Single-cell genomics reveals metabolic strategies for microbial growth and survival in an oligotrophic aquifer

Michael J. Wilkins, David W. Kennedy, Cindy J. Castelle, Erin K. Field, Ramunas Stepanauskas, James K. Fredrickson and Allan E. Konopka

Bacteria from the genus *Pedobacter* are a major component of microbial assemblages at Hanford Site (a largely decommissioned nuclear production complex) in eastern Washington state, USA, and have been shown to change significantly in abundance in response to the subsurface intrusion of Columbia River water. Here we employed single-cell genomics techniques to shed light on the physiological niche of these microorganisms. Analysis of four *Pedobacter* single amplified genomes (SAGs) from Hanford Site sediments revealed a chemoheterotrophic lifestyle, with the potential to exist under both aerobic and microaerophilic conditions via expression of both aa3-type and cbb3-type cytochrome c oxidases. These SAGs encoded a wide range of both intra- and extracellular carbohydrate-active enzymes, potentially enabling the degradation of recalcitrant substrates such as xylan and chitin, and the utilization of more labile sugars such as mannose and fucose. Coupled to these enzymes, a diversity of transporters and sugar-binding molecules were involved in the uptake of carbon from the extracellular local environment. The SAGs were enriched in TonB-dependent receptors, which play a key role in uptake of substrates resulting from degradation of recalcitrant carbon. Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-Cas mechanisms for resisting viral infections were identified in all SAGs. These data demonstrate the potential mechanisms utilized for persistence by heterotrophic microorganisms in a carbon-limited aquifer, and hint at potential linkages between observed *Pedobacter* abundance shifts within the 300 Area (in the south-eastern corner of the site) subsurface and biogeochemical shifts associated with Columbia River water intrusion.

INTRODUCTION

Within shallow groundwater aquifers, indigenous microbial populations catalyse a broad range of biogeochemical cycles that have implications for domestic, industrial and agricultural water usage. Biologically catalysed processes include organic contaminant degradation, the release of certain toxic metals such as arsenic and the immobilization of some soluble metals such as uranium (Anderson et al., 2003; Islam et al., 2004; Maymó-Gatell et al., 1997; Williams et al., 2011). Relative to surface environments, these subsurface systems are perceived to be relatively stable and undergo less frequent perturbation. However, despite these general observations, dynamic temporal and spatial shifts within microbial populations have been reported in certain aquifers (Ginige et al., 2013; Lin et al., 2012b; Zhou et al., 2012a); the factors responsible for this variability are currently poorly understood. The majority of analyses of microbial populations have focused on the identification of microbial taxa, via a marker gene such as 16S rRNA. In addition, the ability to generate and analyse shotgun metagenomic data has allowed metabolic and physiological traits to be inferred (Hemme et al., 2010; Wrighton et al., 2012). These datasets have proved critical for understanding (1) how microbial populations interact with their environment, and (2) the factors responsible for driving shifts in microbial population structures over space and time. More recently, single-cell genomics has become a widely applicable research tool, enabling direct rather than

Correspondence
Michael J. Wilkins
wilkins.231@osu.edu

Received 7 October 2013
Accepted 9 December 2013
inferred examination of the metabolic potential of uncultured microbial cells (Stepanauskas, 2012).

The Hanford Site in eastern Washington state, USA, was operated during the Second World War and Cold War for plutonium production, as part of the US nuclear weapons programme. The site is now mostly decommissioned, and work at the 300 Area (in the south-eastern corner of the site) (Fig. 1) has focused on the role of biogeochemical cycling in contaminant fate and transport (Peretyazhko et al., 2012; Um et al., 2010; Zachara et al., 2013). The unconfined aquifer beneath the 300 Area is an oligotrophic environment characterized by oxic conditions in the upper Hanford Formation (O₂ concentrations are frequently measured at 10 mg l⁻¹) and limited nutrient concentrations (Lin et al., 2012a). Dissolved organic carbon concentrations typically average around 1 mg l⁻¹. Despite these geochemical conditions, biomass concentrations of 10⁴ to 10⁵ cells ml⁻¹ have been measured (J. Stegen, personal communication), with cell doubling times as fast as 2 h, similar to rates reported in other aquifer systems (Ayuso et al., 2010). The aquifer is mainly present within highly transmissive Hanford Formation sediments, with the hydraulic gradient directly affected by the height stage of the nearby Columbia River (Bjornstad et al., 2009). For the majority of the year, groundwater discharges into the Columbia River; however, under high river stage conditions associated with mountain snowpack runoff in spring, flow direction is reversed and Columbia River water intrudes into the subsurface and groundwater elevation increases. Using a 16S rRNA gene pyrosequencing approach, Lin et al. (2012b) tracked planktonic microbial communities within Hanford Formation groundwater over a nine-month sampling period. Dynamic shifts in microbial community structure were reported, often associated with seasonal Columbia River water intrusion (Lin et al., 2012b). However, the absence of functional information for specific taxa prohibited a deeper understanding of the physical or biogeochemical controls that led to these temporal dynamics.

