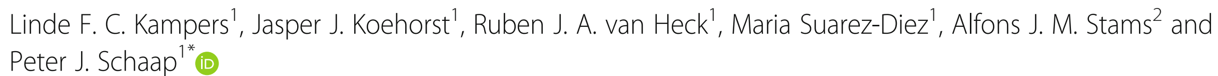
BMC Microbiology



RESEARCH ARTICLE Open Access

A metabolic and physiological design study of *Pseudomonas putida* KT2440 capable of anaerobic respiration

Linde F. C. Kampers[1], Jasper J. Koehorst[1], Ruben J. A. van Heck[1], Maria Suarez-Diez[1], Alfons J. M. Stams[2] and Peter J. Schaap[1*]

Abstract

Background: *Pseudomonas putida* KT2440 is a metabolically versatile, HV1-certified, genetically accessible, and thus interesting microbial chassis for biotechnological applications. However, its obligate aerobic nature hampers production of oxygen sensitive products and drives up costs in large scale fermentation. The inability to perform anaerobic fermentation has been attributed to insufficient ATP production and an inability to produce pyrimidines under these conditions. Addressing these bottlenecks enabled growth under micro-oxic conditions but does not lead to growth or survival under anoxic conditions.

Results: Here, a data-driven approach was used to develop a rational design for a *P. putida* KT2440 derivative strain capable of anaerobic respiration. To come to the design, data derived from a genome comparison of 1628 *Pseudomonas* strains was combined with genome-scale metabolic modelling simulations and a transcriptome dataset of 47 samples representing 14 environmental conditions from the facultative anaerobe *Pseudomonas aeruginosa*.

Conclusions: The results indicate that the implementation of anaerobic respiration in *P. putida* KT2440 would require at least 49 additional genes of known function, at least 8 genes encoding proteins of unknown function, and 3 externally added vitamins.

Keywords: *Pseudomonas*, Anaerobic respiration, Anaerobic fermentation, Computational design, Bioinformatics, Microbial lifestyle engineering

Background

Pseudomonas putida KT2440 is a HV1-certified [1], genetically accessible [2–7] and metabolically versatile [8, 9] species, which makes it an interesting adaptable industrial workhorse [10–12]. However, its strict aerobic lifestyle is an industrial disadvantage [4, 13–16] as the strict requirement for dissolved O2 results in increased costs of large-scale cultivation and may lead to unstable production rates due to inadequate local oxygen supply caused by oxygen fluctuations. Its strict aerobic nature also excludes production of O2-sensitive enzymes, pathway intermediates or target products.

Most *Pseudomonas* species are facultative anaerobes and use an inorganic compound such as nitrate as alternate terminal electron receptor. This includes species closely related to the *P. putida* KT2440 strain, such as *P. fluorescens* and *P. denitrificans*. Only one *Pseudomonas* species is capable of anaerobic fermentation: *Pseudomonas aeruginosa* [17–20]. *P. aeruginosa* is capable of arginine fermentation and pyruvate fermentation, although the latter only leads to prolonged survival under anoxic conditions, not to growth [18–20].

* Correspondence: peter.schaap@wur.nl
[1]Laboratory of Systems and Synthetic Biology, Wageningen University and Research Centre, Stippeneng 4, 6708, WE, Wageningen, The Netherlands
Full list of author information is available at the end of the article

© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

As there is a relatively short evolutionary distance between the strict aerobic *P. putida* KT2440 and facultative anaerobic *Pseudomonas* species compared to other anaerobic bacteria [21], it could be reasoned that a minimal set of adaptations would be required to change an aerobic *Pseudomonas* species into an facultative anaerobic one. Through the implementation of a rational engineering cycle, this strain could be adapted to a facultative anaerobic lifestyle. In an attempt to obtain a *P. putida* KT2440 derived strain capable of anaerobic fermentation a Design, Build, Test, Learn-engineering cycle [22] was performed in earlier work [23] to obtain an *P. putida* KT2440 strain capable of anaerobic fermentation. Using genome metabolic models (GSMs) iJP962 and iJP746 combined with a protein domain comparison (PDC) between six aerobic *Pseudomonas putida* strains including KT2440 and six facultative anaerobic *Pseudomonas* strains, three key enzymes were selected and included in the final design: acetate kinase (encoded by *ackA*), dihydroorotate dehydrogenase (*pyrK-pyrD B*) and ribonucleotide triphosphate reductase class III (*nrdD-nrdG*). This design was built, and the resulting recombinant strain showed growth under micro-oxic conditions [23]. Earlier work already described an increase in survival rates upon introduction of solely acetate kinase [4, 14], and since the model predictions used in the design only considered full anoxic conditions, survival rates of the recombinant strains under anoxic conditions needed to be tested.

Here, we (i) determined the survival rates of the previously constructed recombinant strains under anoxic conditions, (ii) identified limitations for anaerobic growth through respiration, and (iii) composed a new design for a recombinant *P. putida* KT2440 capable of anaerobic respiration. In pursuit of this goal we expanded upon earlier work using the current wealth of genome data available on *P. putida* and other *Pseudomonas* species by inclusion of 1628 strains in an extensive comparison of the protein domain content [24]. Random forest, a machine learning method, was used to identify key protein domains associated with "anaerobic growth". Transcriptome data of the *Pseudomonas aeruginosa* type strain PA14 cultures grown in 14 different conditions [25] were also considered and integrated with previous and newly obtained GSM simulation results to compose a final design.

Methods

Bacterial strains and cultivation conditions

Bacterial strains and plasmids are listed in in Table S1. For plasmid construction see previous work [23]. *E. coli* CC118λpir was used for cloning procedures and plasmid maintenance, and was routinely cultivated at 37 °C in aerated conditions in LB medium (10 g/l tryptone, 10 g/l NaCl and 5 g/l yeast extract), optionally containing antibiotics for selection (50 μg/ml kanamycin or 50 μg/ml ampicillin). For solid medium, 15 g/l agar was added to the medium. *P. putida* KT2440 was routinely cultivated under oxic conditions at 30 °C in LB medium. Experiments were performed in De Bont minimal medium [26] (3.88 g/l K_2HPO_4, 1.63 g/l $NaH_2PO_4 \cdot 2H_2O$, 2.00 g/l $(NH_4)_2SO_4$, 0.1 g/l $MgCl_2 \cdot 6H_2O$, 10 mg/l EDTA, 2 mg/l $ZnSO_4 \cdot 7H_2O$, 1 mg/l $CaCl_2 \cdot 2H_2O$, 5 mg/l $FeSO_4 \cdot 7H_2O$, 0.2 mg/l $Na_2MoO_4 \cdot 2H_2O$, 0.2 mg/l $CuSO_4 \cdot 5H_2O$, 0.4 mg/l $CoCl_2 \cdot 6H_2O$, 1 mg/l $MnCl_2 \cdot 2H_2O$), with 20 g/l gluconic acid as the sole carbon source. In previous work, different carbon sources were tested for optimal performance [23]. Gluconic acid was used for optimal growth by eliminating ATP consumption for substrate uptake due to passive membrane transport. The medium was supplemented with 50 μg/ml kanamycin when indicated. Precultures were prepared aerobically overnight at 200 rpm at 30 °C.

Anoxic survival experiment

Oxygen gradients served to allow the recombinant strains to grow in micro-oxic conditions as described in [23]. Anoxic cultivation of *P. putida* KT2440 recombinants unpassed or passed over oxygen gradients was performed at 30 °C in 50 ml glass 20 mm aluminium crimp cap vials with rubber stoppers (Glasgerätebau Ochs Laborfachhandel e.K.) in 30 ml DeBont with 20 g/l gluconic acid, 1 mg/l resazurin and 50 μg/ml kanamycin as selection marker for recombinant strains. Where indicated, a 1000x diluted vitamin mix was added (0.02 g/l biotin, 0.2 g/l nicotinamide, 0.1 g/l p-aminobenzoic acid, 0.2 g/l thiamin, 0.1 g/l pantothenic acid, 0.5 g/l pyridoxamine, g/l cyanocobalamin, 0.1 g/l riboflavin). Before inoculation, the vials were gas exchanged with CO_2/N_2. Inoculation was done with aerobically pre-cultured bacterial sample at an OD_{600} of 0.05. Approx. 8 h after inoculation, the resazurin became completely colourless, indicating obtainment of anaerobic conditions. Samples were taken using sterile CO_2 flushed 1.5″ needles (BD Microlance) and 3–5 ml syringes (ThermoFisher) to avoid O_2 exposure. Anoxic conditions were ensured as the resazurin turned from colourless to bright pink within seconds in extracted samples. Survival rates were analysed by colony forming units (CFU) determination. A dilution series was made and five drops of 10 μl per dilution were applied onto LB-agar plates without selection marker, which were incubated o/n at 30 °C. Colonies were counted manually, and photos were taken of the plates. Gram-staining was performed as an intermediate check of culture purity, according to manufacturers' instructions (Gram-staining kit Machery-Nagel, Germany), while genome sequencing was applied to ensure culture purity at the start and end of every experiment.

Statistical analysis
Experiments were independently repeated six times with biological triplicates in each separate experiment. Figures represent the mean values of corresponding biological triplicates and the standard deviation. The level of significance of the differences when comparing results was evaluated by means of analysis of variance (ANOVA), with $\alpha = 0.05$.

