Published in: Bioresource Technology 156 (2014) 84–91

Anodic microbial community diversity as a predictor of the power output of microbial fuel cells

James P. Stratford[a†*], Nelli J. Beecroft[a*], Robert C.T. Slade[b], Andre Gruning[c], and Claudio Avignone-Rossa[a‡]

[a]Department of Microbial and Cellular Sciences, [b]Department of Chemistry, and [c]Department of Computing

University of Surrey

Guildford, Surrey GU2 7XH, UK

†Current address: School of Biology, University of Nottingham, Nottingham NG7 2RD, United Kingdom

*JPS and NJB contributed equally to this manuscript

‡Corresponding author:
Claudio Avignone-Rossa
e-mail: c.avignone-rossa@surrey.ac.uk
Tel: +44 1483 686457
Fax: +44 1483 686401

Abstract

The relationship between the diversity of mixed-species microbial consortia and their electrogenic potential in the anodes of microbial fuel cells was examined using different diversity measures as predictors. Identical microbial fuel cells were sampled at multiple time-points. Biofilm and suspension communities were analysed by denaturing gradient gel electrophoresis to calculate the number and relative abundance of species. Shannon and Simpson indices and richness were examined for association with power using bivariate and multiple linear regression, with biofilm DNA as an additional variable. In simple bivariate regressions, the correlation of Shannon diversity of the biofilm and power is stronger ($r=0.65$, $p=0.001$) than between power and richness ($r=0.39$, $p=0.076$), or between power and the Simpson index ($r=0.5$, $p=0.018$). Using Shannon diversity and biofilm DNA as predictors of power, a regression model can be constructed ($r=0.73$, $p<0.001$). Ecological parameters such as the Shannon index are predictive of the electrogenic potential of microbial communities.

1. Introduction

Microbial fuel cells (MFCs) are a promising technology for the generation of energy and treatment of wastewaters (Sun et al 2010; Rahimnejad et al 2012). These devices are fed substrates which are processed by the metabolism of the microorganisms present in the anode (either single species or multi-species consortia). As the system is kept under anaerobic conditions, organisms will use the anode of the MFC as their terminal electron acceptor (Logan et al 2006). A number of microbial species has been observed to be electrogenic, and representatives of most classes of bacteria have been reported to be present in the microbial communities used in MFCs (Logan et al 2006; Nimje et al 2012). Many different parameters affect the power output and performance of MFCs. In addition to internal resistance, cathode performance and proton transfer, the power output of an MFC will also depend on the efficiency with which the community (usually attached to the anode as a biofilm) transfers electrons to the anode, and the potential link between community structure and functioning in MFCs has been recently shown for the first time (Wrighton et al 2010). While a great variety of organisms have been shown to be electrogenic, the taxonomic structure of communities is highly variable with no established taxonomic rules as to what constitutes an electrogenic organism. It has also been observed that there is no ideal electrogenic consortium, and observed trends include a tendency for enriched mixed electrogenic communities to contain a larger proportion of β-Proteobacteria (Chae et al 2009). While communities have been characterised in previous work, generally studies focus on variation in substrate, reactor design, or monocultures of novel organisms (Di Lorenzo et al 2010; Kiely et al 2011; Wang et al 2011). Interestingly,

ecological models to discover predictive relationships between fundamental community ecology and power output in replicate systems have not been used to a large extent in MFC research (Gruning et al., submitted). Diversity is a logical starting point for this type of analysis, as is an assessment of the value of different methods for its measurement. It has been observed that diversity varies significantly with design, and links have been made between more powerful fuel cell architecture and increased diversity. The systems that have been compared were not identical in design and operating parameters, lending the situation to interpretation (Kim et al 2011; Sun et al 2010).

The diversity and abundance in microbial communities can be assessed by a number of methods commonly used in ecological studies. The Shannon and Simpson indices of diversity are two of the most widely used measures for biodiversity in macro ecology (Keylock and Lane 2005; Hejda et al 2009), while Richness, a measure of species number, is also usually employed (Wilson and Brennan 2004; Grunewald 2006). The Shannon index has been used in the analysis of microbial communities (Gafan et al 2005; Steen et al 2010). Interestingly, no examples could be found in microbial ecology of biodiversity being examined as a predictive independent variable for a performance measure rather than a dependent measure in its own right. This is likely to be because most ecological analyses assume biodiversity is an inherently desirable outcome, rendering it the dependent variable in analysis. Despite the important role of microbial communities associated with the anode, there has not been a systematic quantitative analysis of the predictive power of different measures of diversity on the output of microbial fuel cells.

