The ISME Journal (2015) 9, 101–114

www.nature.com/ismej

ORIGINAL ARTICLE

Life on the edge: functional genomic response of *Ignicoccus hospitalis* to the presence of *Nanoarchaeum equitans*

Richard J Giannone[1,5], Louie L Wurch[2,5], Thomas Heimerl[3], Stanton Martin[1,4], Zamin Yang[1], Harald Huber[3], Reinhard Rachel[3], Robert L Hettich[1] and Mircea Podar[1,2]

[1]*Oak Ridge National Laboratory, Oak Ridge, TN, USA;* [2]*Department of Microbiology, University of Tennessee, Knoxville, TN, USA;* [3]*Lehrstuhl für Mikrobiologie und Archaeenzentrum, Universität Regensburg, Regensburg, Germany and* [4]*SAS Institute Inc., Cary, NC, USA*

The marine hyperthermophilic crenarchaeon *Ignicoccus hospitalis* supports the propagation on its surface of *Nanoarchaeum equitans*, an evolutionarily enigmatic archaeon that resembles highly derived parasitic and symbiotic bacteria. The cellular and molecular mechanisms that enable this interarchaea relationship and the intimate physiologic consequences to *I. hospitalis* are unknown. Here, we used concerted proteomic and transcriptomic analyses to probe into the functional genomic response of *I. hospitalis* as *N. equitans* multiplies on its surface. The expression of over 97% of the genes was detected at mRNA level and over 80% of the predicted proteins were identified and their relative abundance measured by proteomics. These indicate that little, if any, of the host genomic information is silenced during growth in the laboratory. The primary response to *N. equitans* was at the membrane level, with increases in relative abundance of most protein complexes involved in energy generation as well as that of several transporters and proteins involved in cellular membrane stabilization. Similar upregulation was observed for genes and proteins involved in key metabolic steps controlling nitrogen and carbon metabolism, although the overall biosynthetic pathways were marginally impacted. Proliferation of *N. equitans* resulted, however, in selective downregulation of genes coding for transcription factors and replication and cell cycle control proteins as *I. hospitalis* shifted its physiology from its own cellular growth to that of its ectosymbiont/parasite. The combination of these multiomic approaches provided an unprecedented level of detail regarding the dynamics of this interspecies interaction, which is especially pertinent as these organisms are not genetically tractable.
The ISME Journal (2015) **9,** 101–114; doi:10.1038/ismej.2014.112; published online 11 July 2014

Introduction

In every environment, few if any microbes live in spatial or functional isolation. The types of interactions between microbes can be complex and range from competitive to cooperative, syntrophic, even reaching obligate codependence (Hibbing *et al.*, 2010; Moissl-Eichinger and Huber, 2011; Wrede *et al.*, 2012; Morris *et al.*, 2013). Holistic studies of such interspecies relationships are limited by difficulties in identifying and stably maintaining symbiotic microbial systems in the laboratory (Orphan, 2009). Among those, the marine hyperthermophiles *Ignicoccus hospitalis* and *Nanoarchaeum equitans* represent the first described interspecific association between two archaea (Huber *et al.*, 2002). Species of the genus *Ignicoccus* have been isolated or detected based on rRNA sequences from marine hydrothermal systems around the globe (Huber *et al.*, 2000; Flores *et al.*, 2011, 2012). They are free-living, obligate hyperthermophilic chemolithoautotrophs, fixing CO_2 using the energy derived from reducing elemental sulfur with molecular hydrogen (Huber *et al.*, 2000, 2008). *N. equitans* has been codetected in the same environments but, while limited data suggest it might confer *Ignicoccus* some ecological advantage (McCliment *et al.*, 2006), there is no laboratory experimental evidence or genomic-based inference of a positive impact on *I. hospitalis* (Jahn *et al.*, 2008; Podar *et al.*, 2008a). Superficially, *N. equitans* resembles an ectoparasite, strictly relying on physical contact with its host *I. hospitalis* (Jahn *et al.*, 2008). Lacking almost all genes required for primary metabolism and energy, *N. equitans* must obtain its small-molecule cellular precursors (lipids, amino acids, sugars, nucleotides) from its host, *I. hospitalis* (Waters *et al.*, 2003;

Correspondence: M Podar, Biosciences Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA.
E-mail: podarm@ornl.gov
[5]These authors contributed equally to this work.
Received 18 January 2014; revised 30 May 2014; accepted 5 June 2014; published online 11 July 2014

Jahn *et al*., 2004, 2007). The mechanisms by which *N. equitans* recognizes its host, establishes a physical cell contact and mediates the acquisition of metabolites and energy sources are unknown. A related system involving an uncultured nanoarchaeote from a geothermal spring was described at the genomic level (Podar *et al*., 2013), suggesting that these symbiotic/parasitic archaea are widespread in thermal environments and may use common mechanisms for molecular transfers.

With a combined repertoire of only ~2000 predicted protein-encoding genes and bearing evidence of genome streamlining and coevolution (Waters *et al*., 2003; Podar *et al*., 2008a), *I. hospitalis* and *N. equitans* engage in what may be the genetically simplest interspecies association known so far. Previous whole-cell proteomic analysis revealed that these organisms constitutively express most genes and inferred the relative abundance of over 75% of the predicted encoded proteins in the genomes of *I. hospitalis* and *N. equitans* (Giannone *et al*., 2011). While our previous study was limited to one, stationary phase stage of the association, the present study focuses on using high-density microarrays and mass spectrometry proteomics to monitor the temporal dynamics of the association using relative changes in mRNA and protein abundance as a proxy to inferring molecular responses of *I. hospitalis* to proliferation of *N. equitans* on its cell surface.

Materials and methods

Cultivation of I. hospitalis and N. equitans

I. hospitalis KIN4/I (DSM 18386) was cultured either singly or in coculture with *N. equitans* at 90 °C in a 300 l bioreactor at the University of Regensburg Archaea Center using 0.5 × SME medium (Per liter: NaCl, 13.85 g; $MgSO_4.7H_2O$, 3.5 g; $MgCl_2.6H_2O$, 2.75 g; KH_2PO_4, 0.5 g; $CaCl_2.2H_2O$, 0.38 g; KCl, 0.33 g; $(NH_4)_2SO_4$, 0.25 g; NaBr, 0.05 g; H_3BO_3, 0.015 g; $SrCl_2.6H_2O$, 7.5 mg; KI, 25 μg; reduced with Na_2S, 0.2 g), excess elemental sulfur (S^o, 5.0 g l^{-1}) and a constant supply of H_2/CO_2 (10–15 l min^{-1}) as described previously (Huber *et al*., 2002; Jahn *et al*., 2008). Briefly, the culture medium was prepared and sterilized directly in the bioreactor, followed by pressurization with H_2/CO_2 at growth temperature and inoculation with 3 l of actively growing cells obtained by cultivation in 1 l bottles (200 ml per bottle). Cultivation experiments were performed for both *I. hospitalis* and *I. hospitalis–N. equitans* in separate bioreactor runs, but bioreactor replication was not possible. The dynamics of *N. equitans* propagation on its host is difficult to control and match between separate experiments. That, combined with the existence of only one bioreactor that can accommodate the growth of these organisms, limited the feasibility of biologic replication. The cultures were sampled at various time points along the ~24 h growth course, covering the transition from relatively mid-log to stationary phase. For each harvest point, 15–30 l samples were collected anaerobically, rapidly cooled and the cells were isolated by centrifugation, frozen in liquid nitrogen and stored in aliquots at −80 °C (yields of 1–2 g wet cell pellet per sample). The final sample from each bioreactor (190 l, ~15 g wet cell pellet) was collected and processed in the same manner. Cell densities were determined microscopically and the *N. equitans* frequency was calculated by analyzing 50 random *I. hospitalis* cells from the coculture.

Scanning electron microscopy

Separate small-scale batch cocultures (200 ml) were used to visualize different stages of the association by scanning electron microscopy (SEM). Approximately 20 ml of coculture were gently collected onto a 0.1 μm filter and washed using phosphate-buffered saline (pH 7.2). Cells were then aspirated from the filter and fixed using 3% glutaraldehyde in phosphate-buffered saline for 1 h at room temperature. After fixation, cells were washed with 3 ml of phosphate-buffered saline by syringe filtering and then resuspended in 2% osmium tetroxide in phosphate-buffered saline for 4 h at room temperature. Cells were then collected by centrifugation at 10 000 *g* for 5 min and washed three times in deionized water. On the final wash, cells were aliquoted onto a $5 \times 5\,mm^2$ silicon chip (Ted Palla, Redding, CA, USA) and allowed to settle for 15 min, and then adsorbed onto the surface of the chip. The chips were dehydrated through immersion in increasing ethanol concentration series (50, 75, 85, 90 and 95%) for 10 min each followed by 100% for 15 min and dried in a Ladd Critical-Point Dryer (Ladd Research, Williston, VT, USA). The samples were gold-coated using an SPI sputter coater and examined on a Zeiss Auriga FIB-SEM (Carl Zeiss SMT GmbH, Oberkochen, Germany) at the University of Tennessee Advanced Microscopy and Imaging Center. While the samples used for SEM were not the ones used for functional genomics, they intended to convey the spatial distribution of *N. equitans* on its host and were physiologically equivalent.