Genome annotation
Information on the oxygen requirements of 16,989 *Pseudomonas* strains was obtained from the Gold database [27]. Per species, extensive literature research was performed to validate their aerobicity (Data S5). One thousand six hundred twenty-eight Genomes of facultative anaerobic and strict anaerobic strains from the *Pseudomonas* genus were obtained from the European Nucleotide Archive repository in March 2015 [28]. All genomes were de-novo annotated in SAPP [29] using Prodigal for gene prediction (version 2.6) [30], 2010] and InterProScan version 5.4–47.0 [31] for functional annotation using Pfam [32].

Comparisons of protein domain content
The positions (start and end on the protein sequence) of the protein domains and their order in a protein when multiple domains were present, were used to identify domain architecture (i.e. combinations of protein domains). Protein domain architectures were labelled by the ordered list of Pfam identifiers as described in [33]. Protein domain architectures identified in each genome sequence were stored in a matrix. From this a binarized domain architecture presence-absence matrix was extracted and used as input for principal component analysis using the standard R-package prcomp and hierarchical clustering using the standard R-package hclust.

Gene persistence
The persistence of a gene in a taxonomic group or group of genomes can be defined as

$$Persistence = \frac{N(orth)}{N}$$

where N (orth) is the number of genomes carrying a given ortholog and N is the number of genomes considered [24]. For the set of 1628 considered genomes. Orthologous genes were identified through identity of protein domain architectures considering copy number. Resulting protein domain contents were analysed through protein domain comparison (PDC).

Feature selection using random forest
The random forest classification algorithm was used to classify the genome sequences in aerobic and facultative anaerobic species with the goal to identify the domains (features) responsible for the separation in these two groups (feature selection). Three hundred randomly selected genomes from aerobic and anaerobic *Pseudomonas* species were selected to train random forest models. The process was repeated one hundred times. The resulting 100 different models were used to weigh 5831 protein domains from both aerobic and anaerobic *Pseudomonas* species. Variable selection was used to identify the most influential domains for classification in aerobic and facultative anaerobic strains, yielding 100 Gini coefficients, representing the importance of a protein domain for separation per protein domain. Gini coefficients were combined into the cumulative Gini coefficient. The resulting protein domains were separated into aerobic/anaerobic specific protein domains before further analysis.

Transcriptome data analysis
A publicly available *P. aeruginosa* transcriptome data set was retrieved from GEO database (accession number GSE55197) [25]. This dataset contains 47 samples corresponding to 14 environmental conditions, including changes in growth temperature, growth stage, osmolarity, concentration of ions in the media, and surface attachment and anaerobic respiration. For every gene, the log2 fold change of its expression values was calculated in comparing every possible condition with anaerobic respiration. Missing or infinity values arising from genes with very low counts in some condition(s) were imputed to 0 or ± 4, according to the significance of the differential expression (False discovery rate, $fdr < 0.05$). Normalization, fold change computations and differential expression analysis were performed using the R package DESeq [34].

Genome-scale metabolic models
In this study we used the *P. putida* genome-scale metabolic models (GSMs) iJP962, iJN746 and iJN1411 [3, 5, 35]. iJN1411 was obtained directly from the authors [35]. GSM simulations were performed as described in [23], with uptake rates of up to 1000 mmol $gdw^{-1} h^{-1}$ (gdw: grams dry weight) of copper, cobalt, iron, protons, water, sodium, nickel, ammonia, phosphate, sulphate, and nitrate, a maximal glucose uptake rate of 6.14 mmol $gdw^{-1} h^{-1}$, based on experimentally measured uptake rates [36]. Thus, the in silico medium composition mimics the De Bont minimal medium used for the in vivo experiments.

Results

Insertion of acetate kinase in P. putida KT2440

Previous designs to obtain *P. putida* strains surviving anoxic conditions were conceptually based on the hypothesis that survival in anoxic conditions was prevented by a lack of energy conservation and redox balancing [4, 13–16]. Expression of the acetate kinase gene from *P. aeruginosa* and *E. coli* was reported to result in an extended survival under anoxic conditions [4, 14]. Expression of the acetate kinase gene (*ackA*) from *E. coli* combined with class I dihydroorotate hydrogenase (*pyrK-pyrD B*) and class III ribonucleotide triphosphate reductase (*nrdD-nrdG*) from *L. lactis* successfully led to growth under micro-oxic conditions [23].

To determine the tolerance to anoxic conditions of a *P. putida* KT2440 recombinant strain enriched with *ackA, pyrK-pyrD B* and *nrdD-nrdG* and of a negative control strain carrying an empty plasmid to anoxic conditions and to analyse the effect of an adaptation over oxygen gradients as performed earlier [23], an 18-day anoxic survival experiment was performed. After inoculation at a standardized cell density under oxic conditions, cultures were incubated overnight in capped gas-exchanged vials in oxygen-depleted medium (see Materials and Methods). To optimise growth gluconic acid was used as the main carbon source. In a previous design, gluconic acid was shown to offer better results compared to a series of other carbon sources [23]. The survival rate was determined by performing colony forming unit (CFU) counts at set time points over a period of 18 days, with T0 being the start of the experiment in anoxic conditions (supplementary Figures S1, S2, S3, S4, S5). The results showed that in anoxic conditions there is no significant difference in survival rates between the negative control and any of the recombinant strains tested (ANOVA $\alpha = 0.05$). Under these conditions, only the positive control, *E. coli* BW25113 harbouring an empty plasmid, survived.

Design requirements for a P. putida KT2440 derivative strain capable of anaerobic respiration

The failure of the previous, fermentative, design [23] to grow under anoxic conditions could be explained by the heavy reliance on the two state of art genome-scale models (GSMs) used in this design, which currently do not include an accurate representation of the complete redox balance and its intricate involvement in the metabolism. Additionally, while the protein domain comparison performed in the previous study showed apparent differences between aerobic and anaerobic strains in availability of protein domains, this analysis was performed on a limited set of strains.

Many facultative anaerobic *Pseudomonas* species are incapable of anaerobic fermentation, but rather perform anaerobic respiration. The close phylogenetic distances between some of these facultative anaerobic *Pseudomonas* species and *P. putida* KT2440 may suggest that acquiring a facultative anaerobic lifestyle via anaerobic respiration would require less genetic changes. To come to a rational design of *P. putida* KT2440 capable of anaerobic respiration, the previous methods were thus expanded upon by (i) using significantly more facultative anaerobic and aerobic *Pseudomonas* strains for domain analysis, (ii) inclusion of iJN1411, the latest metabolic reconstruction of *P. putida* KT2440 [35], and (iii) incorporation of an elaborate transcriptome analysis of anaerobic respiration of *P. aeruginosa* strains grown under anoxic conditions in comparison with 13 aerobic growth conditions [25]. Inclusion of such transcriptome data would show gene regulation due to growth under anoxic conditions, improving the design as it complements genome-based methods.

For protein domain comparisons, the Pfam domain content of *P. putida* KT2440 was compared with 1627 other *Pseudomonas* strains with fully sequenced genomes. For each strain, a literature search was performed to determine oxygen requirements, yielding 344 obligate aerobic strains including KT2440 and 1284 facultative anaerobic strains. Strain specific differences in protein domain content were visualised using principal component analysis (PCA), and hierarchical clustering using domain presence/absence as input (Fig. 1). Both the PCA and the hierarchical clustering show a separation between several facultative anaerobic strains and the rest of the considered strains (among which *P. putida* KT2440). However, it should be noted that only a small fraction of the total variance is explained by the first two principal components. This separation is also apparent in the dendrogram, suggesting that significant differences could be found in protein domain content.

We assumed that domains essential for anaerobic respiration are highly persistent in facultative anaerobic strains but show a lower persistence in obligate aerobic strains. The strategy to obtain this protein domain core is outlined in Fig. 2. A "long list" of anaerobic protein domains was generated by comparing domain persistence between aerobic versus anaerobic strains. First a 95% persistence threshold was applied, to obtain a "domain core" of domains present in at least 95% of the genomes of "aerobic" strains and in the "anaerobic" strains analysed. These aerobic and anaerobic domain cores were used as input for subsequent comparative analysis and for the first list were split into "shared between aerobic and anaerobic species" (Shared domain core), "specific for aerobic species" (Aerobe specific domain core) and "specific for anaerobic species" (Anaerobe specific domain core) creating a long list of 427 anaerobe specific protein domains. A second long list was created by