63 The objective of this study is to analyse the predictive power of the Shannon index, the
64 Simpson index and the richness of the anodic microbial communities with regard to
65 the power output of replicate MFCs over the course of 91 days. As different measures
66 of diversity weigh low and high abundance components differently, it is interesting to
67 analyse which notion of diversity best captures those structural aspects that support
68 community function. Richness, the inverse Simpson index and the Shannon index are
69 members of one family of indices ${}^{q}D$ (Hill 1973; Jost 2006; Tuomisto 2010), with the
70 parameter q determining how a particular index weighs relative differences in the
71 component abundances. Lower values of q emphasise the contribution of low
72 abundance components to diversity, while higher q values highlight the contribution of
73 predominant components. The case of q=0 corresponds to species abundance being
74 ignored completely and only their presence or not is taken into account: ${}^{0}D$ is the
75 species richness. The case of q=1 gives all component abundances their "natural"
76 weight; an example of ${}^{1}D$ is the Shannon index. Finally ${}^{2}D$ is the (inverse) Simpson
77 index, a measure of the diversity of the system. The numerical value of all indices ${}^{q}D$ is
78 interpreted as the "effective species number", ESN (Tuomisto 2010), and all ${}^{q}D$ yield
79 value N if N species are present with equal abundance. Hence, it can be seen as the
80 number of evenly distributed species required to produce the observed index value.
81 The ESN is considered to represent a measure of "true diversity", and the different ${}^{q}D$
82 allow comparisons between variables to be carried out on a scale using the same units.

83 Richness is the simplest measure of community diversity, and is defined as the number
84 of distinct species present within a given community. Richness does not take account
85 of the relative abundances of the species present or any other quantifiable properties

of the community. Hence the effective number of species for this measure is the actual number of species. Of the three measures considered here, Richness gives the least information about the distribution of species.

The Shannon index is a measure of information entropy (Cover and Thomas 1991), and describes not only the presence and absence of species but also the information contained in their relative abundances. The sum of the product of the probabilities with their natural logarithms gives the value of the Shannon index for a community.

$$H' = -\sum_{i=1}^{R} P_i \ln P_i$$

To convert the Shannon index into the effective species number the exponential is taken.

$$ESN_H = {}^1D = expH'$$

The Simpson index λ is the sum of the squares of the fractional species abundances, and is used to assess diversity.

$$\lambda = \sum_{i=1}^{R} P_i^2$$

where $\boldsymbol{P_i}$ is the fractional abundance of the *i*th species.

The Simpson index value represents the probability that two organisms chosen at random from the community will be of the same species. Throughout this work, the inverse has been used

$$ESN_G \coloneqq {}^2\boldsymbol{D} = \frac{\mathbf{1}}{\boldsymbol{\lambda}}$$

104

105

2. Materials and methods

2.1 MFC setup and operation

Four replicate continuously sucrose-fed microbial fuel cells (MFC1, MFC2, MFC3, and MFC4) were used. The single-chamber MFCs consisted of anode chambers (9 cm^3) and cover plates made of Perspex, with stainless steel metal plates serving as a contact between the cathode and the electrical circuit. The anode electrode contained a carbon fibre veil (PRF Composite Materials, UK) with polyvinyl alcohol binder, with a geometric area of 32 cm^2, which was placed inside the anode chamber and connected to an electrical circuit with an insulated Ni/Cr wire (Advent Research Materials, UK) knitted across the multi-layered anode. The air-breathing cathode consisted of type A carbon cloth (9 cm^2, E-TEK) coated with 4 mg cm^{-2} of Pt black catalyst with polytetrafluoroethylene binder. The platinum side of the cathode was painted with 0.5–1.0 mg cm^{-2} of Nafion perfluorinated ion-exchange ionomer (5% w/v dispersion in lower aliphatic alcohols and H_2O, Aldrich). A Nafion-115 proton-exchange membrane (20 cm^2, DuPont) separated anode chamber from the cathode.

MFCs were started up by suspending anaerobic digester sludge (sieved through 0.6-mm mesh) in sucrose-containing medium (per litre: NH_4Cl, 0.31 g; NaH_2-$PO_4 \cdot H_2O$, 5.38 g; Na_2HPO_4, 8.66 g; KCl, 0.13 g (pH 7.0) (Kim et al 2007), supplemented with trace element (12.5 mL) and vitamin (12.5 mL) solutions (Lovley et al. 1984) at a 10% volume ratio. The concentration of sucrose in the medium was 5 gL^{-1} in batch operation and 0.1 gL^{-1} in continuously fed MFCs. The MFCs were operated in batch-mode during the

initial enrichment period (approximately 2 weeks). During that time, the anodic suspension was repetitively replaced (five times) initially by mixing (1:9) anodic suspension with fresh N_2-purged sucrose-containing medium and then after 1 week by replacing the entire volume of the anodic suspension with fresh medium. MFCs were operated in batch-mode until repeatable cycles of voltage generation were observed. In continuous-mode, medium was supplied to MFCs at a flow rate of 0.18 mL min^{-1} and purged with N_2 gas. The MFCs were operated at room temperature (21–22°C). MFC voltage was monitored using an Arbin BT2000 battery tester (Arbin Instruments, USA) controlled with MITS Pro software (Arbin Instruments) across a fixed external resistance of 40 kΩ. Polarisation curves were recorded with decreasing external resistance (700 kΩ–500 Ω) and measuring the decrease in voltage.