Gene expression microarray analysis

Cell pellets were homogenized in Trizol (Invitrogen, Carlsbad, CA, USA) and total RNA was purified in triplicate from each sample using the PureLinkRNA Kit (Invitrogen) with on-column removal of contaminating DNA. RNA quality and concentration were determined with an Agilent Bioanalyzer (Agilent, Santa Clara, CA, USA). For each sample, 10 μg RNA were converted to cDNA with the ds-cDNA Synthesis Kit (Invitrogen) and labeled with the One Color DNA Labeling Kit (Roche NimbleGen Inc., Madison, WI, USA). A high-density gene expression microarray (1plex, 385k) containing 60mer oligonucleotide probes (up to 20 different probe per gene), for the protein encoding genes of *I. hospitalis*, was designed and manufactured by Roche NimbleGen

Inc. Hybridization of the cDNAs to the arrays and washing was performed according to the manufacturer's instructions followed by scanning with a Surescan HR DNA Microarray Scanner (Agilent). All of the arrays were performed in triplicate and the images were quantified using NimbleScan 2.6 (Roche NimbleGen Inc.). Individual array raw data was $\log_2$ transformed and imported into the statistical analysis software JMP Genomics 3.0 (SAS Institute, Cary, NC, USA). The combined data was normalized using one round of Loess normalization. Distribution analyses were performed both before and after normalization as a quality control step. An analysis of variance was performed to determine differential gene expression levels via direct comparisons between time-point samples and between *I. hospitalis* versus *I. hospitalis–N. equitans* samples. A false discovery rate cutoff using an α level of 0.05 was used to correct for the multiple testing problem, as described previously (Wilson *et al.*, 2013a). The microarray data have been deposited in NCBIs Gene Expression Omnibus (Edgar *et al.*, 2002) under the accession number GSE57033.

Proteomic analysis

Time-course samples from both the *I. hospitalis* culture and the *I. hospitalis–N. equitans* coculture were prepared for proteomic analysis as described previously (Giannone *et al.*, 2011). Briefly, cell pellets (duplicate for each sample) were resuspended in sodium dodecyl sulfate lysis buffer, boiled for 5 min and pulse-sonicated. The resulting whole-cell extract was assayed by BCA analysis and 3 mg protein was precipitated by trichloroacetic acid, pelleted, washed and air-dried. The protein pellet was then resuspended in urea–dithiothreitol to maintain a reduced and denatured state, cysteines blocked by iodoacetamide treatment and proteins digested to peptides via two 20 μg additions of sequencing-grade trypsin (Promega, Madison, WI, USA) as detailed previously. Proteolyzed samples were then salted, acidified and centrifuged through a 10 kDa filter (Vivaspin 2; GE Healthcare) to collect correctly sized tryptic peptides followed by peptide quantification using a BCA assay. The resulting tryptic peptides were used for liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis.

To further enhance proteome coverage compared with our previous study, especially with regard to hydrophobic or trypsin-incompatible proteins/protein regions, the high-molecular-weight, hydrophobic/un(der)digested protein material from the above tryptic peptide sample filtering step was resuspended in 125 μl of 0.5% sodium dodecyl sulfate plus 8 M urea and boiled for 5 min. α-Chymotrypsinogen (Sigma, St Louis, MO, USA) was added in two independent aliquots at a 1:5 (w/w) protease to protein ratio: first aliquot diluted with 125 μl UB (100 mM Tris-HCl, pH 8.0) plus 10 mM $CaCl_2$ and allowing the reaction to proceed for 4 h at room temperature followed by a second aliquot diluted with 250 μl UB with digestion proceeding overnight at room temperature. Chymotryptic peptides were then salted and acidified, refiltered by centrifugation, quantified and used for LC-MS/MS analysis.

For each sample, 100 μg of peptides were loaded with the aid of a pressure cell onto a biphasic MudPIT back column containing both strong-cation exchange and reversed-phase resins and separated as detailed previously (Giannone *et al.*, 2011). Briefly, loaded samples were washed offline, and then placed inline with an in-house pulled nanospray emitter packed with 15 cm of reversed-phase resin. Peptides were then separated and analyzed with an 11-pulse MudPIT LC-MS/MS protocol over a 24-h period using a hybrid LTQ-Orbitrap-XL (Thermo Fisher, Waltham, MA, USA) MS. MS analysis parameters were as follows: data-dependent acquisition, one full scan (two microscans each) followed by five MS/MS scans (two microscans each), Orbitrap mass analyzer was set to 15K resolution, LTQ isolation window = 3 m/z, dynamic exclusion window, duration and max = 3 m/z, 60 s and 500, respectively. Each peptide sample was analyzed in technical duplicate over 24-h LC-MS/MS runs.

Peptides were matched to MS/MS spectra using MyriMatch v. 2.1 (Tabb *et al.*, 2007), filtered and assembled using IDPicker v. 3.0 (Ma *et al.*, 2009), spectral counts (SpC) tabulated, balanced and normalized as described previously (Giannone *et al.*, 2011). Using a minimum of two distinct peptides per protein and adjusting minimum SpC per protein to achieve protein-level false discovery rates of $<5\%$ (peptide-level false discovery rate $\sim 1\%$), over 1.4 million spectra confidently assigned to proteins (average 149 823; relative standard deviation 12% for *I. hospitalis* and 158 142; relative standard deviation 5% for *I. hospitalis–N. equitans*). Normalized SpC (nSpC) were used to derive individual protein relative abundance across samples/time points, clustered based on their relative trend patterns and compared across both the pure and cocultures. To measure the relative change in abundance over time, each individual protein (or groups of proteins/protein complexes) was assigned a relative slope value that describes the general linear fit through all individual time-point nSpC values, with each slope normalized by the average nSpC across all time points to put trends/slopes on a comparable scale. Although abundances may not change in a strictly linear manner, such a single-value metric allows comparisons of all proteins and identifies noisy trends. This trend-based analysis was used because it was difficult to ascertain absolute differences in abundance through semi-quantitative proteomics, particularly when comparing different proteins. To compare the abundance trends between selected protein complexes, nSpC values were standardized across the time series by reassigning values to represent the number of standard deviations from the row mean.

Results and discussion

Cultivation of I. hospitalis–N. equitans for functional genomics

When grown in pure culture, *I. hospitalis* displayed exponential growth to a density of $\sim 2 \times 10^7$ cells per ml, followed by gradual transition to stationary phase between 3 and 5×10^7 cells per ml, reaching a maximum density of 7×10^7 cells per ml after 24 h (Figure 1). As described previously (Jahn *et al.*, 2008), the coculture of *I. hospitalis* with *N. equitans* was characterized by an initial lag phase, in which *I. hospitalis* proliferated to $\sim 10^6$ cells per ml but were rarely attached with *N. equitans* cells. When harbored an average of one *N. equitans* (at $\sim 5 \times 10^6$ cells per ml), *I. hospitalis* growth slowed considerably while its symbiont continued to proliferate, reaching an average of about 7 cells per host cell and with a total cellular density resembling that of the pure *I. hospitalis* culture. These distinct growth profiles between the pure culture and the coculture occur even though $H_2/CO_2/S^o$ are not limiting during this interval. This indicates that even a single *N. equitans* cell can trigger inhibition of its host's cell division through mechanisms yet to be discovered (Jahn *et al.*, 2008). Therefore, sampling of