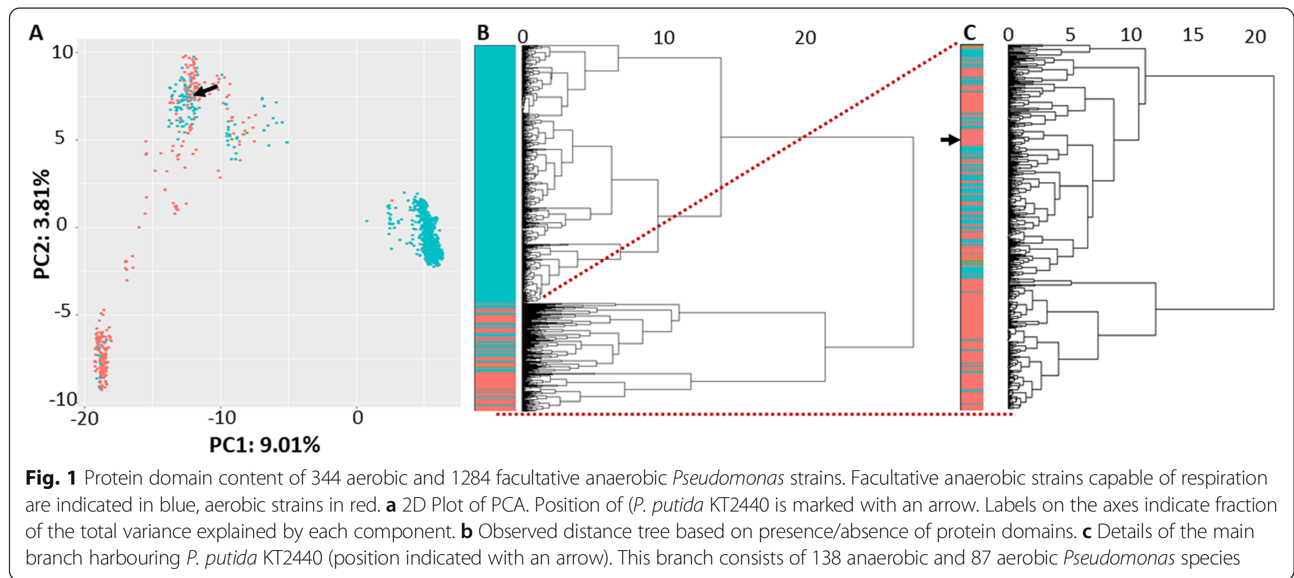


Fig. 1 Protein domain content of 344 aerobic and 1284 facultative anaerobic *Pseudomonas* strains. Facultative anaerobic strains capable of respiration are indicated in blue, aerobic strains in red. **a** 2D Plot of PCA. Position of (*P. putida* KT2440 is marked with an arrow. Labels on the axes indicate fraction of the total variance explained by each component. **b** Observed distance tree based on presence/absence of protein domains. **c** Details of the main branch harbouring *P. putida* KT2440 (position indicated with an arrow). This branch consists of 138 anaerobic and 87 aerobic *Pseudomonas* species

A Genome domain comparisons
Aerobic *Pseudomonas* genomes (344)
Anaerobic *Pseudomonas* genomes (1284)
Annotate genomes
Annotate genomes
Extract protein domains (1029)
Extract protein domains (1182)
All annotated protein domains
Cluster based on genome similarity
95% threshold
1% threshold
90% threshold
14
1015
167
170
1012
17
274
755
427
248
934
95
Compared and combined lists of protein domains
Anaerobic specific protein domains
Aerobic specific protein domains

B Transcriptomics
Transcriptomics data on *P. aeruginosa* in 14 different conditions
Calculate the fold change of all possible pairwise comparisons
Filter based on differential regulation >4
Cluster based on anoxic condition
Cluster based on all conditions
71
74
19
11

C GSM simulations
GSM iJP962
GSM iJN1411
Identify essential oxygen-dependent reactions
Identify essential oxygen-dependent reactions
ackA
pyrK-pyrD B
nrdD-nrdG
Pyoverdine
Active B6
Lipoate
Active B12
LPS
PQQ
Protoheme
Ubiquinol 8
Heme

Potential limitations to anaerobic growth in *P. putida*
Random Forest Analysis
Design
Upregulated genes
Downregulated genes
Missing genes or supplements

Fig. 2 Overview of in silico approaches to identify limitations to anaerobic respiration in P. putida. **a** Comparative genomics workflow. Genomes of the P. putida group and the anaerobic Pseudomonas group were systematically annotated using SAPP [24, 29], the protein domains were extracted, and both all domains or only the domains common to all anaerobic Pseudomonas species (the core domains) were selected using a 95% persistence threshold. Analysis was performed on the whole set of genomes (left) or a genome cluster of closely related strains (right). Each of these methods resulted in a list of protein domains related to an aerobic lifestyle (purple) or an anaerobic lifestyle (light green). **b** Transcriptome analysis. **c** GSM simulations. GSM iJP962 [5] and iJN1411 [35] were expanded with indicated reaction sets and tested for anaerobic growth under anaerobic conditions. Colours indicate final implementation in the design (green). Model and genome base predictions were combined to obtain a final design

the same input but searching for the reverse, a separation based on domains with a very low persistency in aerobic or anaerobic strains. For this a no more than 1% threshold was applied creating a long list of 167 anaerobe specific protein domains.

The dendrogram presented in Fig. 1 indicated a possible early branch split between a large group of exclusively anaerobic *Pseudomonas* strains and a mixed group, including *P. putida* KT2440, containing 138 facultative anaerobic and 87 obligatory aerobic *Pseudomonas* strains (Fig. 1 panel C). Using this split, two “restricted” lists were built by comparing domain persistence as outlined above, but now evaluating only *Pseudomonas* strains present in the mixed branch. For the restricted lists, a 90% persistence threshold and a 1% persistence threshold were used, creating two anaerobic species-specific protein domains lists of 170 and 248 domains, respectively. The four different lists of protein domains essential for anaerobic growth were compared and manually further annotated. Results are summarized in Table 1 and Fig. 2.

As outlined in the Materials and Methods section, the domain content of the facultative anaerobic and the obligatory aerobic *Pseudomonas* strains were used to train a random forest classifier with the goal to identify those domains (features) that are mostly responsible for classification. Gini coefficients and cumulative Gini coefficients for each domain are provided in Data S9. From the 5831 domains that were used as input for the classifier, 5 have a cumulative Gini coefficient ≥ 100, as summarized in Table 1. Gini scores were added as weight to the four protein domain lists derived above.

Transcriptome data obtained from *P. aeruginosa* PO14 grown under 14 different environmental conditions including anoxic conditions [25] was re-analysed for genes that were consistently differentially expressed during anaerobic respiration (see the Materials and Methods section for details). By calculating for every gene, the log2fold change of its expression values in every possible condition compared with anaerobic respiration, 175 protein domains were identified. A heatmap was used to visualise up- and down-regulated genes under anoxic conditions. Regulation due to anoxic growth was considered to be significant when the same behaviour (up- or down-regulation) was observed in at least 7 of the 13 pair-wise comparisons and a fold change of at least 4 was observed in at least three of these comparisons. Protein domain architectures corresponding to the selected locus tags were identified. Based on the differential expression and similar efforts in literature [13], 22 genes encompassing 35 protein domains were selected.

Genome-scale models were used to simulate anoxic conditions. The absence of any reaction products impeding growth due to the simulated lack of oxygen were pinpointed and traced back to proteins and their encoding genes that either need oxygen as a substrate or that cannot be made without oxygen present, resulting in substrates that could thus not be produced under anoxic conditions. Genes and substrates were verified through literature analysis to be essential for growth (Table 1).

Design considerations

A comparison was made between the different lists obtained (Table 1) and previous efforts [4, 13, 14, 23] resulting in an extensive overview of the many hurdles to overcome to build a *P. putida* KT2440 strain capable of anaerobic respiration. The various lists were compared by evaluating the function of each gene starting with the encoded domain annotation, checking for domain co-existence in operonic structures, comparing metabolic functions with GSM data and with gene regulation data. The importance of each protein domain was determined using the random forest analysis (Data S9) as input. In this way, different lists could be combined and reduced to a list of 57 genes. Furthermore, a supplement of 3 vitamins is required.

The selected genes can be separated into various categories based on their functions: Nitrogen metabolism (45 domains in 35 genes), Hydrogenases (9 domains in 9 genes), Cytochrome C (1 domains in 1 genes), Pyrimidine and amino acid biosynthesis (6 domains in 3 genes if 3 vitamins added), ATP production (1 domains in 1 genes), and Domains of Unknown Function (indirectly associated with anaerobic respiration) (8 domains).