2.2 Analysis of microbial community composition

Peak power density and community composition were analysed at 14, 28, 56, 72 and 91 days, except for MFC3 for which data was available for time points 25, 40, 54, and 75 days. To determine community composition at each time point, samples of biofilm (1 cm^2) and bacterial suspension (1 ml) were removed from the anode and the anodic chamber, respectively. DNA was directly extracted from samples using FastDNA Spin Kit for Soil (MP Biomedicals, UK). For sampling of the anode electrode, each MFC was temporarily disassembled in an aseptic environment, a 1-cm^2 anode sample cut out using a sterile scalpel and the MFC reassembled. Prior to DNA extraction, the anode suspension samples were centrifuged (10,000×g, 5 min), washed three times with 1 mL

PBS and resuspended in 100 µL of nuclease free water. For DNA extraction, samples were placed into 2.0 ml tubes containing a lysing matrix (mixture of ceramic and silica particles) and samples homogenized in the FastPrep® instrument. Following lysis, samples were centrifuged to pellet anode electrode fibres, cell debris and lysing matrix. DNA was purified from the supernatant with a silica-based procedure using filters provided with the kit. The amount and quality of DNA was measured using LabTech-Nanodrop ND100 spectrophotometer.

The partial bacterial 16S rRNA genes were amplified using the bacteria-specific forward primer 341F (*E. coli* 16S rRNA positions 341-357) (Muyzer et al 1993) and the universal reverse primer 907R (*E. coli* 16S rRNA position 907-926) (Muyzer et al 1996). A GC-clamp was added to the forward primer at the 5'-end to stabilize the melting behaviour of the DNA fragments in the DGGE (Muyzer et al 1993). PCR reaction mixtures contained 1× Taq PCR buffer (10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3 at 20ºC), 200 µM dNTP, 0.2 µM each primer, 0.025 U μL^{-1} of Taq DNA polymerase (Roche, UK), 400 ng μL^{-1} of bovine serum albumin (BSA, Fermentas, Canada) and nuclease-free water (Promega, UK) to a final volume of 50 µL, to which 1 µL of template was added. PCR was carried out using a GeneAmp PCR System 9700 (PE Applied Biosystems, USA) with the following program: 95ºC for 5 min; 25 cycles of 94ºC for 0.5 min, 50ºC for 1 min and 72ºC for 2 min; followed by final extension at 72ºC for 7 min. These bacterial community samples were analysed by denaturing gradient gel electrophoresis (DGGE). The relative species abundances in the communities were inferred from the relative band intensities calculated by dividing the

peak area of a band by the sum of peak areas of all bands in a lane (excluding chimaeras). DNA was extracted from bands for species identification by sequence analysis.

A complete description of the methods and further technical details are given elsewhere (Kim et al 2011; Beecroft et al 2012).

2.3 Statistical analysis

2.3.1 Variables

The three diversity measures (Richness, Shannon index, and Simpson index) were calculated according to the formulae presented above. Calculations were carried out for all microbial fuel cells, each with 4 or 5 time points, for both the anode biofilm and the anode compartment suspension. All diversity values were converted into an effective species number with the unit of species equivalents. This was to ensure that diversity data were in the same units, preventing the inadvertent comparison of linear to non-linear indices which can cause overestimation of correlations for non-linear indices due to reduced variance (spread). The DNA content of the biofilm (μg cm^{-2}) or the suspension (μg cm^{-3}) was included as an approximation of "total community size". Quantification of the DNA content of the biofilm is useful when comparing the effects of diversity on other variables, as larger communities could produce more electrons. Another theoretical angle on the inclusion of this variable is as an approximation of the size of the sample used to determine community diversity. Larger sampling sizes may

194 have an effect on the predictive power of the diversity estimate, and it is therefore
195 intuitively useful to consider this variable when attempting to use diversity to predict
196 community performance. Biomass was not measured directly due to the difficulty of
197 separating intact cells from the carbon fibre anode. The DNA content of a microbial
198 community has been used as a proxy for the number of cells present in the community
199 (Kubota et al 2009).
200

201 **2.3.2 Regression modelling**

202 To establish a specific relation between power (as the dependent variable) and the
203 DNA content and the diversity measures as the independent variables, a linear model
204 was used as the first approximation. The linear regression was performed using PASW
205 Statistics 18, Release Version 18.0.0 (SPSS Inc., Chicago, IL). As variables may correlate
206 with each other, diagnostics for co-linearity were also carried out for all models.
207

3. **Results and Discussion**

The communities studied are dynamic and undergo large changes in power output and diversity across the observation period. This allows the determination as to what extent diversity is associated with increased or decreased power output across all communities sampled. The use of identical fuel cell modules operating simultaneously and connected to the same measuring system allows a meaningful analysis of community development while minimising variation due to differences between the cells, or measurement equipment. A longitudinal approach allows an in depth investigation as to which of the community parameters may be associated with increases or decreases in fuel cell power output.