Within the data collected by Lin et al. (2012b), the relative abundance of Pedobacter 16S rRNA gene sequences ranged from ~0 % to ~14 %, suggesting that this genus constitutes a major component of Hanford sediment microbial communities (Fig. 2). Despite the isolation of more than 40
Pedobacter species from a diverse range of soil and sediment environments (An et al., 2009; Gordon et al., 2009; Hoang et al., 2013; Jeon et al., 2009; Jung et al., 2012; Kwon et al., 2011; Luo et al., 2010; Margesin et al., 2003; Margesin & Zhang, 2013; Muurholm et al., 2007; Oh et al., 2013; Roh et al., 2008; Steyn et al., 1998), little is known about the ecophysiology of these species. All Pedobacter species grow under aerobic, heterotrophic conditions, and have been shown to utilize a wide range of carbon compounds in the laboratory, including xylose, mannose, ribose and glucose. Interestingly, many of these species show ability to grow over a wide range of temperatures, from 1 to 37 °C (Margesin et al., 2003; Margesin & Zhang, 2013; Zhou et al., 2012b). Common genus characteristics include the production of the enzymes catalase and alkaline phosphatase. Building on these studies, we hypothesized that Pedobacter within the Hanford subsurface would also probably play a key role in carbon cycling. Given the oligotrophic nature of the 300 Area aquifer, we anticipated that these organisms might employ a range of strategies to maximize available substrates. Finally, a better understanding of the physiology and metabolism of these species would be used to better develop hypotheses linking shifts in Pedobacter abundances with geochemical trends associated with Columbia River water intrusion in the Hanford subsurface.

METHODS

Cell selection. Hanford aquifer sediments recovered from a depth of 30 ft (9.1 m) in well 2-25 were mixed with a synthetic Hanford groundwater (Lin et al., 2012a) containing 1% pyrophosphate to aid cell desorption from mineral surfaces. Sediment was incubated in this material at room temperature under gently shaking conditions for 2 h. Samples of the liquid phase were removed, mixed with glycerol to a final concentration of 20%, and flash frozen in liquid N2 before being shipped on dry ice to the Single Cell Genomics Center at Bigelow Laboratory for Ocean Sciences.

FACS. The sample was vortexed vigorously for 30 s and centrifuged at 2000 r.p.m. 270 g for 1 min. The supernatant was diluted 1:10 with synthetic groundwater, stained with 5 mM (final concentration) SYTO-9 (Invitrogen) for 10–30 min and pre-screened through a 40 μm mesh-size cell strainer (BD). Microbial cells were sorted with a MoFlo (Beckman Coulter) flow cytometer equipped with a CyClone robotic arm for droplet deposition into 384-well plates. The cytometer was triggered on side scatter. Sorting gate was delineated using side scatter, green fluorescence (SYTO-9) and red fluorescence signals. The ‘single 1 drop’ mode was used for maximal sort purity, which ensures the absence of non-target particles within the target cell drop and the adjacent drops. Under these sorting conditions, sorted drops contain a few tens of picolitres of sample surrounding the target cell (Sieracki et al., 2005), so non-target DNA is very low or absent. The accuracy of 10 μm fluorescent bead deposition into the 384-well plates was verified by microscopically examining the presence of beads in the plate wells. Of the two to three
plates examined each sort day, <2% of wells were found not to contain a bead and only <0.5% of wells were found to contain more than one bead, indicating very high purity of single cells. In addition, we verified the lack of DNA contamination in the sheath fluid and in sheath fluid lines by performing real-time multiple displacement amplification with the processed sheath fluid as the template.