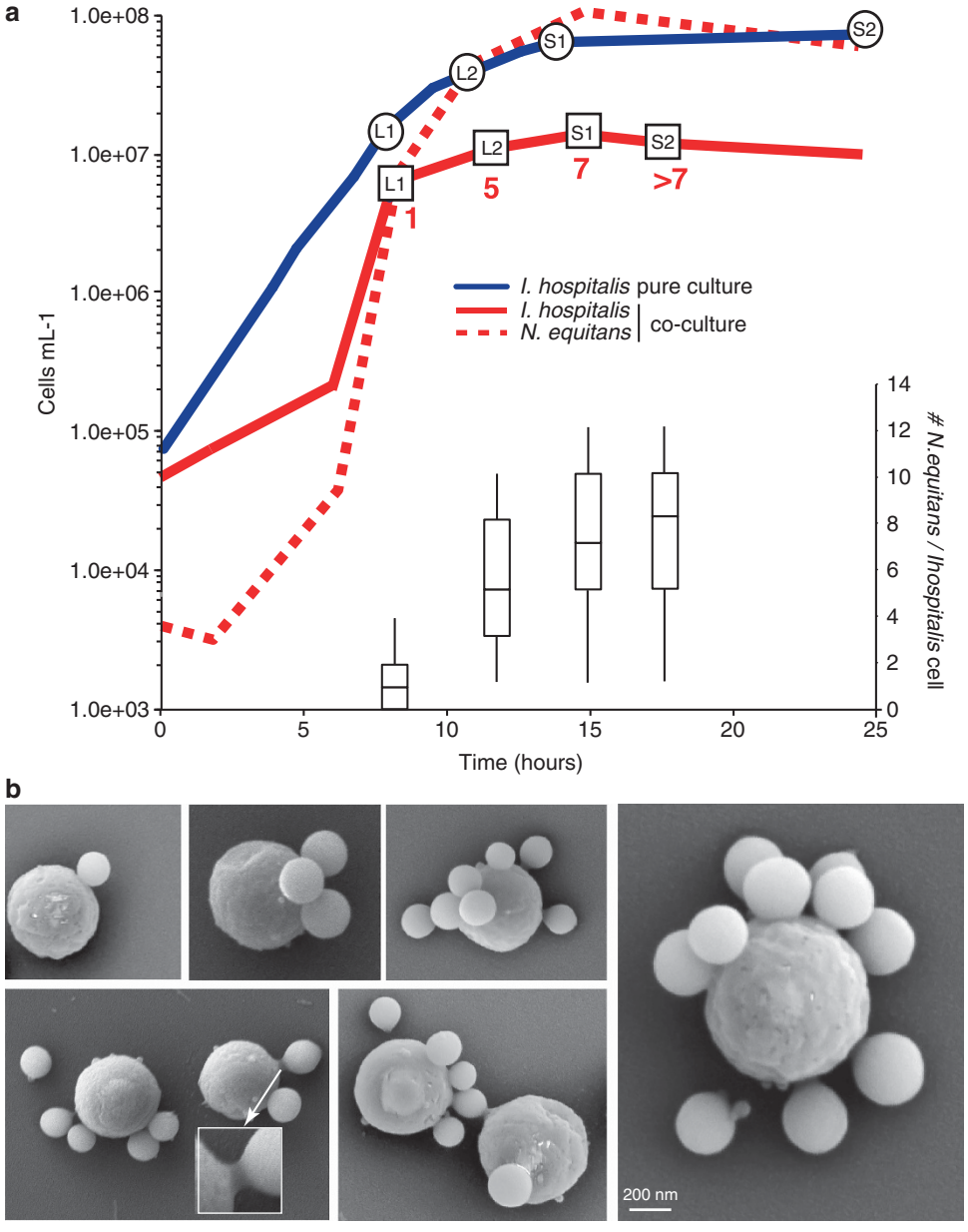


Figure 1 (**a**) Growth curves of *I. hospitalis* and *I. hospitalis–N. equitans* in the large-scale fermentor. L1, L2, S1 and S2 indicate stages used for functional genomic analysis; red numbers indicate average of *N. equitans* cells present per *I. hospitalis* cells in the coculture. The insert box-and-whiskers plot shows the frequency distribution of *N. equitans* per host cell at each of the sampled time points. (**b**) SEMs illustrating different stages of *N. equitans* colonization of *I. hospitalis* and cellular features including a close-up of the interspecies membrane contact. The samples for microscopy were taken from independent, bottle cultures.

both pure and cocultures (Figure 1) aimed at capturing genetic and molecular events associated with colonization and the proliferation of *N. equitans* on its host's surface and distinguishing them from cellular processes that mark the transition between the exponential and the stationary phases of growth. Toward the end of cocultivation, *N. equitans* cells outnumber those of its host by a factor of 10 or more, with many cells free in the culture medium, even though at every coculture stage there are *I. hospitalis* carrying a variable number of symbionts/parasites (Figure 1). Therefore, while *N. equitans* proliferation follows the same overall dynamics among independent cultivation experiments, precise matching of coculture stages for biologic replication was not feasible. We therefore compared the time-point samples using abundance trends as proxies for changes in gene expression and relative protein abundance, each relative to the other, and we also related them to previously determined *I. hospitalis* and *I. hospitalis–N. equitans* proteome profiles (Giannone *et al.*, 2011).

Overall, the transcripts of 1404 genes and the proteins encoded by 1154 genes were identified, representing ~97% of the transcriptome and ~80% of the proteome predicted from the *I. hospitalis* genome (Supplementary Table S1). This remarkably high coverage confirms a constitutively expressed, streamlined genome with few silenced genes, at least under laboratory growth conditions. Relative abundance differences between individual gene products observed during growth as a pure culture and in coculture with *N. equitans* were calculated and integrated based on known or inferred functional protein complexes, cellular structures and processes. Microbial mRNA and protein synthesis, even though interrelated, have distinct control mechanisms and turnover on different time scales (minutes versus hours), observed also in thermophiles (Andersson *et al.*, 2006; Trauger *et al.*, 2008; Sun *et al.*, 2010). In addition, proteomic data (spectra counts) and mRNA expression arrays (hybridization signal) have distinct sensitivities and dynamic ranges, which makes direct comparisons difficult. Therefore, we analyzed the *I. hospitalis* proteome and transcriptome both independently and in correlation with one another.

Proteomics of the I. hospitalis–N. equitans interaction

As a strict chemolithoautotroph, *I. hospitalis* has no metabolic alternatives to CO_2 fixation using energy derived by H_2/S^o respiration. This likely explains the very high fraction of expressed genes, as this organism cannot use other growth conditions and its genome does not encode for distinct physiologic alternatives (Huber *et al.*, 2000; Paper *et al.*, 2007; Podar *et al.*, 2008a). We have also previously shown that *N. equitans* has a relatively small effect on its host's proteome when the coculture was analyzed at the terminal stage and that the association may not require a specialized set of genes that are exclusively induced or repressed by cocultivation (Giannone *et al.*, 2011). To identify relative quantitative variations in specific proteins, protein complexes, metabolic pathways and functional categories (as classified under archaeal clusters of orthologous genes) (Wolf *et al.*, 2012) that are associated with expansion of *N. equitans*, we now analyzed each *I. hospitalis* protein's nSpC (representing the relative abundance of each protein in the entire measured proteome) during growth in pure culture versus coculture. Three phases of *N. equitans* association (empirically defined here as initial contact, active propagation, saturation), represented by an average 1, 5, 7 or more *N. equitans* cells per *I. hospitalis* cells were used to assess proteome differences based on changes in relative abundance of each protein/protein complex as compared with the pure *I. hospitalis* culture. It should be emphasized that although host–symbiont/parasite dynamics implies that *N. equitans* cells in the starting coculture are able to attach to free *I. hospitalis* cells, that process has not been demonstrated experimentally but necessarily takes place during the early cocultivation phase.

The overall abundances of proteins involved in central metabolic processes (the biosynthesis of amino acids, lipids, nucleotides and carbohydrates) were relatively constant and unaffected by the presence of *N. equitans* (Figure 2). This suggests that even though both the pure culture and the coculture reached stationary phases, the *I. hospitalis* cells continued to be metabolically active. Importantly, the low impact of *N. equitans* suggests that *I. hospitalis* is able to cope with the increased demand of metabolic precursors imparted by and transferred to its companion without upregulating most of its biosynthetic enzymatic machinery, an indication that most biosynthetic rates are not limited by protein levels. A gradual increase in the levels of enzymes involved in the CO_2 fixation pathway was nevertheless observed in the coculture, suggesting a direct response to an increased demand in key metabolic intermediates (pyruvate, acetyl-CoA, oxoglutarate and oxaloacetate) as *N. equitans* density increased. Similarly, several enzymes involved in nitrogen assimilation and amino acid interconversion were selectively enriched, most notably glutamine and asparagine synthase, which provide key precursors to other biosynthetic pathways (Figure 3).

Although the relative abundance of most proteins that assemble into the basal replication and transcription complexes were minimally affected by the presence of *N. equitans*, several of the 40 predicted transcriptional regulators suffered a decrease in abundance by 10-fold or more (Igni_840, Igni_882 and Igni_701) and a few others increased (Igni_122 and Igni_839) (Figure 2 and Supplementary Table S1). A similar effect was