3.1. Bivariate relationships

In Table 1 the results of the linear bivariate correlations between each of the diversity indices and either the peak power output of the MFCs (Table 1 A) or the community age (Table 1 B) are shown.

3.1.1 Community age

The setup of microbial fuel cells involves inoculating the sterile anode chamber with cultures or environmental samples of microorganisms which may have electrogenic activity. Planktonic microbes subsequently attach to the electrode and form a biofilm; this biofilm is then able to transfer electrons to the anode. After inoculation, the initial community evolves with time (acclimation), resulting in an increase in electrical output

until a constant value is observed. It is the purpose of this work to demonstrate which biological variables are most associated with these increases in power output.

A number of changes take place as biofilm communities emerge, including increases in community diversity (Shannon: r= 0.52, p=0.013; richness: r=0.63, p=0.002; Simpson ESN: r=0.37, p=0.095), as shown in Table 1A. The richness of the suspension in the anode compartment is also observed to increase significantly with the passage of time (Richness: r=0.67, p<0.001; Shannon ESN: r=0.24, p=0.29; Simpson ESN: r=0.018, p=0.94), but this is not the case for other measures of diversity. Power output from the fuel cells also increases with time (r=0.7, p=<0.001) as does total biofilm DNA (r=0.42, p=0.05).

3.1.2 Diversity

The correlation coefficient is greater between peak power output and the Shannon ESN than for the other measures of diversity (Table 1B). The correlation between power output and the Shannon ESN is also larger (r=0.65, p<0.001) as compared with that of either the Simpson index or richness (r=0.5, p=0.018 and r=0.39, p=0.076 respectively).

Most interestingly both the richness of the biofilm and suspension were more strongly correlated with time than biofilm Shannon diversity or biofilm DNA content yet richness has no significant association with power output (Table 1A). Fig 1 shows the linear regression of community age with power and Shannon diversity of the biofilm, respectively.

Results from the bivariate correlations (Table 1A and 1B) show that while the metrics correlate, they have different correlations with the power output of anodic communities.

3.2 Relationship between suspension and biofilm diversity

Figure 2 shows the linear regression between the Shannon ESN diversity of the suspension in the anode compartment and the biofilm. A significant correlation exists between all diversities of these two populations of cells (Shannon: $r=0.65$, $p=<0.001$; Richness: $r=0.69$, $p=<0.001$; Simpson: $r=0.5$, $p=0.018$). This association is most likely to be due to migration of organisms between the two populations. One of the most striking observations which can be made from the bivariate correlations (Table 1A) is the lack of a significant correlation between the suspension diversities and power output which coincides with a strong significant positive correlation between biofilm diversities and power output.

3.3 Relationships between different measures of diversity

The different measures of biofilm diversity are strongly correlated with one and another (Figure 3); richness correlates with the Shannon ESN ($r=0.83$, $p=<0.001$) and with the Simpson ESN ($r=0.74$, $p<0.001$), and the Shannon ESN correlates with the Simpson ESN ($r=0.91$, $p=<0.001$).

3.4 Community DNA

The amount of DNA recoverable per square centimetre of biofilm is significantly associated with higher power outputs ($r=0.56$, $p=<0.007$) from MFCs (Table 1B). This is

likely to be due to denser biofilms containing larger numbers of cells and therefore greater quantities of DNA. A biofilm with a greater population of cells should have a larger metabolic activity and therefore would have a higher capacity to donate electrons to the anode.

3.5 Linear regression models

Multiple linear models are widely used in ecology, econometrics and psychology to study the interactions of a number of variables within a complex system (Pires et al 2008). Here they are used to relate the biofilm and suspension diversity for all community samples to the peak power density (expressed in W. m^{-3}) of the MFCs.

In total three regression models were constructed, all of which contain two predictor variables. The two community variables were biofilm diversity, as measured in terms of one of the three diversity measures, and biofilm DNA per cm^2; the dependent variable was power density.