Nitrogen metabolism

Most *Pseudomonas* species capable of anaerobic respiration do so using nitrite or nitrate as alternative terminal electron acceptor. Of the 49 known genes found vital for anaerobic respiration, 35 are either directly or indirectly involved in nitrogen metabolism. With nitrate as the final electron acceptor in anaerobic respiration the largest amount of energy can be conserved when compared to other final electron acceptors such as sulphate, iron (III), manganese (II), or selenate [37]. *P. putida* KT2440 lacks the nitrate/nitrite respiration pathway, which was resolved in earlier studies by inserting either a Nir-Nar or a Nor plasmid [13]. This resulted in extended survival under anoxic conditions, but not to growth. Our transcriptomics and protein domain analysis indicated that the combination of the operons of both the Nir-Nar and the Nor operon are required (Table 2). The operons include genes of the Nif, RhfH, Nqr, Rnf, Dau, Nar, Nir and Nor, and Moa protein families, and are required for energy conservation, cofactor biosynthesis, amino acid biosynthesis, nitrogen metabolism, nitrate-, nitrite- and nitrogen transporters, nitrate-, nitrite-, nitric oxide and

Table 1 Respiratory design of a facultative anaerobic P. putida KT2440. Number of unique protein domains obtained

Method	# Unique protein domains
Genome domain comparisons	
Input aerobic domain core	1029
Input anaerobic domain core	1182
Long list, 1628 strains [95% threshold]	
Shared domain core	755
Aerobic specific domain core	274
Anaerobic specific domain core	427
Long list, 1628 strains [1% threshold]	
Shared domain core	1015
Aerobic specific domain core	14
Anaerobic specific domain core	167
Restricted list, 225 strains [90% threshold]	
Shared domain core	1012
Aerobic specific domain core	17
Anaerobic specific domain core	170
Restricted list, 225 strains [1% threshold]	
Shared domain core	934
Aerobic specific domain core	95
Anaerobic specific domain core	248
Transcriptome analysis	175
GSM simulations	18
RandomForest [input]	5831
Domains with a cumulative Gini coefficient ≥ 20	360
Domains with a cumulative Gini coefficient ≥ 100	5

nitrous oxide reductases and several regulatory proteins (Table 2 and Additional file 18). Of the 45 protein domains or 35 genes we identified within this category, only 15 genes had been previously found (*narK1, narK2, narG, narH, narJ, narI narX, narL, nirF, nirQ, nirM, nirS, nirJ, nirL* within the Nir-Nar operon, *norC, norB, norD, nosR* within the Nor operon) [13]. Ureohydrolases such as Arg1, SpeB, HutG and Pah facilitate the ammonia to urea conversion, with urea as the principle product of nitrogen excretion.

Hydrogenases

Included in the anaerobic respiration design are 9 hydrogenase subunits: HupH, HypA-F, HyaE, and HybE. Hydrogenases catalyse the reversible oxidation of molecular hydrogen, fulfilling a regulatory role in balancing the redox state. The redox state of the cell and the availability of O2 are regulatory signals in facultative anaerobic species [38]. [FeFe] And [NiFe]-hydrogenases are widely distributed under anaerobic species. These hydrogenases are only produced under anoxic conditions, and most [NiFe]-hydrogenases are inactivated by oxygen, only to be re-activated under reducing conditions [39].

P. putida KT2440 lack hydrogenases necessary for proton reduction or coupling H_2 oxidation to energy yielding processes under anoxic conditions, and the necessary hydrogenase chaperones, assembly, maturation and formation proteins (Table 3).

None of these genes have been recognised in previous research for their importance in anaerobic respiration.

Cytochrome C

Included in the anaerobic respiration design are 3 C-type cytochromes. C-type cytochromes account for a vital step in ATP bio-generation via the proton motive force (Table 4). Anaerobically, cytochrome C 551 (NirN), C 552 transfer electrons to nitrite reductase (NirS) and nitric-oxide reductase (NorB, NorC). The importance of NirN and NirC (the precursor of NirN) was demonstrated in [13] (Table 2).

The PDC also indicates the need for cytochrome C 552 (Tables 2, 4). The enzyme cytochrome C nitrite reductase (C 552), amongst other important functions, catalyses the six-electron reduction of nitrite to nitrogen as one of the key steps in denitrification. Nitrogen is then reduced to ammonium in the ammonification pathway. C552 thus participates in the anaerobic energy metabolism of dissimilatory nitrate ammonification.

In addition, cobalamin-independent methionine synthase is important. This methionine synthase is responsible for precursor formation of C 551 that can be produced without using vitamin B12 (see Pyrimidine and amino acid biosynthesis, Table 5). This might be a key component for anaerobic growth, since both the protein domain analysis and the GSM iJN1411 [35] predict that, amongst other vitamins, the active form of vitamin B12 can only be bio-generated in the presence of oxygen in *P. putida* KT2440.

Pyrimidine and amino acid biosynthesis

Included in the anaerobic respiration design are 2 genes involved in pyrimidine and amino acid synthesis, and additional bottlenecks that can be solved by adding 3 vitamins to the medium. Earlier GSM simulations with iJP962 indicated that alternate genes must be inserted for dihydroorotate dehydrogenase and ribonucleotide triphosphate reductase type II for pyrimidine and ultimately DNA and RNA biosynthesis [23]. Both the protein domain analysis and GSM simulations using the iJN1411 metabolic model predicted that cobalamin (vitamin B12), pyridoxal-5-phosphate (vitamin B6) and menaquinone (vitamin K2) cannot be produced under anoxic conditions.

Table 2 Respiratory design of a facultative anaerobic P. putida KT2440. Genes related to nitrogen metabolism

InterPro Entry	PFAM	Name	Abbreviation	Function	Source
	PF00491	**ureohydrolases**		Ammonia to urea conversion	PDC
		Arginase	ArgI	L-arginine + H(2) O < => L-ornithine + urea	PDC
		Agmatinase	SpeB	Agmatine + H(2) O < => putrescine + urea	PDC
		Formimidoylglutaminase	HutG	N-formimidoyl-L-glutamate + H(2) O < => L-glutamate + formamide	PDC
IPR015868	PF04960	Glutaminase		Glutamine + H2O → Glutamate + NH3	PDC, T
IPR000292	PF01226	Formate/nitrite transporter		Transporter	PDC, T
IPR025736	PF13556	PucR C-terminal helix-turn-helix domain	PucR	PucR-like transcriptional regulator	PDC, T
IPR000825	PF01458	**Uncharacterized protein family (UPF0051)**		Chaperone proteins for nitrogenase production	PDC, T
			NifS	Metallocluster formation Nitrogenase	PDC, T
			NifU	Metallocluster formation Nitrogenase	PDC, T
IPR005346	PF03658	Ubiquitin	RnfH family	Electron transport	PDC, T
	PF02508	Rnf-Nqr subunit, membrane protein	Rnf-Nqr	Nitrogen fixation	PDC, T
	PF03116	**Nqr2 family**	Nqr2	Nitrogen fixation	PDC, T
		RnfD family	RnfD	Nitrogen fixation	PDC, T
		RnfE family	RnfE	Nitrogen fixation	PDC, T
	PF03060	Nitronate monooxygenase		Nitrogen metabolism	PDC, T
IPR010349	PF06089	L-asparaginase II		Hydrolysis of L-asparagine to L-aspartate and ammonium.	PDC, T
PA3862	PF02423		DauB	NAD(P)H-dependent anabolic L-arginine dehydrogenase	PDC, T, Steen et al., 2012
PA3863	PF01266		DauA	FAD-dependent catabolic D-arginine dehydrogenase	PDC, T, Steen et al., 2012
PA3864	PF08348		DauR	Transcriptional regulator of the dauBAR operon	PDC, T, Steen et al., 2012
	PF13309				PDC, T, Steen et al., 2012
PA14_13750	PF07690	Nitrite extrusion protein (putative)	NarK2	Membrane protein; Transport of small molecules	PDC, T, Steen et al., 2012
PA14_13770	PF07690	Nitrite extrusion protein	NarK1	Membrane protein; Transport of small molecules	PDC, T, Steen et al., 2012
PA3875	PF14710		NarG	Energy metabolism	PDC, T, Steen et al., 2012
	PF00384				PDC, T, Steen et al., 2012
	PF01568				PDC, T, Steen et al., 2012
PA14_13800	PF13247	Nitrate reductase	NarH	β-subunit, Energy metabolism	PDC, T, Steen et al., 2012
	PF14711				PDC, T, Steen et al., 2012
PA14_13810	PF02613	Nitrate reductase	NarJ	Δ-chain, Energy metabolism	PDC, T, Steen et al., 2012
PA14_13830	PF02665	Nitrate reductase	NarI	γ-chain, Energy metabolism	PDC, T, Steen et al., 2012
PA3878	PF02518	two-component sensor	NarX		PDC, T, Steen et al., 2012

Table 2 Respiratory design of a facultative anaerobic P. putida KT2440. Genes related to nitrogen metabolism *(Continued)*

InterPro Entry	PFAM	Name	Abbreviation	Function	Source
	PF00672				PDC, T, Steen et al., 2012
	PF07730				PDC, T, Steen et al., 2012
	PF13675				PDC, T, Steen et al., 2012
PA3879	PF00072	two-component response regulator	NarL	Response regulator	PDC, T, Steen et al., 2012
	PF00196				PDC, T, Steen et al., 2012
PA14_ 13850	PF04055	Heme d1 biosynthesis protein	NirJ	Heme d1 Biosynthesis	PDC, T, Steen et al., 2012
PA0516	PF02239	Heme d1 biosynthesis protein	NirF	Heme d1 Biosynthesis	PDC, T, Steen et al., 2012
PA0514		Heme d1 biosynthesis protein	NirL	Heme d1 Biosynthesis	PDC, T, Steen et al., 2012
PA0520	PF07728		NirQ	Regulatory protein	PDC, T, Steen et al., 2012
	PF08406		CbbQ	post-translational activation of Rubisco -- photosynthesis	PDC, T, Steen et al., 2012
			NorQ,	Post-translational activation of Rubisco -- photosynthesis	
PA14_ 06750	PF13442	nitrite reductase precursor	NirM	Cofactor biosynthesis, Energy metabolism	PDC, T, Steen et al., 2012
PA3870	PF00994	molybdopterin biosynthetic protein A1	MoaA1	Biosynthesis of cofactors, prosthetic groups and carriers	PDC, Steen et al. 2012
	PF03453				PDC, Steen et al. 2012
	PF03454				PDC, Steen et al. 2012
PA14_ 13260	PF00994	molybdopterin biosynthetic protein B1	MoaB1		PDC
	PF00394	**Multicopper oxidase**			PDC, T
PA0519	PF13442	Nitrate reductase	NirS	Energy metabolism	PDC, T, Steen et al., 2012
		Nitrate reductase	NirS	Energy metabolism	PDC, T, Steen et al., 2012
	PF02239	Nitrate reductase	NirS	Energy metabolism	PDC, T, Steen et al., 2012
	PF05940	**NnrS protein**			PDC, T
			NirK	Reduction of nitrite to nitrous oxide	PDC, T
			Nor	Reduction of nitrite to nitrous oxide	PDC, T
PA14_ 06810	PF00034	nitric-oxide reductase	NorB-NorC	Subunit B, C	PDC, T, Steen et al., 2012
PA14_ 06830	PF00115	nitric-oxide reductase	NorB-NorC	Subunit B, C	PDC, T, Steen et al., 2012
PA14_ 06840	PF00092		NorD	Putative dinitrification protein	PDC, T, Steen et al., 2012
PA14_ 20230	PF04205		NosR	Regulatory protein for N2O reductase	PDC, T