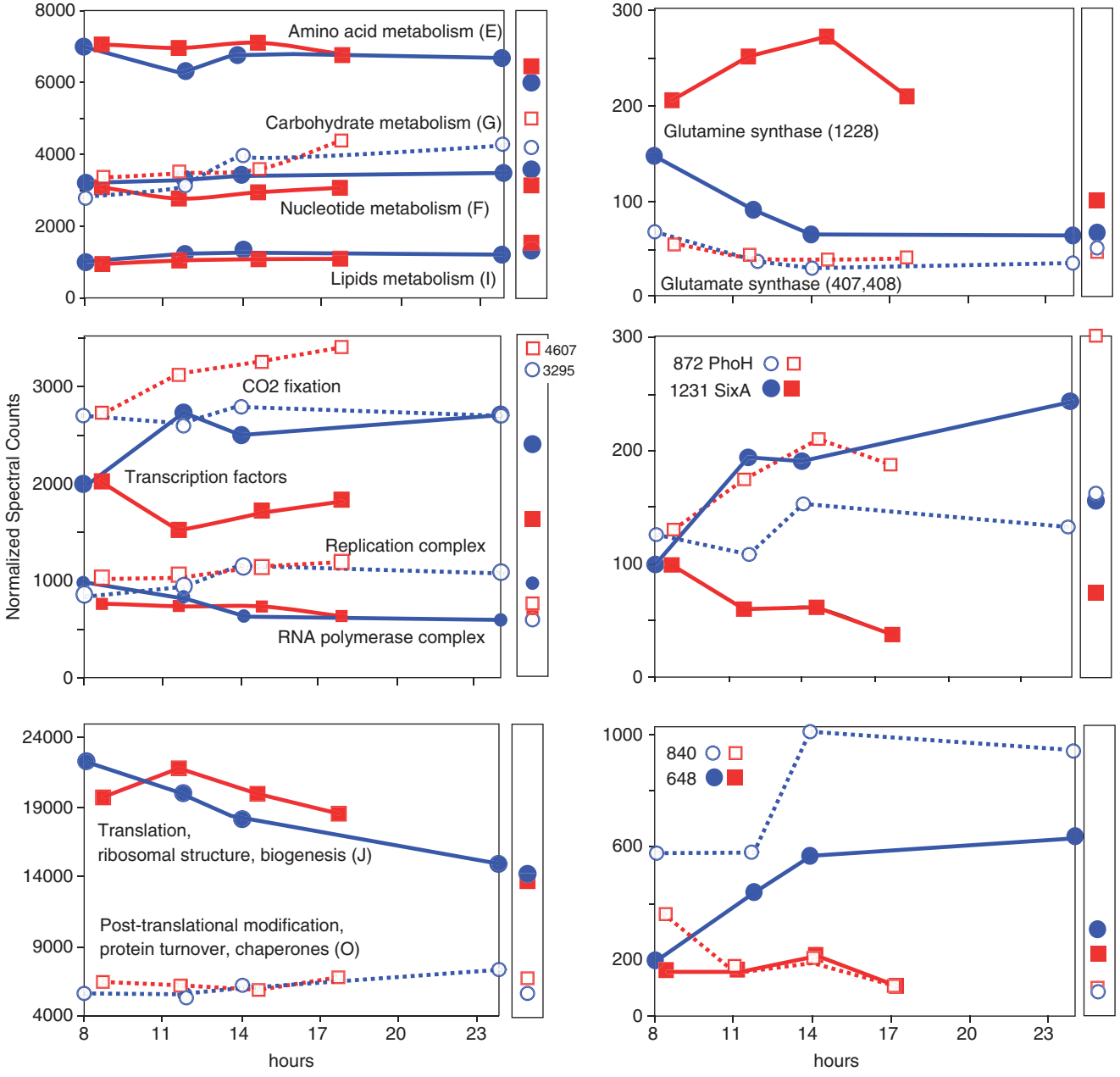


Figure 2 Dynamics of individual *I. hospitalis* proteins, protein complexes and functional protein categories (based on archaeal clusters of orthologous genes classification) in pure culture (blue symbols) and coculture with *N. equitans* (red symbols), based on nSpCs. The side panels indicate the relative abundance of those proteins reported in Giannone *et al.* (2011) normalized to the proteome scale used here. With regard to complexes and functional categories, the constituent proteins that detail the abundance trends are identified in Supplementary Table S1.

detected for proteins predicted to be involved in posttranscriptional regulation and signaling, such as a Igni_1231 (phosphohistidine phosphatase SixA), Igni_1217 (PII-like signaling protein) and Igni_648 (a CBS domain protein), which increased sharply in the pure culture, but decreased or remained constant under proliferation of *N. equitans*, whereas Igni_872 and Igni_1324 (phosphate and carbon starvation-inducible proteins) displayed the opposite trend. Interestingly, several predicted regulators were only detected in the coculture and were most abundant during the early stages of colonization (Igni_99, Igni_702 and Igni_971). The regulatory network in *I. hospitalis* is unknown, but while we cannot predict the genes controlled by those factors and their downstream effects, proliferation of *N. equitans* appears to trigger specific responses in its host. Because several major cellular processes that signal cellular injury and defense activation (translation, protein folding and turnover, CRISPR-CAS system) are relatively unchanged, cocultivation with *N. equitans* does not appear to resemble viral infection or accelerate culture aging.

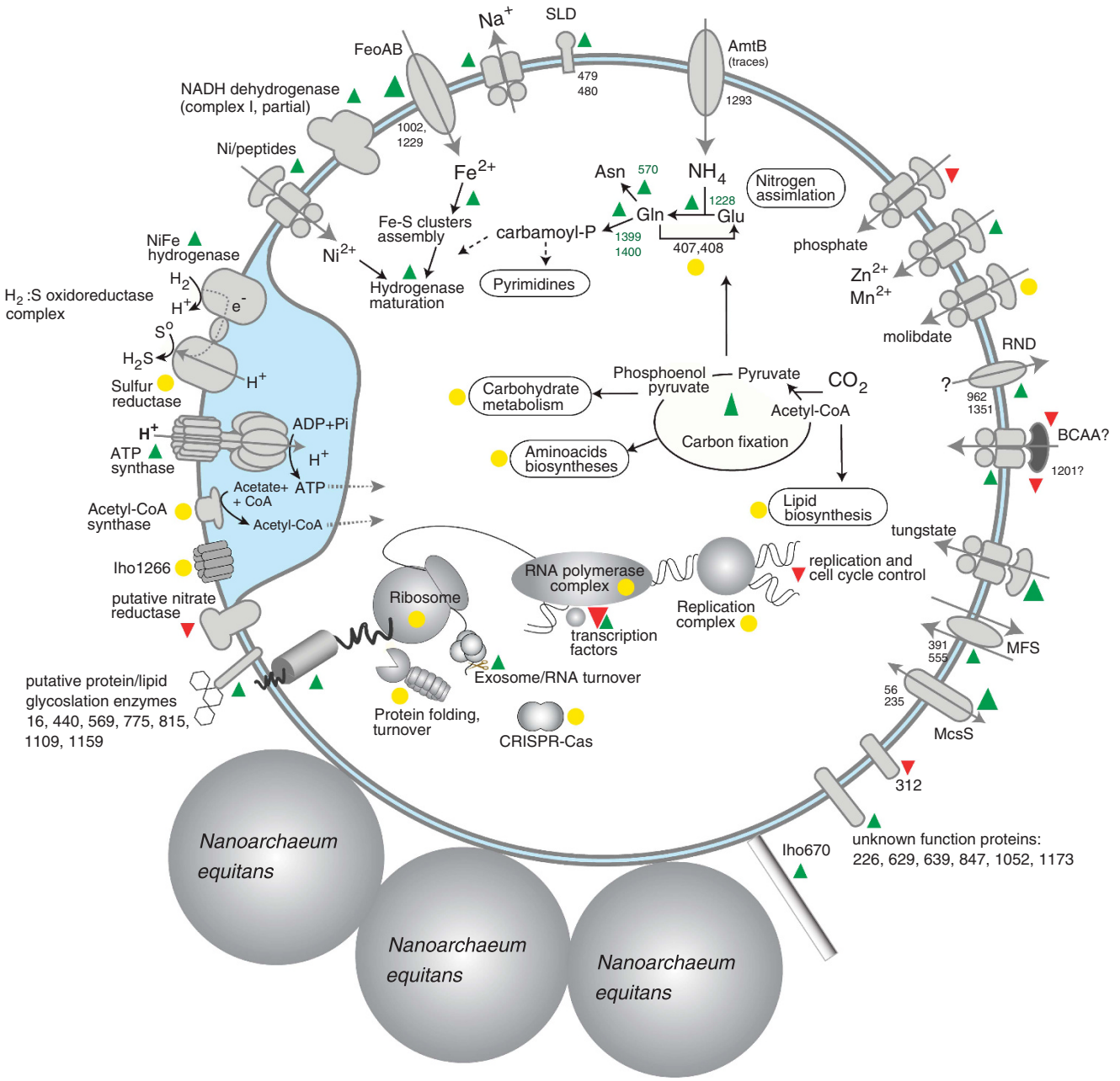


Figure 3 Cellular map diagrammatic sketch representing some of the major *I. hospitalis* metabolic processes and protein complexes and their inferred responses to *N. equitans* colonization (green, upward arrow: increase; red downward arrow: decrease; yellow circle: stable). The double-membrane organization of *I. hospitalis* is represented. Proteins/complexes for which the specific localization (outer or inner membrane) is unknown are depicted spanning both membranes. The known localization of some proteins/complexes on the outer membrane is depicted.