**Multiple displacement amplification (MDA).** Bacterial cells were deposited into 384-well plates containing 0.6 μl per well of TE buffer. Plates were stored at −80°C until further processing. The cells were lysed and their DNA was denatured using cold KOH (Raghunathan et al., 2005). Genomic DNA from the lysed cells was amplified using MDA (Dean et al., 2002) in a final volume of 10 μl. The MDA reactions contained 2 U RepliHi polymerase μl⁻¹ (Epicentre), 1 × reaction buffer (Epicentre), 0.4 mM each dNTP (Epicentre), 2 mM DTT (Epicentre), 50 mM random hexamers with the two 3’-terminal nucleotide bonds phosphorothioated (IDT) and 1 μM SYTO-9 (Invitrogen) (all final concentration). The MDA reactions were run at 30°C for 12–16 h, and then inactivated by 15 min incubation at 65°C. The amplified genomic DNA was stored at −80°C until further processing. We refer to the MDA products originating from individual cells as SAGs. The instruments and the reagents were decontaminated for DNA prior to sorting and MDA setup, as previously described (Stepanauskas & Sieracki, 2007; Woyke et al., 2011). Cell sorting and MDA setup were performed in a high-efficiency particulate absorption-filtered environment. As a quality control, the kinetics of each MDA reaction was monitored by measuring the SYTO-9 fluorescence using FLUOstar Omega (BMG). The critical point (Cp) was determined for each MDA reaction as the time required to produce half of the maximal fluorescence. The Cp is inversely correlated with the amount of DNA template (Zhang et al., 2006). Of the 384 wells, 315 were dedicated for single cells, 66 were used as negative controls (no droplet deposition) and three received 10 cells each (positive controls). The Cp values were significantly lower in one-cell wells than in zero-cell wells (P<0.05; Wilcoxon twosample test) in each microplate.

16S rRNA gene amplification. The MDA products were diluted 50-fold in sterile TE buffer. Then, 0.5 μl aliquots of the dilute MDA products served as templates in 5 μl real-time PCR screens targeting bacterial subunit rRNA genes using primers 27F and 907R (Casamayor et al., 2000; Lane, 1991). Forward (5’-GTAACGGGAC- GCCCAGT-3’) or reverse (5’-CAGGAAACAGCTATGACC-3’) 16S rRNA gene primers were used in all PCR reactions to obtain the 5’ end of each primer pair to aid direct sequencing of the PCR products. All PCRs were performed using LightCycler 480 SYBR Green I Master mix (Roche) in a LightCycler 480 II real-time thermal cycler (Roche). The real-time PCR kinetics and the amplicon melting curves served as proxies detecting successful target gene amplification. New, 20 μl PCRs were set up for the PCR-positive SAGs and the amplicons were sequenced from both ends using M13 targets and Sanger technology by Beckman Coulter Genomics. Eleven of 17 PCR-positive SAGs were sequenced by RAST server (Aziz et al., 2008). Protein localization information was determined using the PSORTb 3.0 software (http://www.psort.org/psortb/) (Yu et al., 2010), while protein family information was obtained via Pfam searches (http://pfam.sanger.ac.uk/) (Punta et al., 2012). Genome size was estimated using single-copy gene lists published by Raes et al. (2007). Orthologous proteins in other sequenced Pedobacter species were inferred using the SPOCS tool (Curtis et al., 2013). Clustered regularly interspaced short palindromic repeat (CRISPR)-Cas sequences were identified using the CRISPRFinder tool (http://crispr.u-psud.fr/) (Grissa et al., 2007). 3D protein structure predictions were generated by SWISS-MODEL based on protein alignment and secondary structure prediction (Schwede et al., 2003). SWISS-MODEL is an automated protein homology-modelling server (Bordoli et al., 2009). The alignment mode was utilized for a first approach based on a user-defined target–template alignment. Conservation of key catalytic residues and the secondary structure for each model were manually inspected. The four Hanford *Pedobacter* SAGs are publicly available at the US Department of Energy Integrated Microbial Genomes (IMG) site: http://img.jgi.doe.gov/.