T Trancriptomics, *PDC* Protein Domain Comparisons, *GSM* Genome Scale Modelling [30, 31];. Printed in bold are classes of genes, the genes belonging to that class listed directly underneath

Table 3 Respiratory design of a facultative anaerobic P. putida KT2440. Genes encoding for hydrogenases

InterPro Entry	PFAM	Name	Abbreviation	Function	Source
IPR027394	PF14720	NiFe/NiFeSe hydrogenase		Reversible oxidation of molecular hydrogen	PDC
	PF09459	Ethylbenzene dehydrogenase		Anaerobic degradation of hydrocarbons	PDC
IPR001501	PF00374	Nickel-dependent hydrogenase		Catalysis the reversible activation of hydrogen	PDC
IPR001109	PF01455	HupF/HypC family	HupF/HypC	[NiFe]-hydrogenase and other nickel metalloenzymes synthesis	PDC
IPR000671	PF01750	Hydrogenase maturation protease		Hydrogenase maturation	PDC
IPR002780	PF01924	Hydrogenase formation	HypA	Hydrogenase formation	PDC
IPR006894	PF04809	Hydrogenase expression protein	HupH	Hydrogenase synthesis, C-terminal domain	PDC
IPR010893	PF07449	Hydrogenase-1 expression protein	HyaE	Hydrogenase assembly	PDC
IPR023994	PF11939	Chaperone for [NiFe]-hydrogenase assembly	HybE	[NiFe] hydrogenases assembly chaperone	PDC
	PF13237	4Fe-4S dicluster domain		Mediate electron transfer	PDC
IPR000688	PF01155	Metallochaperone	hypA	[Ni,Fe]-hydrogenase and urease chaperone	PDC
IPR007038	PF04955	HupE / UreJ protein	HupE / UreJ	Hydrogenase / urease accessory proteins	PDC

T Trancriptomics, *PDC* Protein Domain Comparisons, *GSM* Genome Scale Modelling [30, 31];

Crespo et al. showed that class II RNRs depend on adenosylcobalamin or vitamin B12 (cobalamin) to generate its radical independently of oxygen [40]. Cobalamin is a complex essential cofactor for many enzymes mediating methylation, reduction, and intramolecular rearrangements, and for methionine synthase. There is a recognised distinction between aerobic and anaerobic generation of cobalamin [41, 42]. The routes differ in terms of cobalt chelation (via CobNST complex in the aerobic pathway, via precorrin-2 with CbiK in the anaerobic pathway) and oxygen requirements. The enzymes CobI, CobG, CobJ, CobM, CobF, CobK, CobL, CobH, CobB and CobNST form the aerobic pathway. CbiK, CbiL, CbiH, CbiF, CbiG, CbiD, CbiJ, CbiET, CbiC and CbiA form the anaerobic route [31, 41, 43]. Surprisingly, the protein domain comparison yielded none of the enzymes of the anaerobic pathway for vitamin B12 synthesis, but instead CobT and CbtB, both described as important for the aerobic pathway [41]. According to the extensive analysis, these specific protein domains linked to these genes are not present in aerobic species analysed but only in anaerobic species. It was found that in the anaerobic bacterium *Eubacterium limosum*, CobT functions as an activator for a range of lower ligand substrates including DMB, determining cobamide diversity. The specific function of CbtB is unknown [41, 42].

Vitamin B6 is required for a wide variety of processes [44]. There are many vitamin B6-dependent proteins involved in amino acid biosynthesis, amino acid catabolism, antibacterial functions, iron metabolism, carbon metabolism, nucleotide utilization, cofactors for biotin, folate and heme, NAD biosynthesis, cell wall metabolism, tRNA modification, regulation of gene expression and biofilm formation.

Vitamin K2 is responsible for electron transport during anaerobic respiration. However, knock-out experiments in *E. coli* showed that upon loss of menaquinone and vitamin K1 only 3% of theoretical yield was obtained, but this was instantly revived to 44% upon supplementing of vitamin K1 or vitamin K2 [45], indicating vitamin K1 can partially make up for the loss of vitamin K2.

Rather than inserting all missing genes, in a minimal design setup, these vitamins can be supplemented to the medium (indicated in Table 5 with *). To determine any immediate effect on growth or survival rates, medium supplementation of these vitamins was tested, monitoring performance of all recombinant strains under anoxic conditions. In parallel a survival experiment without the

Table 4 Respiratory design of a facultative anaerobic P. putida KT2440. Genes related to cytochrome C

InterPro Entry	PFAM	Name	Abbreviation	Function	Source
IPR003321	PF02335	Cytochrome c552		Cytochrome C oxidase bio generation	PDC

T Trancriptomics, *PDC* Protein Domain Comparisons, *GSM* Genome Scale Modelling [30, 31];

vitamin mix was done. No difference in growth or survival rates was found (Figure S4, Figure S5, Data S10, Data S11).

Lastly within this category, transcriptional regulator DNR was found, its importance in anaerobic respiration described in earlier research [13].

ATP generation

Of the 49 genes of known function required for anaerobic respiration, only one is involved in ATP generation. The protein domain analysis, transcriptomics data and metabolic modelling with iJP962 and iJN1411 all indicated that ATP production remains one of the main bottlenecks to tackle. Earlier work came to the same conclusion and this was tackled by insertion of genes for acetate production [4, 14]. The recombinant strain best performing in previous micro-oxic research included acetate kinase (*ackA)* for ATP generation through acetate production, which was therefore included in the previous design [23] (Table 6). *Pseudomonas putida* KT2440 has its own functional phosphate acetyltransferase (*pta*).

Domains of unknown function

The protein domain analysis additionally included 270 unique protein domains of unknown function occurring in the genomes of facultative anaerobic strains but not in aerobic strains. Based on contextual information, eight were identified as important for anaerobic respiration by their physical co-localisation with genes required for anaerobic respiration. These were therefore included in the design (Table 7). Similarly, 28 protein domains of unknown function were associated with virology factors or immunity and on these ground were excluded from the design. The remaining 234 protein domains of unknown function provided no direct contextual hints and thus cannot not be conclusively excluded from the design.

Discussion

No extended survival under anoxic conditions after acetate kinase expression

Our previous rational design [23] was based on two genome-scale models and genome domain comparison analysis of six facultative anaerobic *Pseudomonas* species compared to six obligatory aerobic *Pseudomonas putida* species. Under micro-oxic conditions, the addition of acetate kinase, dihydroorotate dehydrogenase and class II ribonucleotide triphosphate reductase leads to growth.

In our hands there was however no extended survival under anoxic conditions of the recombinant strains upon introduction of *ackA*. It is extremely challenging to acquire anoxic conditions. Both the medium and the headspace must be treated to completely remove oxygen from the start of the experiment, otherwise oxygen depletion takes up to 12 h. Further, the medium must be prepared with L-cysteine or sodium thioglycollate to actively remove oxygen. Without these precautions, the medium is very easily oxygenated. Small stopper-capped vials are preferred strongly over screw-cap vials, in which oxygen leaks frequently occurred [23]. Addition of the redox-sensitive phenoxazine dye resazurin functions to detect aerobic respiration. Resazurin is generally used as an oxygen indicator as the colour changes from dark purple (high oxygen levels) to pink (low oxygen levels) to transparent (below detectable oxygen levels, as determined by micro-electrode at 0.01 g/l dissolved oxygen [23]). However, resazurin cannot be applied to distinguish between micro-oxic and anoxic conditions.

The lack of improvement in survival rates under anoxic conditions can easily be explained when contemplating the novel design assembled in this research, as numerous essential factors such as an alternative electron acceptor or an anaerobically active cytochrome-C were missing.

Technical design issues

The aim of this research was to determine the requirements for anaerobic respiration, using the industrially interesting workhorse *P. putida* KT2440 as a concept organism. This fundamental question resulted in a design that required 49 genes of known function and 8 genes encoding protein domains of unknown function, resulting in almost 60 genes to be included in the genome. For solely industrial applications, it might be more beneficial to opt for a facultative anaerobic strain from the start. However, this design does offer a fundamental insight in the elaborate change in genotype to achieve an anaerobic lifestyle.