Membrane processes dominate the proteomic response of I. hospitalis to N. equitans

The most significant changes in the *I. hospitalis* proteome triggered by *N. equitans* were at the membrane level. The combined relative abundance of membrane proteins increased by 50% in the coculture while it remained essentially unchanged in the pure culture (Figures 3 and 4). Major categories of proteins contributing to these dynamics included proteins involved in energy generation and conversion, transport functions and membrane architecture. The ATP synthase, H_2:sulfur oxidoreductase and the putative NADH dehydrogenase complex displayed a steady increase reaching twofold higher levels in the presence of *N. equitans*. Overall, LC-MS/MS identified 26 of the 29 predicted subunits for these four important energy-generating complexes across both pure and the coculture, with 23 identified in both. Of those subunits, 22 displayed a consistent response to the presence of *N. equitans*, reaching an approximate twofold increase, while in the pure culture they

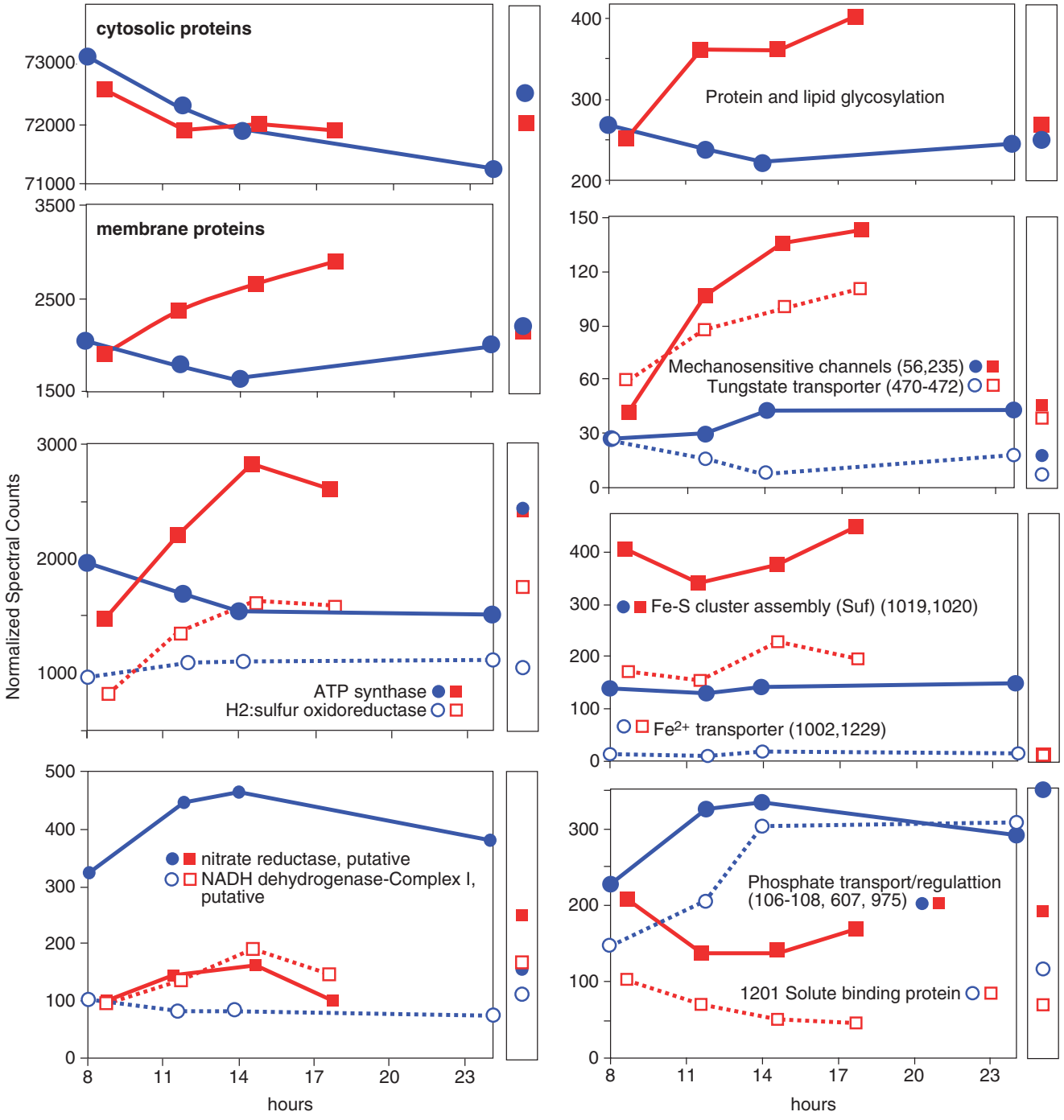


Figure 4 Dynamics of *I. hospitalis* proteins and protein complexes linked to membrane processes in pure culture (blue symbols) and coculture with *N. equitans* (red symbols), based on nSpCs.

remained relatively unchanged. Standardized abundance trends for each of the four major energy-generating complexes and for their individual subunits show high similarly with one another and over time (Supplementary Figure S1). As these complexes work in tandem to provide energy for the cell, this supports the conclusion that their increased relative abundance correlates with *N. equitans* proliferation and likely reflects the metabolic and energetic demand imposed on *I. hospitalis* by its ectosymbiont. Since in the *I. hospitalis* pure culture these complexes were relatively unchanged, the cellular energetic balance does not appear limited by protein abundance in the absence of *N. equitans*. The concordance between these complexes and their subunits provides confidence in data robustness in the absence of independent cultivation replicates. One membrane complex showed a different trend: the putative nitrate reductase, encoded by a

four-gene operon (Igni1377-1380) was present at about threefold lower levels in the coculture. Because our current knowledge suggests that *Ignicoccus* is not able to use nitrite or nitrate as electron acceptors (Huber *et al.*, 2000), understanding the biologic function of that complex and the significance of its decrease in coculture with *N. equitans* will depend on future direct enzymatic measurements.

N. equitans attaches to and interacts with *I. hospitalis* through specific, physical cell–cell contact. The actual contact site between *I. hospitalis* and *N. equitans* is small, having an area of about 1250 nm^2, or 1/400th of the *N. equitans* surface (Junglas *et al.*, 2008). Membrane-embedded proteins within this site may have key roles in this interaction by providing specificity and directly mediating transfer of small molecules. However, this is not to exclude an alternative hypothesis where partial interspecies membrane fusion and direct cytoplasmic contact between the cells could provide the means for metabolite transfer. To this end, previous thin section electron micrographs have revealed membrane 'sticking' and stretching between the cells (Junglas *et al.*, 2008), which is also evident in some of the SEM images (Figure 1).

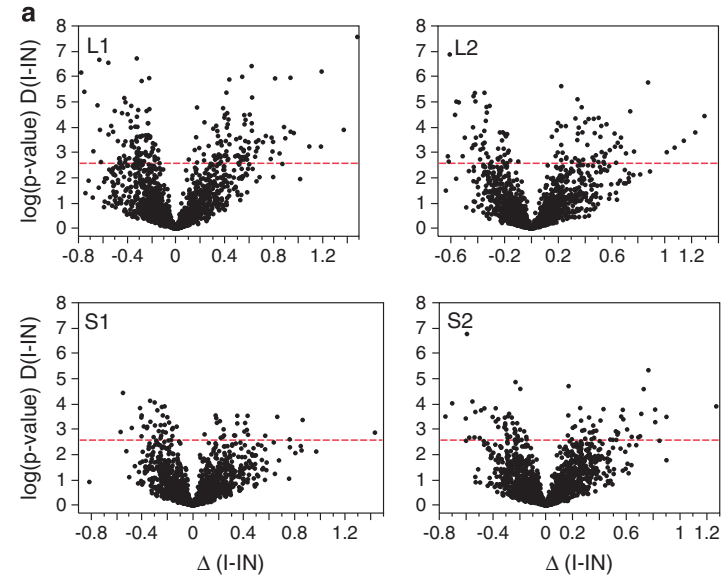
Previously, using purified *N. equitans* cells we could not detect a significant transfer of *I. hospitalis* proteins to its symbiont/parasite, which suggests most if not all metabolites are acquired, athough a yet unidentified mechanism (Giannone *et al.*, 2011). A major question is whether or not proteins involved in this interspecies interaction are constitutively expressed in *I. hospitalis* or are upregulated upon multiplication of *N. equitans*. Several predicted membrane proteins with unknown functions were highly elevated (e.g. Igni_226) as well as the fiber protein 'Iho670' (Müller *et al.*, 2009; Yu *et al.*, 2012) but the predicted pore-forming protein 'Iho1266' (Burghardt *et al.*, 2007) remained constant (Supplementary Table S1). Interestingly, enzymes predicted to be involved in protein and lipid glycosylation, some with membrane-anchor regions, are induced in the coculture, which possibly elevates membrane sugar decoration and may impact interspecies interaction. The accumulation of *N. equitans* cells on the surface of its host also triggers a sharp, nearly fivefold increase in the level of mechanosensitive channels (Igni_56 and Igni_235). Such channels are known to mediate adaptive cellular response to mechanical and osmotic stress in various organisms (Sukharev and Sachs, 2012; Wilson *et al.*, 2013b), and, in the case of the *I. hospitalis–N. equitans* system, they may take part in an interspecies membrane complex.