**RESULTS AND DISCUSSION**

We analysed four single amplified genomes (SAGs) obtained from Hanford Site sediment: SCGC AC-286-M16, SCGC AC-286-M17, SCGC AC-286-P17 and SCGC AC-288-E13 (hereafter M16, M17, P17 and E13). Phylogenetic analysis of 16S rRNA gene sequences from the Hanford SAGs placed these organisms within the genus *Pedobacter*. Assembly size of these SAGs ranged from 0.8 to 2.4 Mb (Table S1). Genome completeness was estimated using conserved single copy gene identifications (Raes et al., 2007), suggesting 20–75% genome recovery. The SAGs exhibited 97.4–99.0% similarity to each other at the 16S rRNA gene sequence level. Linking to common characteristics of cultured *Pedobacter* species, genes encoding catalase were detected in two SAGs, but no evidence of alkaline phosphatase was identified in the genomic data.
Results for the utilization of carbon compounds are presented below.

**Carbohydrate utilization**

Genomic reconstructions indicated that the Hanford *Pedobacter* SAGs are chemoheterotrophic, obligately aerobic bacteria, with the ability to degrade and oxidize a range of organic carbon compounds, and survive under both oxic and microaerophilic conditions. To survive in the relatively oligotrophic Hanford environment, the *Pedobacter* SAGs encode a wide range of mechanisms for degrading both complex and simple carbon substrates, and transporting these compounds into the cell. Carbohydrate-active enzymes accounted for approximately 5% of the ORFs per SAG; a similar percentage is reported in the sequenced *Pedobacter* strains *Pedobacter saltans* DSM 12145 and *Pedobacter* sp. BAL39. Nearly half of these enzymes were classified within glycoside hydrolase (GH) families, and had functions associated with the degradation of polymers (xylan and chitin), as well as the metabolism of monosaccharides such as rhamnose, mannose and fucose (Fig. 3, Table 1). Other carbohydrate-active enzymes contained carbohydrate-binding modules, while others were classified within the polysaccharide lyase family. Protein localization inferences suggested that a subset of these enzymes are secreted into the extracellular environment where they can catalyse the depolymerization of particulate or mineral-adsorbed organic material. This pool of secreted enzymes may be larger; for each SAG, ~30–60% of enzymes with unknown cellular localizations (Fig. 3) contained a signal peptide, suggesting a role in secretion from the cell.

Following the breakdown of more recalcitrant carbon substrates, *Pedobacter* micro-organisms must transport labile carbohydrates into the cell for metabolism. This may be achieved via a series of mechanisms that enable efficient carbon uptake in the oligotrophic Hanford environment. The *Pedobacter* SAGs contain large numbers of ABC-transporter systems, permease proteins, and TonB-dependent receptor (TBDR) proteins. While TBDRs have traditionally been implicated in Fe(III) and vitamin B12 uptake (Chimento et al., 2005; Postle & Kadner, 2003), recent work has elucidated a role for TBDRs in the transport of sugars across the outer membrane of Gram-negative bacteria (Blanvillain et al., 2007; Neugebauer et al., 2005). Blanvillain et al. (2007) identified a small but phylogenetically diverse group of bacteria that over-represented TNBRs within their genomes. As a common characteristic of this group, all shared the ability to degrade complex carbon sources. The *Pedobacter* SAGs fall within this group, with 18 TNBRs (proteins containing Pfam domains PF07715 and PF00593) encoded within the most complete M17 SAG genome. The TBDR/Mbp genome size ratio of 7.5 places SAG M17 clearly within the ‘over-represented’ group (Blanvillain et al., 2007). For comparison, across the microbial groups studied by Blanvillain et al. (2007) 27% of bacteria had no TBDR proteins, and 43% possessed between one and 13 TBDRs. Interestingly, over-representation of TNBRs is also observed in two additional sequenced *Pedobacter* taxa, *Pedobacter* sp. BAL39 and *P. saltans*, suggesting that this trend may apply to *Pedobacter* generally. Although specific substrates for many of these transport systems cannot be elucidated from genomic sequence alone, genes encoding multiple periplasmic maltose/maltodextrin binding proteins were identified, together with fucose permease proteins that were predicted to localize to the cytoplasmic membrane.

All the enzymes necessary for the degradation and utilization of xylan were detected within the SAGs (Fig. 4), including endo-1,4-xylanases for the breakdown of xylan into xylose subunits, and xylose isomerase for the conversion of xylose into xylulose. An encoded extracellular solute binding protein has ~30% amino acid similarity to the xylo-oligosaccharide binding XynE protein in *Geobacillus stearothermophilus* (Shulami et al., 2007), suggesting a role for this protein in xylan/xylose uptake into the cell. Xylulose generated via these reactions is subsequently metabolized via the pentose phosphate and Embden–Meyerhof–Parnas pathways. Xylan degrading potential was also detected in *P. saltans* genomic data, but not *Pedobacter* sp. BAL39.