To enable an anaerobic lifestyle, previous designs included the introduction of between 3 and 24 genes in *P. putida* KT2440 genome [4, 13, 14, 23] but our in silico methods suggests that approximately three times more genes are required. Novel methods developed specifically for integration of large operons or multiple genes like yTREX [46] allow incorporation of up to 14 genes at the time in *P. putida*. Albeit an elaborate effort, we consider the inclusion of almost 60 genes in the *P. putida* KT2440 genome to be technically possible, especially since there are constantly new developments in the field of genome engineering.

The GSMs indicated that to produce vitamin B6 and vitamin K2, oxygen is indirectly used. Although anaerobic alternative routes could be implemented, supplementation of these vitamins in the medium significantly reduces the design. Furthermore, while the design is based on gene presence or absence in anaerobic strains

Table 5 Respiratory design of a facultative anaerobic P. putida KT2440. Genes related to pyrimidine and amino acid biosynthesis

InterPro Entry	PFAM	Name	Abbreviation	Function	Source
IPR002751	PF01891	Cobalt uptake substrate-specific transmembrane region	Vitamin B12 [a]	VitB12 Biosynthesis	PDC
IPR006538	PF06213	Cobalamin biosynthesis protein CobT	Vitamin B12 [a]	VitB12 Biosynthesis	PDC
	PF09489	Probable cobalt transporter subunit (CbtB)	Vitamin B12 [a]	VitB12 Biosynthesis	PDC
	PF10531	SLBB domain	Vitamin B12 [a]	Vit B12 uptake	PDC
		adenosylcobalamin	Vitamin B12 [a]	VitB12 Biosynthesis	GSM
	PF02621	Menaquinone biosynthesis	Vitamin K2 [a]	VitK2 biosynthesis	PDC
		Pyridoxal-5-phosphate	Vitamin B6 [a]	Pyridoxal-5-phosphate biosynthesis	GSM
		Dihydroorotate dehydrogenase	PyrK-PyrD B	Pyrimidine production	GSM, [23]
		Ribonucleotide triphosphate reductase type II	NrdD-NrdG	Pyrimidine production	GSM, [23]
PA0527	PF00027	transcriptional regulator	DNR	Transcriptional regulators	Steen et al. 2012
	PF13545				Steen et al. 2012

T Trancriptomics, *PDC* Protein Domain Comparisons, *GSM* Genome Scale Modelling; [a]can be added as vitamin to medium [30, 31];

of the same species, the design could be slimmed down by a futher elimination of proposed genes based on more elaborate analyses of their gene activity in facultative anaerobic species when compared to aerobic species. This would for instance include a selection of nitrogen sources the strain would grow on and would serve to eliminate all nitrogen fixation genes up to that point.

The 57 genes in our design do not consider the 234 genes of unknown function, which complicate the task even further. Without knowing their function, these genes cannot be excluded from the design. At least eight of these were found to be closely associated with genes required for growth in anoxic conditions in literature and/or through their physical co-localisation [31]. The crucial roles that genes of unknown function might play was demonstrated by Hutchison and colleagues [47], who in their attempt to make a minimal bacterial genome, unexpectedly found 149 genes of unknown function to be essential for growth.

Many of the genes found in the design are associated with metal chelation and transport, including many hydrogenases and genes required for vitamin biosynthesis. It should be considered that changes in oxygen availability drastically alters metal bioavailability as extensively reviewed in [48]. Interestingly, we found multiple strict aerobic *Pseudomonas* strains which upon literature research proved able of nitrogen fixation, but not of anaerobic respiration. This might suggest that the evolutionary road between the two first includes nitrogen fixation (or excludes ammonification) before including ammonification (or excluding nitrogen fixation). Either way, experimental validation in anaerobic strains is required to proof which genes required for ammonification or and nitrogen fixation are essential for anaerobic respiration.

The new design compared to previous designs

We predicted that for anaerobic growth both the Nir-Nar and Nor operons are vital. There do exist *Pseudomonas* species that naturally have only one of these operons and are capable of nitrate to nitrite transformation. However, these strains respire nitrate under oxic conditions only, and have been shown to be incapable of growth in anoxic conditions [49, 50]. If *P. putida* KT2440 would be enriched with both the denitrification pathway and the ammonification pathway it could reduce nitrate or nitrite to ammonium, which could then be assimilated to organic compounds, transforming *P. putida* KT2440 in a diazotroph of agronomic importance [51].

The most prevalent anaerobic dissimilatory nitrate respiration regulator DNR is a key transcription factor

Table 6 Respiratory design of facultative anaerobic P. putida KT2440. Genes related to ATP generation to add for a P. putida KT2440 capable of anaerobic respiration

InterPro Entry	PFAM	Name	Abbreviation	Function
	Acetate kinase	AckA	ADP to ATP conversion by acetate production	PDC, GSM, Kampers et al. under review

T Trancriptomics, *PDC* Protein Domain Comparisons, *GSM* Genome Scale Modelling [30, 31];

Table 7 Respiratory design of facultative anaerobic P. putida KT2440. Domains of unknown function to add for a P. putida KT2440 capable of anaerobic respiration

InterPro Entry	PFAM	Name	Abbreviation	Function	Source
	PF09086	**Domain of unknown function**	**DUF1924**		**PDC**
IPR013039	PF07627	Domain of unknown function	DUF1588		PDC
IPR013036	PF07626	Domain of unknown function	DUF1587		PDC
IPR013042	PF07631	Domain of unknown function	DUF1592		PDC
IPR013043	PF07637	Domain of unknown function	DUF1595		PDC
IPR011727	PF09601	Domain of unknown function	DUF2459		PDC
	PF12981	Domain of unknown function	DUF3865		PDC
	PF02026	Domain of unknown function	RyR domain		PDC

T Trancriptomics, *PDC* Protein Domain Comparisons, *GSM* Genome Scale Modelling [30, 31];

obtained from the protein domain comparison. In the facultative anaerobic *E. coli*, knock-out FNR mutants, an ortholog of DNR, were unable to grow by anaerobic respiration under anoxic conditions. By DNA microarray technology it was shown that in *E. coli* 49% of the genes that differ in expression between anoxic and oxic conditions are regulated by FNR [38]. The two-component aerobic respiratory control system (ArcA and ArcB) controls gene transcription in *E. coli* under anoxic conditions. Mutations in this system are known to affect expression of over 30 operons. Most of these are repressed under anoxic conditions, but cytochrome C oxidase and pyruvate formate lyase are activated. In *E. coli*, ArcAB and FNR are deemed essential for anaerobic activation and robustness under micro-oxic conditions [52–55]. To maintain after incorporating the ability of anaerobic respiration, optimal functionality of this strain under oxic conditions, these genes are included in the final design.

We argue that for a lifestyle shift from a strict aerobic lifestyle in *P. putida* KT2440 to an faculatitve anaerobic respiratory lifestyle, all 49 known genes, at least 8 protein domains of unknown function and 3 added vitamins are required. However, increased strain performance under micro-oxic conditions or prolonged survival rates under anoxic conditions already could significantly improve strain robustness in large scale bioreactors with fluctuating oxygen levels. For enhanced performance under micro-oxic conditions, it was demonstrated that increasing ATP production through acetate production already appears to be enough [23]. For prolonged survival rates, however, these key elements include both Nir-Nar and Nor operons for denitrification and ammonification, cytochrome C 552, and external supplementation of the lacking vitamins. This conclusion is supported by previous findings that energy supply and redox balancing are the main bottlenecks in an anaerobic lifestyle [4, 13–16, 23].

Conclusion

Increased ATP generation by insertion of acetate kinase via a plasmid does not lead to prolonged survival rates of *Pseudomonas putida* KT2440 under anoxic conditions. This proves that increased performance under micro-oxic conditions does not guarantee prolonged survival under anoxic conditions. A *P. putida* KT2440 strain capable of anaerobic respiration would require the insertion of at least 57 additional genes into the genome and a supplement of 3 vitamins to the medium. The conversion of a strict aerobic species to a facultative anaerobic lifestyle by anaerobic respiration is a much more elaborate process than was thought before. Especially the function of DUFs and their role in anaerobic respiration must be researched, as it remains unknown how many of these should be added to this design.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12866-020-02058-1.

Additional file 1: Table S1. Bacterial strains and plasmids used in this study .xlsx file with bacterial strains and plasmids listed, including references or sources

Additional file 2: Figure S1. Anoxic survival of P. putida KT2440 transformed strains, grown on De Bont minimal medium with gluconic acid as sole carbon source and kanamycin. The headspace was flushed from oxygen with nitrogen. Survival under anoxic conditions was determined by comparing the number of colony forming units (CFU) over time with the number of CFU at T0. *Escherichia coli* BW25113 was used as positive control and Pseudomonas putida KT2440 with an empty plasmid (pS2213 -) was used as a negative control. Tested strains were Pseudomonas putida KT2440 with acetate kinase (pS2213 ackA) unpassed (p + 0) or passed three consecutive times over oxygen gradients (p + 3), and Pseudomonas putida KT2440 with acetate kinase, dihydroororotate dehydrogenase and ribonucleotide triphosphate reductase type II (pS2213 ackA-(pyrK-pyrD B)-(nrdD-nrdG) unpassed (p + 0) or passed three consecutive times over oxygen gradients (p + 3).