Membrane transporters were another class of molecules selectively impacted by the presence of *N. equitans* and, in some instances, correlated with downstream processes also upregulated based on relative protein abundance (Figures 3 and 4). For example, both Fe^{2+} and Ni^{2+} transporters were up to 10-fold more abundant in the coculture and were linked to a several-fold increase in enzymes involved in iron cluster formation and maturation of the Fe–Ni hydrogenase complex. ABC transporters assumed to be involved in uptake of zinc/manganese or tungsten, major facilitator transporters and the protein translocation machinery were similarly elevated, while a few others (a predicted phosphate transporter and several predicted permeases with unknown specificity) were strongly downregulated by the presence of *N. equitans*. Such transporter abundance changes are likely linked to the increased metabolic demand, and also to what may be surface-level and osmotic effects resulting from aggregation of multiple *N. equitans* on its host. On the other hand, acquisition of metabolic precursors by *N. equitans* clearly requires specific transport mechanisms across all three cellular membranes (two in *I. hospitalis* and one in *N. equitans*) (Rachel *et al.*, 2002), unless a direct cytoplasmic connection between the cells exists. The major facilitator superfamily permeases of *I. hospitalis*, which strongly increased in abundance under high *N. equitans* density, but are unaffected by growth phase, may have a role in that transport. Major facilitator superfamily proteins are membrane transporters ubiquitously present in all three domains of life and serve a variety of functions including transport of simple sugars, oligosaccharides, amino acids, nucleosides and a variety of other metabolites (Pao *et al.*, 1998). Previously, we compared the proteomic changes in *I. hospitalis* linked to *N. equitans*, using a single, late stationary culture stage (Giannone *et al.*, 2011). Because the growth profiles of *I. hospitalis* are different when growing in isolation versus in the presence of its symbiont, that study was unable to differentiate responses to culture aging from those triggered by *N. equitans* proliferation. The temporal proteome response to *N. equitans* analyzed here largely agrees with the prior single time-point results, and also reveals the response triggered by the initial colonization stage, before *I. hospitalis* entering stationary phase (Figures 2 and 4). In addition, the improved coverage of membrane proteins evidenced a strong response not only at the level of bioenergetic complexes but also for transporters and membrane proteins that were previously not detected or were only identified in trace amounts. These findings, while not directly revealing how the two organisms interact and transfer metabolites, identify specific responses of *I. hospitalis* to its symbiont/parasite and point to new avenues for molecular mechanistic interrogation.

Gene expression analysis complements whole-cell proteomics

The response of *I. hospitalis* to *N. equitans* was also investigated at the mRNA level using gene expression microarrays. To identify genes that are differentially expressed as *N. equitans* multiplies on its

host's surface, time points from the coculture experiment were compared with sample L1, the point at which the ratio *Ignicoccus–Nanoarchaeum* was roughly 1:1. In total, 163 genes showed significant up- or downregulation ($\alpha = 0.05$ among technical replicates) in at least one time-point relative to L1 (Figure 5). These genes may be differentially expressed due to the presence of *N. equitans* or due to changes in growth rate and culture stage. The same analysis was conducted using the pure *I. hospitalis* culture, comparing different culture stages with the equivalent mid-log cell population (L1). To this end, 166 genes were either up- or down regulated, indicating genes most



b

Experiment	Comparison	# genes	up-regulated	down-regulated	Total differentially expressed
I.hospitalis + N.equitans	L2:L1	60	27	33	163
	S1:L1	118	43	75	
	S2:L1	121	36	85	
	L2,S1,S2:L1	36	11	25	
I.hospitalis	L2:L1	26	14	12	166
	S1:L1	73	34	39	
	S2:L1	124	51	73	
	L2,S1,S2:L1	10	4	6	

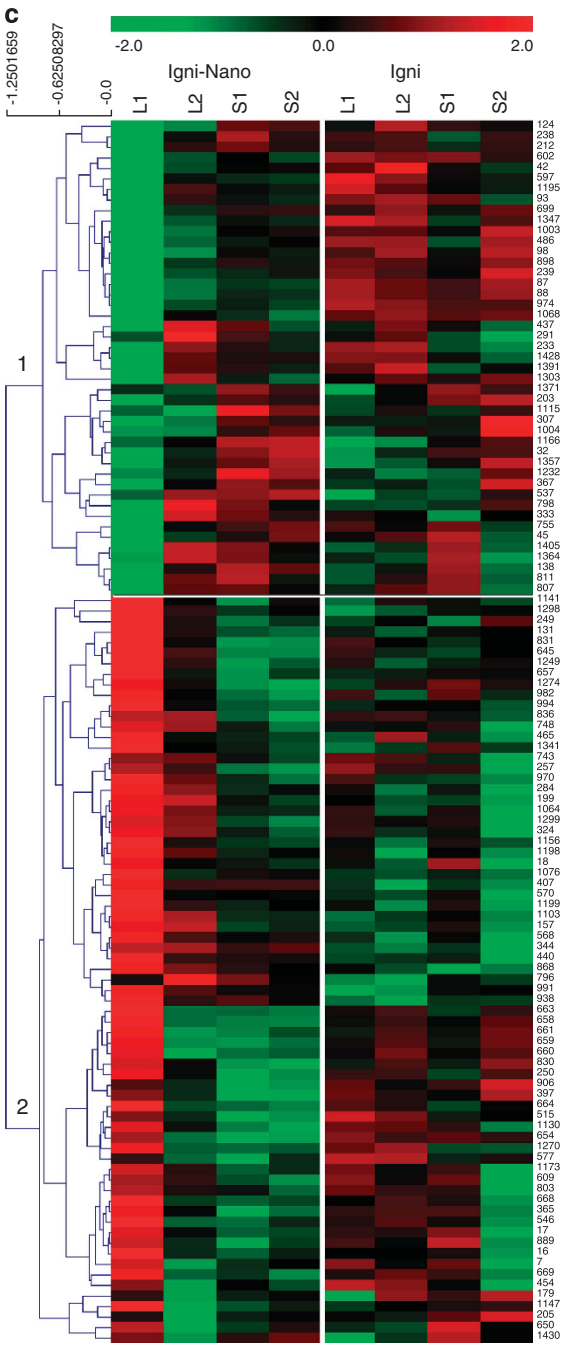


Figure 5 Differentially expressed *I. hospitalis* genes in the presence of *N. equitans* based on mRNA levels. (**a**) Volcano plots of significant differentially expressed genes (above dotted red line) between matched time-point samples of *I. hospitalis* (I) and *I. hospitalis–N. equitans* (IN). (**b**) Summary of the number of *I. hospitalis* genes up- or downregulated relative to initial time point of the culture and coculture. (**c**) Heatmap clusters of relative upregulated (red) or downregulated genes (green) linked to *N. equitans* proliferation. The genes were grouped based on similarity of expression (Pearson's correlation, with average gene linkage hierarchical clustering), with the two major groups reflecting up- or downregulation trends relative to the earliest coculture stage. Numbers on the right refer to the corresponding gene (*ORF*) from the *I. hospitalis* genome (also identified in Supplementary Table S1).

likely responding to growth rate and/or culture stage. The overall differences in gene expression were small, with the overwhelming majority of genes being <1.5-fold different across the different time points. When genes that were up- or downregulated from mid-log to stationary phase in both the culture and the coculture were excluded, a set of 115 genes remained for the coculture set, with 46 genes indicating statistically significant upregulation and 69 being specifically downregulated as *N. equitans* proliferates (Figures 5 and 6). Among the upregulated genes, half encode proteins with unknown function, and 11 of them are predicted to be membrane bound. Two transcriptional regulators were also induced, one (*Igni*_486) from the xenobiotic response element family that also showed elevated response at protein level. The 69 repressed genes represent a broad range of cellular processes including primary metabolism, energetic functions and information processing. Among them, multiple genes encoding key components of the replication machinery and cell division control were repressed, including DNA topoisomerase and reverse gyrase, subunits of the replication initiation complex and the ESCRT (Endosomal Sorting Complex Required for Transport) system. Some of those proteins were not detected or were present at very low levels in the proteome as the cell synthesizes them in relatively few copies, mRNA thus providing a complementary view of this interspecies dynamics. These findings are consistent with previous experiments that showed that even a single attached cell of *N. equitans* significantly restricts its host's cell division (Jahn *et al.*, 2008). Additionally, the growth curve of *I. hospitalis*, which displays a more rapid entry into stationary phase when in coculture with its symbiont (Figure 1), has been described previously (Jahn *et al.*, 2008) and points to a cytostatic effect that *N. equitans* proliferation exerts on its host.

To identify potential correlations between up- or downregulated genes at the mRNA level and the relative abundance of encoded proteins, we first calculated a protein variation index Δp (slope of change in abundance) for both the pure culture (ΔpI) and the coculture (ΔpIN). For proteins with little change in relative abundance that index is small, near zero. Most proteins display an index between -0.2 and 0.2 for both experiments, reflecting small effects of both culture stage and presence of *N. equitans* (Figure 6). Proteins with a similar (positive or negative) index between pure culture and coculture reflect changes linked to growth stage and not to *N. equitans* presence, most showing a decrease in abundance. The impact of *N. equitans* was evidenced at several levels. Proteins with $\Delta pIN > \Delta pI$ ('induced') indicate either a positive effect of *N. equitans* by potential upregulation of their synthesis or stabilization of abundance level during growth relative to the pure culture, where their abundance decreased (negative ΔpI). Proteins with $\Delta pIN < \Delta pI$ ('repressed') are those synthesized at lower rates as *N. equitans* proliferates and ones that are turned over by proteolysis faster than in the pure *I. hospitalis* culture.

