practical Protocol for Genome-Scale Metabolic Reconstructions, Humana Press**:** 1-25.

Oberhardt, M. A., et al. (2011). "Reconciliation of Genome-Scale Metabolic Reconstructions for Comparative Systems Analysis." PLoS Comput Biol **7**(3): e1001116.

Oh, Y.-K., et al. (2007). "Genome-scale Reconstruction of Metabolic Network in *Bacillus subtilis* Based on High-throughput Phenotyping and Gene Essentiality Data." Journal of Biological Chemistry **282**(39): 28791-28799.

Orth, J. D., et al. (2011). "A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism--2011." Mol Syst Biol **7**: 535.

Orth, J. D., et al. (2010). "What is flux balance analysis?" Nat Biotech **28**(3): 245-248.

Ren, Q., et al. (2007). "TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels." Nucleic Acids Research **35**(suppl 1): D274-D279.

Saier, M. H., et al. (2014). "The Transporter Classification Database." Nucleic Acids Research **42**(D1): D251-D258.

Schellenberger, J., et al. (2011). "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0." Nat. Protocols **6**(9): 1290-1307.

Segrè, D., et al. (2002). "Analysis of optimality in natural and perturbed metabolic networks." Proceedings of the National Academy of Sciences **99**(23): 15112-15117.

Sohn, S. B., et al. (2010). "In silico genome-scale metabolic analysis of *Pseudomonas putida* KT2440 for polyhydroxyalkanoate synthesis, degradation of aromatics and anaerobic survival." Biotechnology Journal **5**(7): 739-750.