The regression models were all found to be statistically significant as overall models but with large differences in the significance and association of the component variables. Table 2 summarises the overall model performance, while Table 3 lists the coefficients for the component variables. The strongest relationship with power density is seen for the Shannon index and biofilm DNA model ($r= 0.73$, $p<=0.001$). Table 3 shows the coefficients for the different component variables within the models, with the main coefficient in assessing the value of these models being the beta weight (β). A β weight of 1.0 means that the dependent variable changes one standard deviation with a one standard deviation change in the independent variable

292 in question, when all the other variables in the model are held constant. The
293 observation of significant ($p<0.01$) β weights for the biofilm Shannon index and for the
294 biofilm DNA content within the context of a significant ($p<0.001$) overall model
295 indicates that these variables together are better predictors of power output than they
296 are individually. Biofilm richness and the Simpson index fail to reach statistical
297 significance within their respective models. The only statistically significant predictor
298 outside the biofilm Shannon model is the biofilm DNA content in the richness model
299 ($\beta=0.46$, $p=0.024$). It must be noted that the failure to reach significance within the
300 context of multiple regression models for the Simpson index and the richness suggests
301 that these are too weak as predictors on their own to be significant with the number of
302 cases available for this analysis. This does not mean they are not useful, but only that
303 they are much less strongly associated with power output than the Shannon index
304 when other variables are considered. This difference is quite substantial and suggests
305 that, within the realistic data limitations of studies in this field, the Shannon index
306 represents a much more useful measure than other approaches to measuring diversity
307 as functionally relevant to power output.

308 **3.6 Microbial fuel cell performance and diversity**

309 Microbial fuel cells improved their performance progressively as communities
310 developed on the anode. In this work interactions between the diversity of the
311 suspension and the biofilm are considered with regard to their possible effects on MFC
312 power output. It is observed that the diversity of the anode biofilm has a positive
313 association with the electrogenic potential of the MFC while the diversity of the

planktonic community (i.e. suspension) has no significant effect. This may reflect the fact that only species in the anodic biofilm can participate in electron transfer, while species in suspension may divert available energy from substrate towards other metabolic activities. The association between diversity and community electricity production seems to suggest that community functioning is a determinant of electrogenic potential in microbial communities. The density of the biofilm as measured using the DNA content is also positively associated with the MFC power density. The DNA content of a microbial community is a good proxy for the number of cells present in the community (Kubota et al 2009), and DGGE is a standard, straightforward and widely used molecular method providing reliable results for the analysis of community structure in environmental samples (Dong and Reddy 2010; Ling et al. 2010) and laboratory mixed cultures (Yang et al 2012; Zhang et al 2011). A combination of PCR-DGGE and DNA content analysis has been shown to be useful to monitor changes in the composition and structure of microbial communities (Mallin and Illmer 2008).

Probably the most interesting observation is that the Shannon index of the biofilm community is a very strong predictor of power density and that other indices such as richness and the Simpson index are not significant predictors when biofilm DNA content is considered as well. It is acknowledged that more cases would improve the statistical power of this type of analysis but the case numbers used here are representative of the volume of community data that can be realistically collected for bench-top MFC setups. It is also important to note that while the statistical evidence is very compelling for an association between Shannon diversity and power output, this

work demonstrates association not causation. It is possible that a latent process which may affect power output could also cause increases in both Shannon diversity and microbial density. Further illumination of the ecology of electrogenic communities would be highly advantageous and could offer further strategies for improving MFC productivity (Gruning et al, submitted).

By demonstrating a strong statistical link between the value of the Shannon index and power output in MFCs, it is possible to infer new directions for research into electrogenic communities and to suggest an ecological basis for some published observations. Such directions would involve factors such as pH, which has been observed to increase biodiversity in unrelated systems (Fierer and Jackson 2006) or salinity, which has been linked to a decrease in microbial diversity (Wang et al 2010). It is a rule of biogeography that the larger the area of the community studied the greater will be the diversity of the organisms observed (Kostylev et al 2005). In the particular case of MFCs, those effects add scientific merit to approaching improved efficiency by raising and buffering pH and increasing the porosity and colonisable surface area of MFC anodes. Interestingly, pH has already been observed to increase fuel cell power output although diversity was not suggested as mediating the relationship (Behera et al 2009). Three-dimensional electrodes with large conductive surface areas have also been observed to produce greater power outputs for fuel cells (Di Lorenzo et al 2009). It is tempting to speculate that diversity is a major background effect unifying the interesting results reported in published experimental data.