The genes significantly induced or repressed based on expression microarrays were identified in the Δp plot of the whole cell proteome. As shown in Figure 6, for most of them Δp was small (between -0.2 and 0.2), and indicative of reduced correlative power between the transcriptome and the proteome, as well as reflecting the general small impact of *N. equitans* on both its host transcriptome and proteome. Lack of total correlation between proteomic and microarray data is well known to occur owing to temporal differences in mRNA versus proteins synthesis and turnover, posttranscriptional and posttranslational regulation, as well as to differences in the type of signals and dynamic range between hybridization microarrays and mass spectrometry (Sun *et al.*, 2010; Lee *et al.*, 2011; Vogel and Marcotte, 2012; Haider and Pal, 2013). Nonetheless, some genes did exhibit strong correlations between transcript and protein expression, in particular those repressed in the coculture (Figure 6). Among them are genes involved in replication and cell division control, several transporter subunits and membrane proteins with unknown functions (Igni_312), transferases and regulators. One of them, a RecA-type ATPase (Igni_064) potentially involved in signal transduction is strongly repressed by proliferation of *N. equitans* but is induced during pure culture growth. Similarly, an AAA family ATPase assigned to the cell division ESCRT system (Igni994) tracks the cellular proliferation at both RNA and protein levels, being rapidly repressed in the coculture but showing a progressive decline in the pure culture. It appears therefore that on one hand *Ignicoccus* responds to the presence of *Nanoarchaeum* by increasing the abundance of a wide range of membrane-level proteins and complexes involved in energy generation, transport and maintenance of cellular integrity, with selective upregulation of metabolic components linked to those processes. The other part of its response involves a marked slowdown in cell division, which is seemingly under specific transcriptional and replication controls. Because the overall changes in gene expression patterns are modest, it is still unclear whether the inferred changes in membrane composition result from selective posttranscriptional regulation or changes in membrane protein turnover rates.

Is N. equitans *a parasite?*

The exact nature of the relationship between *N. equitans* and *I. hospitalis* has intrigued microbiologists for over a decade, being described as an 'intimate association' (Jahn *et al.*, 2008). The lack of evidence of any beneficial effect coupled with a negative impact on host cell division, potentially linked to the drain of cellular metabolites, suggests

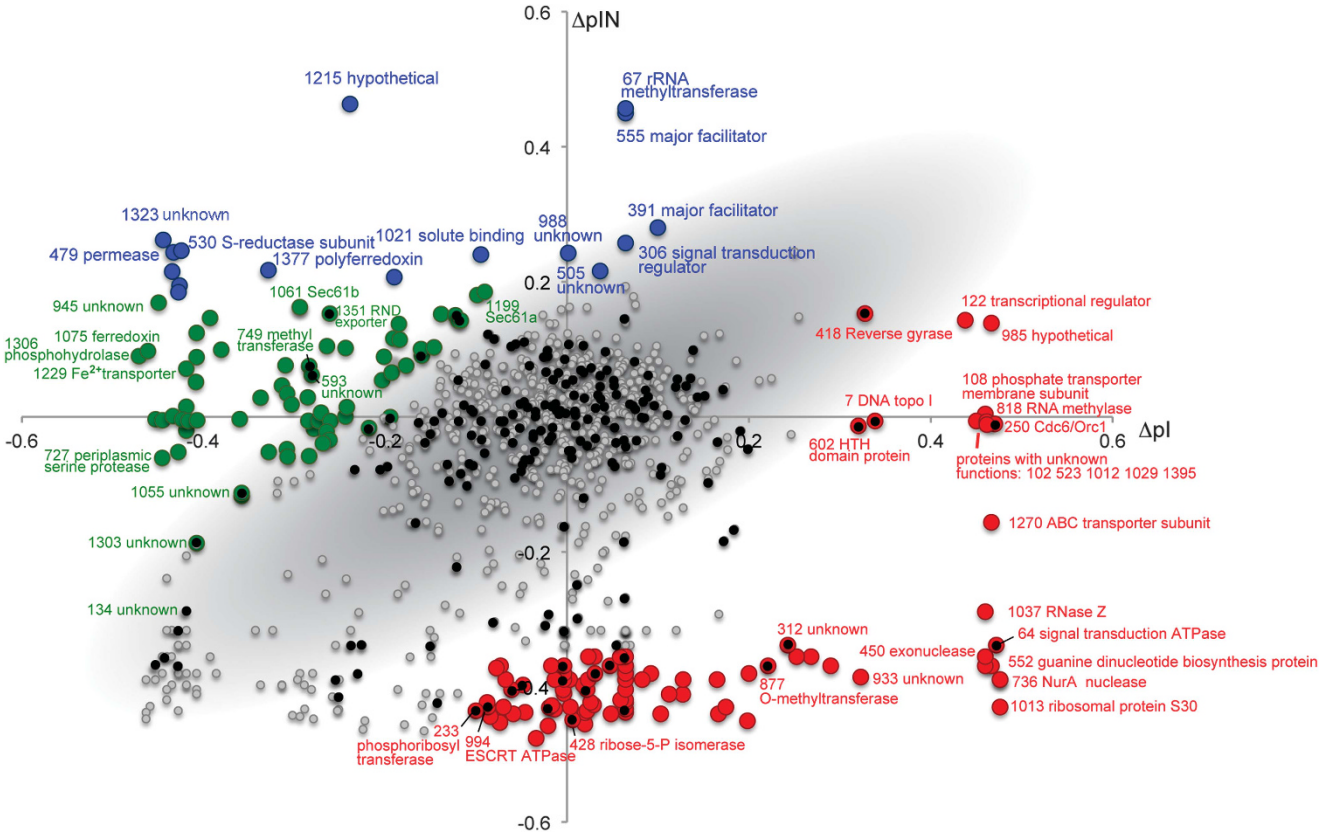


Figure 6 Abundance variation of *I. hospitalis* proteins in pure culture (Δ*pI*) and in coculture with *N. equitans* (Δ*pIN*). Each circle represents an identified protein. Proteins in the shaded diagonal region show little variation between the pure culture and coculture, and those on the outside indicate increasing levels of enrichment (blue) or depletion (red) in the presence of *N. equitans*. Some *I. hospitalis* proteins while not necessarily increasing with proliferation of *N. equitans* are stabilized by its presence relative to its abundance in the pure culture (green). Overlapping black circles indicate significant up- or downregulation of those corresponding genes at the mRNA level.

that *Nanoarchaeum* is a nutritional parasite on *I. hospitalis*. On the other hand, the surprisingly subdued global response at gene expression level to rampant proliferation of *N. equitans* on its surface and the apparent lack of a defense mechanism or stress response suggests that *I. hospitalis* has evolved a resilient metabolism that can effectively cope with that demand without major genomic regulatory changes. The significant streamlining of its genome and the lack of metabolic alternatives to chemolithoautotrophy probably further limits its responsive capacity.

Archaea from the *Ignicoccus* genus have a multitude of rather unique cellular and genomic characteristics, including a double-membrane system separated by a large intermembrane space containing vesicles or tubes budding from the cytoplasmic membrane (Rachel *et al.*, 2002) and multiple genes encoding V4R proteins related to components of the eukaryotic vesicle transport system (Podar *et al.*, 2008b). Those vesicles/tubes migrate and fuse to the outer membrane, which lacks typical S-layer proteins but contain the energy generating protein machinery (Küper *et al.*, 2010; Mayer *et al.*, 2012). It would be tempting to consider that these features may have been exploited by *N. equitans* in its adaptation to use *Ignicoccus* as a host. In this context, it is worth noting that members of the Nanoarchaeaota that are nutritionally and energetically dependent on archaea distinct from *Ignicoccus* have recently been discovered in terrestrial hydrothermal environments (Podar *et al.*, 2013). The fundamental mechanisms that enable cell–cell contact and molecular transfer are likely to be linked to both the host characteristics and the Nanoarchaeota lineage; they lead to parasitic adaptations specific for different hosts, influenced by their gene content, cellular architecture and physiology, as well as the environment. As hyperthermophilic Archaea have been argued to harbor specific genomic, biochemical and membrane level adaptations that enable them to thrive under chronic energy stress (Valentine, 2007), the hosts of Nanoarchaeota appear to have pushed that limit even further to the edge by supporting nutritional and energetic parasites. Comparative physiologic, ultrastructural and molecular studies of such systems should bring us closer to understanding the mechanisms and the evolutionary histories of these archaeal associations.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This research was supported by a grant from the US Department of Energy, Office of Biological and Environmental Research (DE-SC0006654) and by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL). ORNL is managed by UT-Battelle, LLC, for the US Department of Energy. TH, HH and RR were funded by Deutsche Forschungsgemeinschaft. We would like to acknowledge the University of Tennessee Advanced Microscopy and Imaging Center for instrument use, scientific and technical assistance.