Protein orthologue searches across multiple *Pedobacter* genomes were performed to identify patterns and distributions of carbohydrate-active enzymes; results indicated that only five GH enzymes were unique to the Hanford SAGs,

![Fig. 3. Distribution of carbohydrate-active enzymes by enzyme class (a) and predicted cellular localization (b) across the four SAGs. PL, polysaccharide lyase; CBM, carbohydrate binding module; CE, carbohydrate esterase; GT, glycosyltransferase; GH, glycoside hydrolase.](image-url)
while 23 different kinds of GH enzymes found in SAG genomes had orthologues in two other sequenced Pedobacter taxa, *P. saltans* and *Pedobacter* sp. BAL39 (Table S2). These species were isolated from soil (Steyn et al., 1998) and seawater (Pinhassi et al., 1997), respectively, and further support the inference that the capacity for complex carbon and sugar degradation is a conserved trait across these species. In this study, however, the relatively high number of both carbohydrate-active enzymes and mechanisms for carbon substrate uptake encoded within the Hanford SAGs further supports the hypothesis that these mechanisms enable heterotrophic bacteria to thrive within the oligotrophic Hanford subsurface.

Orthologue searches also revealed a large number of hypothetical genes unique to the Hanford SAGs (Table S3); of 549 orthologous genes unique to the SAGs (i.e. present across two or more SAGs, but not the two sequenced *Pedobacter* species), 400 (~72%) were annotated as hypothetical genes. These trends were further pronounced within patterns of genes present in only one genome; of 374 non-orthologous genes in Hanford SAG M17, 232 (~62%) were annotated as hypothetical genes. While these suites of genes may play key roles in shaping ecology and physiology within these species, the role for this ‘biological dark matter’ requires detailed experimental analyses (Kolker et al., 2005).

### Defence mechanisms

Genes encoding orthologous CRISPR-associated Cas proteins (Cas1, Cas2 and Cas9) were detected across three of the SAGs (M17, M16 and P17). The presence of these defence mechanisms suggests a potential strategy for combating environmental disturbances or predation by other microorganisms in the oligotrophic subsurface environment of Hanford.