3.7 Functional relationship between diversity and power output

A number of explanations for the association between diversity and electrogenic potential are apparent from the literature. It has been established that “ecosystem functioning” or the overall metabolic potential of a microbial community is linked to diversity (Bell et al 2005). It is explained that this occurs because each organism has a unique pattern of metabolic utilisation; the greater the diversity, the more efficiently the community can utilise the available substrates to produce energy, including those produced by other microbes. Thus greater diversity allows networks of metabolic interdependency to form within communities, further enhancing the functioning of the community. The advantage in predictive power of the Shannon index (Shannon entropy) when compared to other measures of diversity may also be caused by the degree to which these indices are affected by evenness and low abundance species. Use of richness assumes that all species contribute equally to community functioning. This is unlikely, as the evenness of the overall community will to some extent determine the possible volume of interactions between species. A community of several similarly abundant species will have more possible metabolic interactions than a community where most of the organisms are of a single species. The Simpson index very effectively captures this evenness, as mathematically it represents the probability that two randomly selected individuals in the community will be of the same species. This results in the opposite problem to richness measurements: low abundance species are given very little weight and are assumed to contribute in direct proportion to their (low) abundance. However, species at low abundance could carry out key processes within a metabolic network, and therefore may show effects disproportionate to their abundance. The value of the Shannon index is strongly

382 influenced by both community evenness among dominant species and gives low
383 abundance species greater weighting. Finally, recent work (Haegeman et al. 2013) has
384 shown that indices such as Shannon and Simpson with q=1 and q=2 are better at
385 predicting real population diversity from limited sample data. This may be a major
386 factor in the predictive value of Shannon ESN over richness. It does not by itself explain
387 the difference in the ability to predict MFC power output between the Shannon ESN
388 and Simpson ESN. This constitutes mathematical acknowledgement that low
389 abundance species can make disproportionate contributions to community
390 functioning, while acknowledging the importance of evenness amongst the dominant
391 species.

392 **4. Conclusions**

393 The Shannon diversity index of the anodic microbial consortia in a microbial fuel cell
394 shows a stronger correlation with the electrogenic potential than the correlation
395 observed with other diversity indices. The Shannon index and the DNA content of the
396 community (as an additional variable) can be used as predictors of the power output of
397 the system, obtaining a strong correlation in a regression model. The results presented
398 here demonstrate that ecological parameters such as the Shannon index can be used
399 to predict the electrogenic potential of anodophilic microbial communities to an extent
400 that has not been found with taxonomic variables.

401

402 **6. Acknowledgements**

403 This work was performed under the EPSRC Supergen Biological Fuel Cells Consortium
404 contract EP/H019480/1. AG was supported by EPSRC grant EP/I000992/1. The authors
405 would like to thank Dr M. Victoria Flores for her helpful comments and suggestions.

406

6. References

1. Beecroft, N.J., Zhao, F., Varcoe, J.R., Slade, R.C.T., Thumser, A.E. and Avignone-Rossa, C. (2012) Dynamic changes in the microbial community composition in microbial fuel cells fed with sucrose. Appl. Microbiol. Biotechnol. **93**, 423-437.

2. Behera, M. and Ghangrekar, M. M. (2009). Performance of microbial fuel cell in response to change in sludge loading rate at different anodic feed pH. Bioresour. Technol. **100**, 5114-5121.

3. Bell, T., Newman, J.A, Silverman, B. W., Turner, S. L., and Lilley, A. K. (2005). The contribution of species richness and composition to bacterial services. Nature **436**, 1157-1160.

4. Chae, K.-J., Choi, M.-J., Lee, J.-W., Kim, K.-Y. and Kim, I. S. (2009). Effect of different substrates on the performance, bacterial diversity, and bacterial viability in microbial fuel cells. Bioresour. Technol. **100**, 3518-3525.

5. Cover, T. M. and Thomas, J.A. (1991). Elements of Information Theory. New York: Wiley. ISBN 0-471-06259-6

6. Di Lorenzo, M., Scott, K., Curtis, T. P. and Head, I. M. (2010). Effect of increasing anode surface area on the performance of a single chamber microbial fuel cell. Chem. Eng. J. **156**, 40-48.

7. Dong, X. and Reddy, G. B. (2010). Soil bacterial communities in constructed wetlands treated with swine wastewater using PCR-DGGE technique. Bioresour. Technol. **101**, 1175 – 1182.

8. Fierer, N. and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. Proc. Nat. Acad. Sci. USA **103**, 626-631.

9. Gafan, G. P., Lucas, V. S., Roberts, G. J., Petrie, A., Wilson, M. and Spratt, D. A. (2005). Statistical analyses of complex denaturing gradient gel electrophoresis profiles. J. Clin. Microbiol. **43**, 3971-3978.

433 10. Grunewald, R. (2006). Assessment of damages from recreational activities on
434 coastal dunes of the Southern Baltic Sea. J. Coast. Res. **225**, 1145–1157.

435 11. Haegeman, B., Hamelin, J., Moriarty, J., Neal, P., Dushoff, J., and Weitz, J. S. (2013).
436 Robust estimation of microbial diversity in theory and in practice. ISME J. **7**, 1092-
437 1101.

438 12. Hejda, M., Pyšek, P., and Jarošík, V. (2009). Impact of invasive plants on the species
439 richness, diversity and composition of invaded communities. J. Ecol. **97**, 393–403.

440 13. Hill, M. O. (1973). Diversity and Evenness: A unifying notation and Its
441 consequences. Ecology **54**, 427 – 432.