Additional file 3: Figure S2. Survival experiment of P. putida KT2440 under anoxic conditions. The CFU determination of Pseudomonas putida KT2440 with an empty plasmid (pS2213 -), acetate kinase (pS2213 ackA) or acetate kinase, dihydroororotate dehydrogenase and ribonucleotide triphosphate reductase type II (pS2213 ackA-(pyrK-pyrD B)-(nrdD-nrdG) unpassed (p + 0) or passed three consecutive times over oxygen

gradients (p + 3) survival under anoxic conditions. The experiment was repeated independently six times. All figures share the same legend. (A) Experiment 1 (B) Experiment 2 (C) Experiment 3 (D) Experiment 4.

Additional file 4: Figure S3. Transcriptomics of *Pseudomonas aeruginosa* PA01 in 15 different conditions. (A) Heatmap of up (green) or downregulation (red) of all genes per condition. (B) All upregulated genes per condition. (C) All downregulated genes per condition.

Additional file 5: Figure S4. Survival experiment of P. putida KT2440 under anoxic conditions. The CFU determination of Pseudomonas putida KT2440 with an empty plasmid (pS2213 -), acetate kinase (pS2213 ackA) or acetate kinase, dihydroororotate dehydrogenase and ribonucleotide triphosphate reductase type II (pS2213 ackA-(pyrK-pyrD B)-(nrdD-nrdG) unpassed (p + 0) or passed three consecutive times over oxygen gradients (p + 3) with or without vitamin mix.

Additional file 6: Figure S5. Growth experiment of P. putida KT2440 under anoxic conditions. The OD600 determination of Pseudomonas putida KT2440 with an empty plasmid (pS2213 -), acetate kinase (pS2213 ackA) or acetate kinase, dihydroororotate dehydrogenase and ribonucleotide triphosphate reductase type II (pS2213 ackA-(pyrK-pyrD B)-(nrdD-nrdG) unpassed (p + 0) or passed three consecutive times over oxygen gradients (p + 3) with vitamin mix.

Additional file 7: Data S1. With anaerobic cultivation Analysis 1.

Additional file 8: Data S2. With anaerobic cultivation Analysis 2.

Additional file 9: Data S3. With anaerobic cultivation Analysis 3.

Additional file 10: Data S4. With anaerobic cultivation Analysis 4.

Additional file 11: Data S5. With *Pseudomonas* selection database as obtained via GOLDDatabase, including extra information and sources.

Additional file 12: Data S6. With PDC data.

Additional file 13: Data S7. With Transcriptomics data.

Additional file 14: Data S8. With GSMsimulation data.

Additional file 15: Data S9. With random forest data.

Additional file 16: DataS10. With anaerobic cultivation Analysis 5.

Additional file 17: DataS11. With anaerobic cultivation Analysis 6.

Additional file 18. Potential roles of Dau, Nif, Rnf, RhfH operons in anaerobic respiration.

Acknowledgements
We gratefully acknowledge dr. P. I. Nikèl and dr. R. A. Weusthuis for insightful discussions regarding the metabolic insights and anaerobic experimentation in this project.

Authors' contributions
Conceived and designed the experiments: LFCK/JJK/MSD Performed the anaerobic experiments: LFCK Performed the GSM simulations: RGAvH Performed Comparative genomics: JJK/LFCK Performed transcriptomics analysis: MSD/LFCK Protein annotation: LFCK Final design construction: LFCK/AJMS Work supervision: AJMS/PJS Wrote manuscript: LFCK Revised manuscript: MSD/AJMS/JJK/RGAvH/PJS. The authors read and approved the final manuscript.

Funding
We gratefully acknowledge financial support from the European Horizon 2020 projects EmPowerPutida (Project reference: 635536) and IBISBA (Project reference: 730976) and from the Wageningen University IP/OP project. The funders had no role in study design, data collection and analysis, or preparation of the manuscript. All authors have read and approved the manuscript.

Availability of data and materials
All data generated or analysed during this study is included in this published article and its supplementary information files.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Laboratory of Systems and Synthetic Biology, Wageningen University and Research Centre, Stippeneng 4, 6708, WE, Wageningen, The Netherlands. [2]Laboratory of Microbiology, Wageningen University and Research Centre, Stippeneng 4, 6708, WE, Wageningen, The Netherlands.

Received: 26 May 2020 Accepted: 2 December 2020
Published online: 06 January 2021

References

1. Kampers LFC, Volkers RJM, Martins dos Santos VAP. *Pseudomonas putida* kt2440 is hv1 certified, not gras. Microb Biotechnol. 2019;12(5):845. https://doi.org/10.1111/1751-7915.13443 https://onlinelibrary.wiley.com/doi/pdf/10.1111/1751-7915.13443.
2. Belda E, van Heck RGA, Jos'e Lopez-Sanchez M, Cruveiller S, Barbe V, Fraser C, et al. The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis. Environ Microbiol. 2016;18(10):3403–24. https://doi.org/10.1111/1462-2920.13230.
3. Nogales J, Palsson BO, Thiele I. A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: i JN746 as a cell factory. BMC Syst Biol. 2008;2:79. https://doi.org/10.1186/1752-0509-2-79.
4. Sohn SB, Kim TY, Park JM, Lee SY. *In silico* genome-scale metabolic analysis of *Pseudomonas putida* KT2440 for polyhydroxyalkanoate synthesis, degradation of aromatics and anaerobic survival. Biotechnol J. 2010;5(7):739–50. https://doi.org/10.1002/biot.201000124.
5. Oberhardt MA, Puchalka J, Martins dos Santos VAP, Papin JA. Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. PLoS Comput Biol. 2011;7(3):e1001116. https://doi.org/10.1371/journal.pcbi.1001116.
6. Puchalka J, Oberhardt MA, Godinho M, Bielecka A, Regenhardt D, Timmis KN, et al. Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. PLoS Comput Biol. 2008;4(10):e1000210. https://doi.org/10.1371/journal.pcbi.1000210.
7. van Heck RG, Ganter M, dos Santos VAM, Stelling J. Efficient reconstruction of predictive consensus metabolic network models. PLoS Comput Biol. 2016;12(8):e1005085.
8. Clarke PH. The metabolic versatility of pseudomonads. Antonie Van Leeuwenhoek. 1982;48(2):105–30. https://doi.org/10.1007/BF00405197.
9. Bagdasarian M, Lurz R, Ruckert B, Franklin FCH, Bagdasarian MM, Frey J, et al. Specific-purpose plasmid cloning vectors II. Broad host range, high copy number, RSF 1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. Gene. 1981;16(1–3):237–47. https://doi.org/10.1016/0378-1119(81)90080-9.
10. Nikel PI, de Lorenzo V. *Pseudomonas putida* as a functional chassis for industrial biocatalysis: from native biochemistry to trans-metabolism. Metab Eng. 2018;50:142. https://doi.org/10.1016/j.ymben.2018.05.005.
11. Avendano R, Chaves N, Fuentes P, S'anchez E, Jim'enez JI, Chavarr'ıa M. Production of selenium nanoparticles in *Pseudomonas putida* kt2440. Sci Rep. 2016;6:37155.
12. Poblete-Castro I, Becker J, Dohnt K, Santos VM, Wittmann C. Industrial biotechnology of *Pseudomonas putida* and related species. Appl Microbiol Biotechnol. 2012;93(6):2279–90. https://doi.org/10.1007/s00253-012-3928-0.
13. Steen A, Utkur FO, Borrero-de Acuna JM, Bunk B, Roselius L, Buhler B, et al. Construction and characterization of nitrate and nitrite respiring *Pseudomonas putida* KT2440 strains for anoxic biotechnical applications. J Biotechnol. 2013;163(2):155–65. https://doi.org/10.1016/j.jbiotec.2012.09.015.
14. Nikel PI, de Lorenzo V. Engineering an anaerobic metabolic regime in *Pseudomonas putida* KT2440 for the anoxic biodegradation of 1,3-dichloroprop-1-ene. Metab Eng. 2013;15:98–112. https://doi.org/10.1016/j.ymben.2012.09.006.
15. Schmitz S, Nies S, Wierckx N, Blank LM, Rosenbaum MA. Engineering mediator-based electroactivity in the obligate aerobic bacterium *Pseudomonas putida* KT2440. Microbial Physiol Metab. 2015;6:284. https://doi.org/10.3389/fmicb.2015.00284.