---

**Table 1. Carbohydrate-active enzymes encoded across four Pedobacter SAGs from the Hanford 300 Area subsurface**

<table>
<thead>
<tr>
<th>Defined activity</th>
<th>Pfam domain</th>
<th>Carbohydrate Active enZYmes (CAZY) family</th>
<th>Hanford SAG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>M17</td>
</tr>
<tr>
<td>Cellulose degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-Glucosidase</td>
<td>PF00933</td>
<td>GH3</td>
<td>1</td>
</tr>
<tr>
<td>Cellulase</td>
<td>PF00150</td>
<td>GH5</td>
<td>2</td>
</tr>
<tr>
<td>Endo-1,6-β-glucanase</td>
<td>PF02055</td>
<td>GH30</td>
<td>2</td>
</tr>
<tr>
<td>Debranching enzymes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,4-α-Glucan branching enzyme</td>
<td>PF00128</td>
<td>GH13</td>
<td>1</td>
</tr>
<tr>
<td>α-Glucosidase</td>
<td>PF01055, PF01261, PF04909</td>
<td>GH31, GH97</td>
<td>2</td>
</tr>
<tr>
<td>α-1-Arabinofuranosidase</td>
<td>PF04616</td>
<td>GH43</td>
<td>2</td>
</tr>
<tr>
<td>α-1-Rhamnosidase</td>
<td>PF05592</td>
<td>GH78</td>
<td>3</td>
</tr>
<tr>
<td>Endohemicellulases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endo-1,4-β-xylanase</td>
<td>PF00331, PF04616</td>
<td>GH10, GH43</td>
<td>1</td>
</tr>
<tr>
<td>Amylolytic enzymes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-Amylase</td>
<td>PF00128</td>
<td>GH13, GH70</td>
<td>1</td>
</tr>
<tr>
<td>Amylomaltase</td>
<td>PF02446, PF00128</td>
<td>GH13, GH77</td>
<td>1</td>
</tr>
<tr>
<td>Glucoamylase</td>
<td>PF00723</td>
<td>GH15</td>
<td>1</td>
</tr>
<tr>
<td>Oligosaccharide-degrading enzymes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-1,2-Mannosidase</td>
<td>PF07971</td>
<td>GH92</td>
<td>4</td>
</tr>
<tr>
<td>α-1,6-Mannanase</td>
<td>PF03663</td>
<td>GH76</td>
<td>3</td>
</tr>
<tr>
<td>α-1-Fucosidase</td>
<td>PF00107, PF01120</td>
<td>GH29</td>
<td>1</td>
</tr>
<tr>
<td>β-Galactosidase</td>
<td>PF00703, PF01301, PF02836, PF02929</td>
<td>GH2, GH35</td>
<td>5</td>
</tr>
<tr>
<td>β-Glucanase</td>
<td>PF00722</td>
<td>GH16</td>
<td>0</td>
</tr>
<tr>
<td>β-Hexosaminidase</td>
<td>PF00728, PF00933</td>
<td>GH3, GH20</td>
<td>1</td>
</tr>
<tr>
<td>β-Mannosidase</td>
<td>PF00703</td>
<td>GH2</td>
<td>1</td>
</tr>
<tr>
<td>β-Xylosidase</td>
<td>PF04616</td>
<td>GH43</td>
<td>0</td>
</tr>
<tr>
<td>Glucuronyl hydrolase</td>
<td>PF05147</td>
<td>GH88</td>
<td>0</td>
</tr>
<tr>
<td>Other glycoside hydrolases</td>
<td>PF02055</td>
<td>GH30</td>
<td>0</td>
</tr>
<tr>
<td>Predicted glycoside hydrolase</td>
<td>PF00759, PF04041, PF06439, PF06441, PF07944</td>
<td>GH9, GH125, GH127, GH130</td>
<td>4</td>
</tr>
<tr>
<td>Other degrading enzymes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-Lactamase</td>
<td>PF00144</td>
<td>GH3</td>
<td>4</td>
</tr>
<tr>
<td>Chitinase</td>
<td>PF00704</td>
<td>GH18</td>
<td>1</td>
</tr>
<tr>
<td>Glucosaminidase</td>
<td>PF00144, PF01476</td>
<td>GH3, GH73</td>
<td>1</td>
</tr>
<tr>
<td>Malto-oligosyltrehalose synthase</td>
<td>PF00128</td>
<td>GH13</td>
<td>1</td>
</tr>
<tr>
<td>Malto-oligosyltrehalose trehalohydrolase</td>
<td>PF00128</td>
<td>GH13</td>
<td>1</td>
</tr>
<tr>
<td>Maltose phosphorylase</td>
<td>PF03632</td>
<td>GH65</td>
<td>0</td>
</tr>
</tbody>
</table>

http://mic.sgmjournals.org
P. saltans and Pedobacter sp. BAL39, indicating the development of viral defence mechanisms for these species within the Hanford aquifer. From the operon structure in all three SAGs (cas9–cas1–cas2) we can infer that these Cas proteins form a type II-A or II-B system (Makarova et al., 2011), characterized by the presence of Cas9, which generates CRISPR RNA and cleaves target DNA. CRISPR regions were identified in SAG contigs, with numbers of spacers ranging from four to 47. Across the three SAGs, 70 different spacer sequences were identified (Table S4); only four were shared across all three SAGs, while 40 were shared between SAGs M16 and E13. SAG M17 contained 20 unique sequences (Fig. 5). Of these shared spacer sequences between M16 and E13, the majority are present in a conserved order across the two SAGs. Single nucleotide polymorphisms are responsible for unique spacer sequences in M16 and E13 (Fig. 5, Table S4). From an evolutionary perspective, the pattern and ordering of spacer distribution suggest that SAGs M16 and E13 have a far more recent common ancestor than M17; the high number of unique spacers in M17 indicates that the lineage for this SAG incorporated mostly different foreign DNA into genomic material, relative to the other SAGs considered in this study. From an environmental perspective, the presence and distribution of these genes in the Hanford single cells suggest that viral infection is an active process within the aquifer, and may be a factor controlling microbiobial population dynamics.