442 14. Jost, L. (2006). Entropy and diversity. Oikos **113**, 363–375.

443 15. Keylock, C. J., and Lane, W. (2005). Simpson diversity and the Shannon Wiener
444 index as special cases of a generalized entropy. Oikos **1**, 203–207.

445 16. Kiely, P. D., Cusick, R., Call, D. F., Selembo, P.A, Regan, J. M.and Logan, B. E. (2011).
446 Anode microbial communities produced by changing from microbial fuel cell to
447 microbial electrolysis cell operation using two different wastewaters. Bioresour.
448 Technol. **102**, 388-394.

449 17. Kim, J. R., Beecroft, N. J., Varcoe, J. R., Dinsdale, R. M., Guwy, A. J., Slade, R. C. T.,
450 Thumser A., Avignone-Rossa, C. and Premier, G.C. (2011). Spatiotemporal
451 development of the bacterial community in a tubular longitudinal microbial fuel cell.
452 Appl. Microbiol. Biotechnol. **90**, 1179–1191.

453 18. Kim, J.R., Cheng, S., Oh, S.-E., and Logan, B.E. (2007). Power generation using
454 different cation, anion, and ultrafiltration membranes in microbial fuel cells. Environ.
455 Sci. Technol. **41**, 1004–1009.

456 19. Kostylev, V. E., Erlandsson, J., Ming, M. Y. and Williams, G. A. (2005). The relative
457 importance of habitat complexity and surface area in assessing biodiversity: Fractal
458 application on rocky shores. Ecol. Complex. **2**, 272-286.

459 20. Kubota, K., Ozaki, Y., Matsumiya Y. and Kubo, M. (2009) Analysis of relationship
460 between microbial and methanogenic biomass in methane fermentation. Appl.
461 Biochem. Biotechnol. **158**, 493-501.

462 21. Ling, Z., Kong, J., Jia, P., Wei, C., Wang, Y., Pan, Z., Huang, W., Li, L. Chen, H. and
463 Xiang, C. (2010). Analysis of oral microbiota in children with dental caries by PCR-DGGE
464 and barcoded pyrosequencing. Microbial Ecol. **60**, 677 – 690.

465 22. Logan, B. E., Hamelers, B., Rozendal, R., Schroder, U., Keller, J., Freguia, S.,
466 Aelterman, P., Verstraete, W. and Rabaey, K. (2006). Microbial fuel cells: Methodology
467 and technology. Environ. Sci. Technol. **40**, 5181–5192.

468 23. Lovley, D.R., Greening, R.C., Ferry, J.G. (1984). Rapidly growing rumen
469 methanogenic organism that synthesizes coenzyme M and has a high affinity for
470 formate. Appl. Environ. Microbiol. **48**, 81–87.

471 24. Mallin, C. and Illmer, P. (2008). Ability of DNA content and DGGE analysis to reflect
472 the performance condition of an anaerobic biowaste fermenter. Microbiol. Res. **163**,
473 503—511.

474 25. Muyzer, G., Dewaal, E.C., Uitterlinden, A.G. (1993). Profiling of complex microbial
475 populations by denaturing gradient gel-electrophoresis analysis of polymerase chain
476 reaction-amplified genes-coding for 16S ribosomal-RNA. Appl. Environ. Microbiol. **59**,
477 695–700.

478 26. Muyzer, G., Hottenträger, S., Teske, A., Wawer, C. (1996). Denaturing gradient gel
479 electrophoresis of PCR-amplified 16S rRNA - A new molecular approach to analyse the
480 genetic diversity of mixed microbial communities. In: Molecular Microbial Ecology
481 Manual, Kluwer Academic Publishers. Dordrecht, The Netherlands, pp. 1–23.

482 27. Nimje, V. R., Chen, C.-Y., Chen, H.-R., Chen, C.-C., Huang, Y. M., Tseng, M.-J., Cheng,
483 K.-C. and Chang, Y.-F. (2012). Comparative bioelectricity production from various
484 wastewaters in microbial fuel cells using mixed cultures and a pure strain of
485 *Shewanella oneidensis*. Bioresour. Technol. **104**, 315-323.

28. Pires, J. C. M., Martins, F. G., Sousa, S. I. V., Alvim-Ferraz, M. C. M. and Pereira, M. C. (2008). Selection and validation of parameters in multiple linear and principal component regressions. Environ. Model. Softw. **23**, 50-55.

29. Rahimnejad, M., Ghoreyshi, A. A., Najafpour, G. D., Younesi, H. and Shakeri, M. (2012). A novel microbial fuel cell stack for continuous production of clean energy. Int. J. Hydrogen Energ. **37,** 5992 – 6000.

30. Steen, A. D., Ziervogel, K. and Arnosti, C. (2010). Comparison of multivariate microbial datasets with the Shannon index: An example using enzyme activity from diverse marine environments. Org. Geochem. **41**, 1019 – 1021.