16. Lai B, Yu S, Bernhardt PV, Rabaey K, Virdis B, Kromer JO. Anoxic metabolism and biochemical production in *Pseudomonas putida* F1 driven by a bioelectrochemical system. Biotechnol Biofuels. 2016;9:39. https://doi.org/10.1186/s13068-016-0452-y.
17. Wu M, Guina T, Brittnacher M, Nguyen H, Eng J, Miller SI. The *Pseudomonas aeruginosa* proteome during anaerobic growth. J Bacteriol. 2005;187(23):8185–90. https://doi.org/10.1128/JB.187.23.8185-8190.2005.
18. Glasser NR, Kern SE, Newman DK. Phenazine redox cycling enhances anaerobic survival in *Pseudomonas aeruginosa* by facilitating generation of ATP and a proton-motive force. Mol Microbiol. 2014;92(2):399–412. https://doi.org/10.1111/mmi.12566.
19. Eschbach M, Schreiber K, Trunk K, Buer J, Jahn D, Schobert M. Long-term anaerobic survival of the opportunistic pathogen *Pseudomonas aeruginosa* via pyruvate fermentation. J Bacteriol. 2004;186(14):4596–604. https://doi.org/10.1128/JB.186.14.4596-4604.2004.
20. Trunk K, Benkert B, Quack N, Munch R, Scheer M, Garbe J, et al. Anaerobic adaptation in *Pseudomonas aeruginosa*: definition of the Anr and Dnr regulons. Environ Microbiol. 2010;12(6):1719–33. https://doi.org/10.1111/j.1462-2920.2010.02252.x.
21. Hesse C, et al. Genome-based evolutionary history of pseudomonas spp. Environ Microbiol. 2018;20(6):2142–59.
22. Petzold CJ, Chan LJG, Nhan M, Adams PD. Analytics for metabolic engineering. Front Bioengin Biotechnol. 2015;3:135. https://doi.org/10.3389/fbioe.2015.00135.
23. Kampers LFC, van Heck RGA, Donati S, Saccenti E, Volkers RJM, Schaap PJ, et al. *In silico*-guided engineering of *Pseudomonas putida* towards growth under micro-oxic conditions. Microb Cell Factories. 2019;18(1):179. https://doi.org/10.1186/s12934-019-1227-5.
24. Koehorst JJ, Dam JCJ, Heck RGA, Saccenti E, Santos VAPM, Suarez-Diez M, et al. Comparison of 432 *Pseudomonas* strains through integration of genomic, functional, metabolic and expression data. Sci Rep. 2016;6:38699. https://doi.org/10.1038/srep38699.
25. Dotsch A, Schniederjans M, Khaledi A, Hornischer K, Schulz S, Bielecka A, et al. The *Pseudomonas aeruginosa* transcriptional landscape is shaped by environmental heterogeneity and genetic variation. mBio. 2015;6(4):00749–15. https://doi.org/10.1128/mBio.00749-15.
26. Hartmans S, Smits JP, van der Werf MJ, Volkering F, de Bont JAM. Metabolism of styrene oxide and 2-Phenylethanol in the styrene-degrading Xanthobacter strain 124x. Appl Environ Microbiol. 1989;55(11):2850–5.
27. Reddy TB, Thomas AD, Stamatis D, Bertsch J, Isbandi M, Jansson J, et al. The genomes online database (gold) v. 5: a metadata management system based on a four level (meta) genome project classification. Nucleic Acids Res. 2014;43(D1):1099–106.
28. Amid C, Alako BTF, Balavenkataraman Kadhirvelu V, Burdett T, Burgin J, Fan J, et al. The European nucleotide archive in 2019. Nucleic Acids Res. 2019;48(D1):70–6. https://doi.org/10.1093/nar/gkz1063.
29. Koehorst JJ, van Dam JCJ, Saccenti E, Martins dos Santos VAP, Suarez-Diez M, Schaap PJ. SAPP: functional genome annotation and analysis through a semantic framework using FAIR principles. Bioinformatics. 2018;34(8):1401–3. https://doi.org/10.1093/bioinformatics/btx767.
30. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119. https://doi.org/10.1186/1471-2105-11-119.
31. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–40. https://doi.org/10.1093/bioinformatics/btu031.
32. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The pfam protein families database in 2019. Nucleic Acids Res. 2019;47(D1):427–32.
33. Koehorst JJ, Saccenti E, Schaap PJ, dos Santos VAM, Suarez-Diez M. Protein domain architectures provide a fast, efficient and scalable alternative to sequence-based methods for comparative functional genomics. F1000Research. 2016;5:1.
34. Anders S, Huber W. Differential expression of rna-seq data at the gene level–the deseq package. Heidelberg: European Molecular Biology Laboratory (EMBL); 2012.
35. Nogales J, Gudmundsson S, Duque E, Ramos JL, Palsson BO. Expanding the computable Reactome in *Pseudomonas putida* reveals metabolic cycles providing robustness. bioRxiv. 2017;139121:1. https://doi.org/10.1101/139121.
36. Nikel PI, et al. Pseudomonas putida KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. J Biol Chem. 2015;290(43):25920–32.
37. Carmona M, Zamarro MT, Bl'azquez B, Durante-Rodr'ıguez G, Ju'arez JF, Valderrama JA, et al. Anaerobic catabolism of aromatic compounds: a genetic and genomic view. Microbiol Mol Biol Rev. 2009;73(1):71–133. https://doi.org/10.1128/MMBR.00021-08.
38. Kovacs A, Rakhely G, Balogh J, Maroti G, Fulop A, Kovacs K. Anaerobic regulation of hydrogenase transcription in different bacteria; 2005.
39. Peters JW, Schut GJ, Boyd ES, Mulder DW, Shepard EM, Broderick JB, et al. [FeFe] and [NiFe]-hydrogenase diversity, mechanism, and maturation. Biochimica et Biophys Acta (BBA) Mol Cell Res. 2015;1853(6):1350–69. https://doi.org/10.1016/j.bbamcr.2014.11.021.
40. Crespo A, Blanco-Cabra N, Torrents E. Aerobic vitamin B12 biosynthesis is essential for *Pseudomonas aeruginosa* class II Ribonucleotide Reductase activity during planktonic and biofilm growth. Front Microbiol. 2018;9:1. https://doi.org/10.3389/fmicb.2018.00986.
41. Fang H, Kang J, Zhang D. Microbial production of vitamin B12: a review and future perspectives. Microb Cell Factories. 2017;16:1. https://doi.org/10.1186/s12934-017-0631-y.
42. Moore SJ, Lawrence AD, Biedendieck R, Deery E, Frank S, Howard MJ, et al. Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin b12). Proc Natl Acad Sci. 2013;110(37):14906–11.
43. Roth JR, Lawrence JG, Rubenfield M, Kieffer-Higgins S, Church GM. Characterization of the cobalamin (vitamin B12) biosynthetic genes of salmonella typhimurium. J Bacteriol. 1993;175(11):3303–16.
44. Richts B, Rosenberg J, Commichau FM. A survey of Pyridoxal 5-phosphate-dependent proteins in the gram-positive model bacterium Bacillus subtilis. Front Mol Biosci. 2019;10:6. https://doi.org/10.3389/fmolb.2019.00032.
45. Wissenbach U, Kroger A, Unden G. The specific functions of menaquinone and demethylmenaquinone in anaerobic respiration with fumarate, dimethylsulfoxide, trimethylamine N-oxide and nitrate by *Escherichia coli*. Arch Microbiol. 1990;154(1):60–6.
46. Domrose A, Weihmann R, Thies S, Jaeger K-E, Drepper T, Loeschcke A. Rapid generation of recombinant *Pseudomonas putida* secondary metabolite producers using ytrex. Synthetic Syst Biotechnol. 2017;2(4):310–9. https://doi.org/10.1016/j.synbio.2017.11.001.
47. Hutchison CA, Chuang R-Y, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman MH, et al. Design and synthesis of a minimal bacterial genome. Science. 2016;351:6280.
48. Hong Enriquez RP, Do TN. Bioavailability of metal ions and evolutionary adaptation. Life. 2012;2(4):274–85.
49. Hatayama K, Kawai S, Shoun H, Ueda Y, Nakamura A. *Pseudomonas* azotifigens sp. nov., a novel nitrogen-fixing bacterium isolated from a compost pile. Int J Syst Evol Microbiol. 2005;55(4):1539–44.
50. Yumoto I, Yamazaki K, Hishinuma M, Nodasaka Y, Suemori A, Nakajima K, et al. *Pseudomonas* alcaliphila sp. nov., a novel facultatively psychrophilic alkaliphile isolated from seawater. Int J Syst Evol Microbiol. 2001;51(2):349–55.
51. Pitcher RS, Watmough NJ. The bacterial cytochrome cbb3 oxidases. Biochimica et Biophysica Acta (BBA) Bioenergetics. 2004;1655:388–99. https://doi.org/10.1016/j.bbabio.2003.09.017.
52. Nizam SA, Shimizu K. Effects of arca and arcb genes knockout on the metabolism in *Escherichia coli* under anaerobic and microaerobic conditions. Biochem Eng J. 2008;42(3):229–36.
53. Nikel PI, et al. Metabolic flux analysis of Escherichia coli creB and arcA mutants reveals shared control of carbon catabolism under microaerobic growth conditions. J Bacteriol. 2009;191(17):5538–48.
54. Shalel-Levanon S, San K-Y, Bennett GN. Effect of ArcA and FNR on the expression of genes related to the oxygen regulation and the glycolysis pathway in Escherichia coli under microaerobic growth conditions. Biotechnol Bioeng. 2005;92(2):147–59.
55. Zhu J, et al. Effect of the global redox sensing/regulation networks on Escherichia coli and metabolic flux distribution based on C-13 labeling experiments. Metab Eng. 2006;8(6):619–27.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.