Central metabolism

Following the uptake of carbon substrates into the cell, central metabolism in the Hanford SAGs is inferred to proceed through a complete tricarboxylic acid cycle, with acetyl-CoA generated from pyruvate via a pyruvate dehydrogenase complex, or from acetate via the combined action of acetate kinase and phosphate acetyltransferase. The presence of a 14-subunit NADH dehydrogenase complex, succinate dehydrogenase, F0F1-type ATP synthase and cytochrome c oxidase is evidence of an electron transport chain for generation of ATP via oxidative phosphorylation. Present in all the SAGs studied here are copies of lactate dehydrogenase (EC 1.1.1.28), which catalyses the reversible reaction between pyruvate and lactate. The presence of this gene offers two metabolic possibilities, the first being that lactate is imported into the cell from the local environment, and converted to pyruvate for biosynthetic

![Fig. 4. Inferred mechanism for extracellular xylan degradation and subsequent uptake in Hanford Pedobacter SAGs.](image-url)

Fig. 4. Inferred mechanism for extracellular xylan degradation and subsequent uptake in Hanford Pedobacter SAGs. (1) Activity of extracellular xylanase. (2) Binding of xylose subunits by protein with similarity to xylo-oligosaccharide binding protein XynE, and transport into cytoplasm via ABC transporter. (3) Activity of α-N-arabinofuranosidase and xylose isomerase. (4) Phosphorylation of d-xylulose by xylulose kinase. (5) Metabolism of d-xylulose by pentose phosphate and Embden–Meyerhof–Parnas pathways.

![Fig. 5. CRISPR spacer distribution across three Hanford Pedobacter SAGs.](image-url)

Fig. 5. CRISPR spacer distribution across three Hanford Pedobacter SAGs. Values refer to individual spacer sequences present in each genome, and illustrate the high similarity in spacer composition between E13 and P16, and the dissimilarity of spacer sequences in M17.
Oxygen reduction

Dissolved O₂ concentrations in 300 Area Hanford Formation groundwater are typically at saturation (~10 mg l⁻¹). Despite this, the heterogeneous nature of the Hanford subsurface can result in the development of diffusion-dominated microenvironments (Nguyen et al., 2012), where microaerophilic conditions may exist. Genomic reconstructions of Pedobacter SAGs suggest that the organisms are unable to grow anaerobically, consistent with their detection only from the oxic Hanford Formation. However, genes encoding both aa₃-type and cb₀₂₄₅₋type cytochrome c oxidases were identified across the SAGs, and two additional sequenced Pedobacter taxa, P. saltans and Pedobacter sp. BAL39. The aa₃-type cytochrome c oxidases are typically found in aerobic bacteria, and have an O₂ substrate affinity (Kₘ) approximately six- to eightfold higher than the cb₀₂₄₅₋type oxidases (Morris & Schmidt, 2013). The cb₀₂₄₅₋type oxidases have therefore been implicated in aerobic respiration under microaerophilic conditions; indeed, an O₂ Kₘ value of 7 nM was calculated for a cb₀₂₄₅₋ oxidase protein in *Bradyrhizobium japonicum* (Preigis et al., 1996). Multiple lines of evidence were used to identify the aa₃-type cytochrome c oxidase; protein modelling of subunit I identified the six histidines (H67, H260, H309, H310, H395 and H397) involved in binding the haem-copper binuclear centre, while the detection of a binuclear copper centre (CuA) in subunit II suggests that this oxygen reductase is a cytochrome c oxidase rather than a quinol oxidase (as this centre is absent in quinol oxidases) (Fig. S1). Further analysis of subunit II was unable to identify a c-type haem motif as retrieved in *aa₃*-type oxidases. *P. saltans* also encodes an alternative complex III (ACIII) that functionally replaces the bc₁ complex (complex III in the oxidative phosphorylation chain) (Pereira et al., 2007). As observed in *Rhodothermus marinus* (Riegofo et al., 2010), ACIII genes in *P. saltans* cluster with the low oxygen affinity *aa₃*-type cytochrome c oxidase and thus probably form a functional association under high oxygen concentrations.