References

Andersson AF, Lundgren M, Eriksson S, Rosenlund M, Bernander R, Nilsson P. (2006). Global analysis of mRNA stability in the archaeon *Sulfolobus*. *Genome Biol* **7**: R99.

Burghardt T, Näther DJ, Junglas B, Huber H, Rachel R. (2007). The dominating outer membrane protein of the hyperthermophilic Archaeum *Ignicoccus hospitalis*: a novel pore-forming complex. *Mol Microbiol* **63**: 166–176.

Edgar R, Domrachev M, Lash AE. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**: 207–210.

Flores GE, Campbell JH, Kirshtein JD, Meneghin J, Podar M, Steinberg JI *et al.* (2011). Microbial community structure of hydrothermal deposits from geochemically different vent fields along the Mid-Atlantic Ridge. *Environ Microbiol* **13**: 2158–2171.

Flores GE, Shakya M, Meneghin J, Yang ZK, Seewald JS, Geoff Wheat C *et al.* (2012). Inter-field variability in the microbial communities of hydrothermal vent deposits from a back-arc basin. *Geobiology* **10**: 333–346.

Giannone RJ, Huber H, Karpinets T, Heimerl T, Küper U, Rachel R *et al.* (2011). Proteomic characterization of cellular and molecular processes that enable the *Nanoarchaeum equitans–Ignicoccus hospitalis* relationship. *PLoS One* **6**: e22942.

Haider S, Pal R. (2013). Integrated analysis of transcriptomic and proteomic data. *Curr Genom* **14**: 91–110.

Hibbing ME, Fuqua C, Parsek MR, Peterson SB. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* **8**: 15–25.

Huber H, Burggraf S, Mayer T, Wyschkony I, Rachel R, Stetter KO. (2000). *Ignicoccus* gen. nov., a novel genus of hyperthermophilic, chemolithoautotrophic Archaea, represented by two new species, *Ignicoccus islandicus* sp nov and *Ignicoccus pacificus* sp nov. *Int J Syst Evol Microbiol* **50**(Part 6): 2093–2100.

Huber H, Gallenberger M, Jahn U, Eylert E, Berg IA, Kockelkorn D *et al.* (2008). A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum *Ignicoccus hospitalis*. *Proc Natl Acad Sci USA* **105**: 7851–7856.

Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**: 63–67.

Jahn U, Gallenberger M, Paper W, Junglas B, Eisenreich W, Stetter KO *et al.* (2008). *Nanoarchaeum equitans* and *Ignicoccus hospitalis*: new insights into a unique, intimate association of two archaea. *J Bacteriol* **190**: 1743–1750.

Jahn U, Huber H, Eisenreich W, Hugler M, Fuchs G. (2007). Insights into the autotrophic CO_2 fixation pathway of the archaeon *Ignicoccus hospitalis*: comprehensive analysis of the central carbon metabolism. *J Bacteriol* **189**: 4108–4119.

Jahn U, Summons R, Sturt H, Grosjean E, Huber H. (2004). Composition of the lipids of *Nanoarchaeum equitans* and their origin from its host Ignicoccus sp. strain KIN4/I. *Arch Microbiol* **182**: 404–413.

Junglas B, Briegel A, Burghardt T, Walther P, Wirth R, Huber H *et al.* (2008). *Ignicoccus hospitalis* and *Nanoarchaeum equitans*: ultrastructure, cell–cell interaction, and 3D reconstruction from serial sections of freeze-substituted cells and by electron cryotomography. *Arch Microbiol* **90**: 395–408.

Küper U, Meyer C, Müller V, Rachel R, Huber H. (2010). Energized outer membrane and spatial separation of metabolic processes in the hyperthermophilic Archaeon *Ignicoccus hospitalis*. *Proc Natl Acad Sci USA* **107**: 3152–3156.

Lee MV, Topper SE, Hubler SL, Hose J, Wenger CD, Coon JJ *et al.* (2011). A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. *Mol Syst Biol* **7**: 514.

Ma ZQ, Dasari S, Chambers MC, Litton MD, Sobecki SM, Zimmerman LJ *et al.* (2009). IDPicker 2.0: improved protein assembly with high discrimination peptide identification filtering. *J Proteome Res* **8**: 3872–3881.

Mayer F, Küper U, Meyer C, Daxer S, Müller V, Rachel R *et al.* (2012). AMP-forming acetyl coenzyme A synthetase in the outermost membrane of the hyperthermophilic crenarchaeon *Ignicoccus hospitalis*. *J Bacteriol* **194**: 1572–1581.

McCliment EA, Voglesonger KM, O'Day PA, Dunn EE, Holloway JR, Cary SC. (2006). Colonization of nascent, deep-sea hydrothermal vents by a novel Archaeal and Nanoarchaeal assemblage. *Environ Microbiol* **8**: 114–125.

Moissl-Eichinger C, Huber H. (2011). Archaeal symbionts and parasites. *Curr Opin Microbiol* **14**: 364–370.

Morris BE, Henneberger R, Huber H, Moissl-Eichinger C. (2013). Microbial syntrophy: interaction for the common good. *FEMS Microbiol Rev* **37**: 384–406.

Müller DW, Meyer C, Gurster S, Küper U, Huber H, Rachel R *et al.* (2009). The Iho670 fibers of *Ignicoccus hospitalis*: a new type of archaeal cell surface appendage. *J Bacteriol* **191**: 6465–6468.

Orphan VJ. (2009). Methods for unveiling cryptic microbial partnerships in nature. *Curr Opin Microbiol* **12**: 231–237.

Pao SS, Paulsen IT, Saier MH Jr. (1998). Major facilitator superfamily. *Microbiol Mol Biol Rev* **62**: 1–34.

Paper W, Jahn U, Hohn MJ, Kronner M, Näther DJ, Burghardt T *et al.* (2007). *Ignicoccus hospitalis* sp. nov., the host of '*Nanoarchaeum equitans*'. *Int J Syst Evol Microbiol* **57**: 803–808.

Podar M, Anderson I, Makarova KS, Elkins JG, Ivanova N, Wall MA *et al.* (2008a). A genomic analysis of the archaeal system *Ignicoccus hospitalis–Nanoarchaeum equitans*. *Genome Biol* **9**: R158.

Podar M, Makarova KS, Graham DE, Wolf YI, Koonin EV, Reysenbach AL. (2013). Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. *Biol Direct* **8**: 9.

Podar M, Wall MA, Makarova KS, Koonin EV. (2008b). The prokaryotic V4R domain is the likely ancestor of a key component of the eukaryotic vesicle transport system. *Biol Direct* **3**: 2.

Rachel R, Wyschkony I, Riehl S, Huber H. (2002). The ultrastructure of *Ignicoccus*: evidence for a novel outer membrane and for intracellular vesicle budding in an archaeon. *Archaea* **1**: 9–18.

Sukharev S, Sachs F. (2012). Molecular force transduction by ion channels: diversity and unifying principles. *J Cell Sci* **125**: 3075–3083.

Sun N, Pan C, Nickell S, Mann M, Baumeister W, Nagy I. (2010). Quantitative proteome and transcriptome analysis of the archaeon *Thermoplasma acidophilum* cultured under aerobic and anaerobic conditions. *J Proteome Res* **9**: 4839–4850.

Tabb DL, Fernando CG, Chambers MC. (2007). MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res* **6**: 654–661.

Trauger SA, Kalisak E, Kalisiak J, Morita H, Weinberg MV, Menon AL *et al.* (2008). Correlating the transcriptome, proteome, and metabolome in the environmental adaptation of a hyperthermophile. *J Proteome Res* **7**: 1027–1035.

Valentine DL. (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat Rev Microbiol* **5**: 316–323.

Vogel C, Marcotte EM. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* **13**: 227–232.

Waters E, Hohn MJ, Ahel I, Graham DE, Adams MD, Barnstead M *et al.* (2003). The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci USA* **100**: 12984–12988.

Wilson CM, Rodriguez Jr M, Johnson CM, Martin SL, Chu TM, Wolfinger RD *et al.* (2013a). Global transcriptome analysis of *Clostridium thermocellum* ATCC 27405 during growth on dilute acid pretreated *Populus* and switchgrass. *Biotechnol Biofuels* **6**: 179.

Wilson ME, Maksaev G, Haswell ES. (2013b). MscS-like mechanosensitive channels in plants and microbes. *Biochemistry* **52**: 5708–5722.

Wolf YI, Makarova KS, Yutin N, Koonin EV. (2012). The updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. *Biology Direct* **7**: 46.

Wrede C, Dreier A, Kokoschka S, Hoppert M. (2012). Archaea in symbioses. *Archaea* **2012**: 596846.

Yu X, Goforth C, Meyer C, Rachel R, Wirth R, Schroder GF *et al.* (2012). Filaments from *Ignicoccus hospitalis* show diversity of packing in proteins containing N-terminal type IV pilin helices. *J Mol Biol* **422**: 274–281.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)