The cb₀₂₄₅₋-type oxidase was identified via analysis of the CcoN, CcoO and CcoP subunits required for catalytic activity. The CcoN subunit corresponds to the catalytic subunit homologous to the subunit I of the *aa₃* type oxidase, while the CcoO and CcoP subunits are membrane-bound monomeric (or dimeric) haem cytochromes c. The CcoN and CcoO subunits were fused into one subunit in the Hanford SAGs; similar observations have been reported in members of the Bacteroides, Planctomycetes and Verrucomicrobia (Ducluzeau et al., 2008). A single c-type haem-binding motif (C₂₀₃ₓₓxC₂₀₇H₂₀₈) was detected in the CcoP subunits, while protein modelling identified haem c (CySS550 X₂ C₅₃₅ His₅₅₄ Met₂₂₂) signatures from the CcoO subunits. Finally, conserved histidine residues binding the haem-copper binuclear centre including haems b and b₃ (H344 and H346, and H60, H206, H256 and H257) were identified in the CcoN subunits (Buschmann et al., 2010) (Fig. S2). Potential for microaerophilic growth in *Pedobacter* species is further supported by the presence of genes encoding a cytochrome *bd*-type quinol oxidase in SAG M16, *P. saltans* and *Pedobacter* sp. BAL39. Previous studies with this enzyme have reported an exceptionally high affinity for oxygen (Puustinen et al., 1991), and highest expression levels under microaerophilic growth conditions (Fu et al., 1991; Rice & Hempfling, 1978).

These comparative analyses suggest that potential for microaerophilic growth may be a conserved trait across *Pedobacter* species, with this respiratory versatility probably aiding growth of these species across the wide range of environments from which they have been isolated. For the SAGs analysed here, these data suggest that *Pedobacter* strains growing within the Hanford Formation may have the ability to survive over a range of O₂ concentrations, utilizing different terminal oxidases for fully aerobic or microaerophilic conditions. Coupled to the potential utilization of a wide range of carbon substrates, this trait increases the metabolic flexibility of Hanford *Pedobacter* strains and could enable active metabolism across a range of microenvironments in the 300 Area aquifer.
Potential links to local hydrology

Temporal relative abundances for Pedobacter species in the Hanford 300 Area subsurface were generally highest between December and April, before the effects of snowmelt resulted in Columbia River stage change, river water intrusion into the subsurface and a water table rise of ~1.5 m (Fig. 2) (Lin et al., 2012b). Biogeochemical shifts (e.g., nutrients, carbon influx) associated with river water intrusion into the subsurface may therefore have had a deleterious effect on Pedobacter populations. It has been hypothesized that changes in organic carbon associated with river water intrusion may drive some microbial community shifts; given that dissolved organic carbon concentrations in Columbia River water (~2 mg l⁻¹) are generally double those for Hanford groundwater, one might anticipate increases in abundances of r-strategist (fast growth rate) bacteria as more carbon is introduced into the subsurface system. This was clearly not the case for Pedobacter populations identified by Lin et al. (2012b), suggesting that they are K-strategists that succeed under resource-limited conditions (Frelich et al., 2010). Although originally applied to animals and plants, the r/K selection theory may play a role in understanding microbial dynamics in environments experiencing nutrient and resource fluxes, such as the 300 Area aquifer during Columbia River water intrusion. Increased flux of organic carbon or the introduction of a specific carbon compound during this event may offer an opportunity for fast-growing bacterial species to out-compete potential K-strategists such as Pedobacter species. The ability of specific bacterial groups to rapidly increase in abundance in response to changing organic carbon flux has been previously demonstrated in an analogous environment (Wilkins et al., 2013). Better characterizing chemical (carbon, sulfur, nitrogen species) shifts associated with Columbia River water intrusion into the Hanford aquifer will therefore be essential for identifying geochemical drivers of microbial community change. Other potential variables linked to river intrusion that might negatively impact indigenous micro-organisms include the introduction of foreign microbial species or predators into the ecosystem, or as-yet-unidentified factors.

Here we were able to isolate single cells from complex mixtures of gravel, silt and clay sediments using single-cell manipulation techniques. Amplification and sequencing of genomic DNA from targeted species allowed the physiology and ecology of Pedobacter cells to be inferred. These results revealed a relatively versatile lifestyle for these species within the Hanford aquifer, with roles in carbon cycling and the potential colonization of microaerophilic microenvironments, such as might be found in heterogeneous subsurface aquifers.

ACKNOWLEDGEMENTS

This research was supported by the US Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of the Subsurface Biogeochemistry Research Program’s Scientific Focus Area (SFA) at Pacific North-west National Laboratory (PNL). PNNL is operated for DOE by Battelle under contract DE-AC06-76RLO 1830. We thank two anonymous reviewers for their helpful comments and suggestions.

REFERENCES


Pedobacter growth strategies.

Edited by: W. Röling