31. Sun, M., Tong, Z.-H., Sheng, G.-P., Chen, Y.-Z., Zhang, F., Mu, Z.-X., Wang, H.-L., Zeng, R. J., Liu, X.-W., Yu, H.-Q., Wei, L. and Ma, F. (2010). Microbial communities involved in electricity generation from sulfide oxidation in a microbial fuel cell. Biosens. Bioelectron. **26**, 470 – 476.

32. Tuomisto, H. (2010). A consistent terminology for quantifying species diversity? Yes, it does exist. Oecologia **4**, 853 — 860.

33. Wang, H.-Y., Bernarda, A., Huang, C.-Y., Lee, D.-J. and Chang, J.-S. (2011). Micro-sized microbial fuel cell: a mini-review. Bioresour. Technol. **102**, 235 – 243.

34. Wang, Z.-Y., Xin, Y.-Z., Gao, D.-M., Li, F.-M., Morgan, J. and Xing, B.-S. (2010). Microbial community characteristics in a degraded wetland of the Yellow River Delta. Pedosphere **20**, 466 – 478.

35. Wilson, J. G. and Brennan, M. T. (2004). Spatial and temporal variability in modelled nutrient fluxes from the unpolluted Shannon estuary, Ireland, and the implications for microphytobenthic productivity. Estuar. Coast. Shelf Sci. **60**, 193 – 201.

36. Wrighton, K.C., Virdis, B., Clauwaert, P., Read, S.T., Daly, R.A., Boon, N., Piceno, Y., Andersen, G.L., Coates, J. D. and Rabaey, K. (2010) Bacterial community structure corresponds to performance during cathodic nitrate reduction. ISME J. **4**, 1443 – 1455.

512 37. Yang, S., Du, F., and Liu, H. (2012). Characterization of mixed-culture biofilms
513 established in microbial fuel cells. Biomass Bioenerg. **46**, 531 – 537.

514 38. Zhang, Y., Min, B., Huang, L., and Angelidaki, I. (2011). Electricity generation and
515 microbial community response to substrate changes in microbial fuel cell. Bioresour.
516 Technol. **102**, 1166 –1173.

517

Table 1. Linear bivariate correlations between each of the diversity indices and the peak power output of the cells. **A**: Bivariate correlation with community age; **B**: Bivariate correlation of community variables with power density (W m^{-2}). *: Statistical significance value < 0.05

A.

	Pearson correlation coefficient (R)	Significance (p)
Community age		
Peak power	0.7	<0.001*
Biofilm Shannon ESN	0.52	<0.013*
Biofilm Simpson ESN	0.37	0.095
Biofilm Richness	0.63	0.002
Biofilm DNA (µg cm^{-2})	0.42	0.05*
Suspension Shannon ESN	0.24	0.29
Suspension Simpson ESN	0.018	0.94
Suspension Richness	0.67	0.001*
Suspension DNA (µg cm^{-3})	0.07	0.77

B.

	Pearson correlation coefficient (R)	Significance (p)
Power density		
Biofilm Shannon ESN	0.65	0.001*
Biofilm Simpson ESN	0.50	0.018*
Biofilm Richness	0.39	0.0076*
Biofilm DNA (µg cm^{-2})	0.56	0.007*
Suspension Shannon ESN	0.36	0.097
Suspension Simpson ESN	0.29	0.2
Suspension Richness	0.33	0.14
Suspension DNA (µg cm^{-3})	0.04	0.86

Table 2. Summary of multivariate models using community variables. All multiple regression models take account of biofilm DNA content (μg cm^{-2}). Variance inflation factors (VIFs) were less than 5 for all variables in all models. *: Statistical significance value < 0.05

Model	Correlation coefficient (R)	F	Significance (p)
Shannon (ESN)	0.73	10.89	<0.001*
Simpson (ESN)	0.60	5.13	<0.017*
Richness	0.64	6.45	<0.007*

530 **Table 3.** Multiple regression coefficients for the different component variables within the
531 models, assessed by the β weight. *: Statistical significance value < 0.05.

532

	Coefficients		
	β	**t**	**P**
Shannon model			
Shannon (ESN)	0.51	3.03	<0.007*
Biofilm DNA (μg.cm^{-2})	0.36	2.16	<0.044*
Simpson model			
Simpson (ESN)	0.28	1.34	0.197
Biofilm DNA (μg.cm^{-2})	0.42	2.00	0.061
Richness model			
Richness (ESN)	0.33	1.76	0.095
Biofilm DNA (μg.cm^{-2})	0.46	2.45	0.024*

533

534

535

Figure 1. Linear relationships between time (expressed as the age of the community) and: (top) Peak power output; (bottom) Shannon diversity (ESN) in the biofilm.

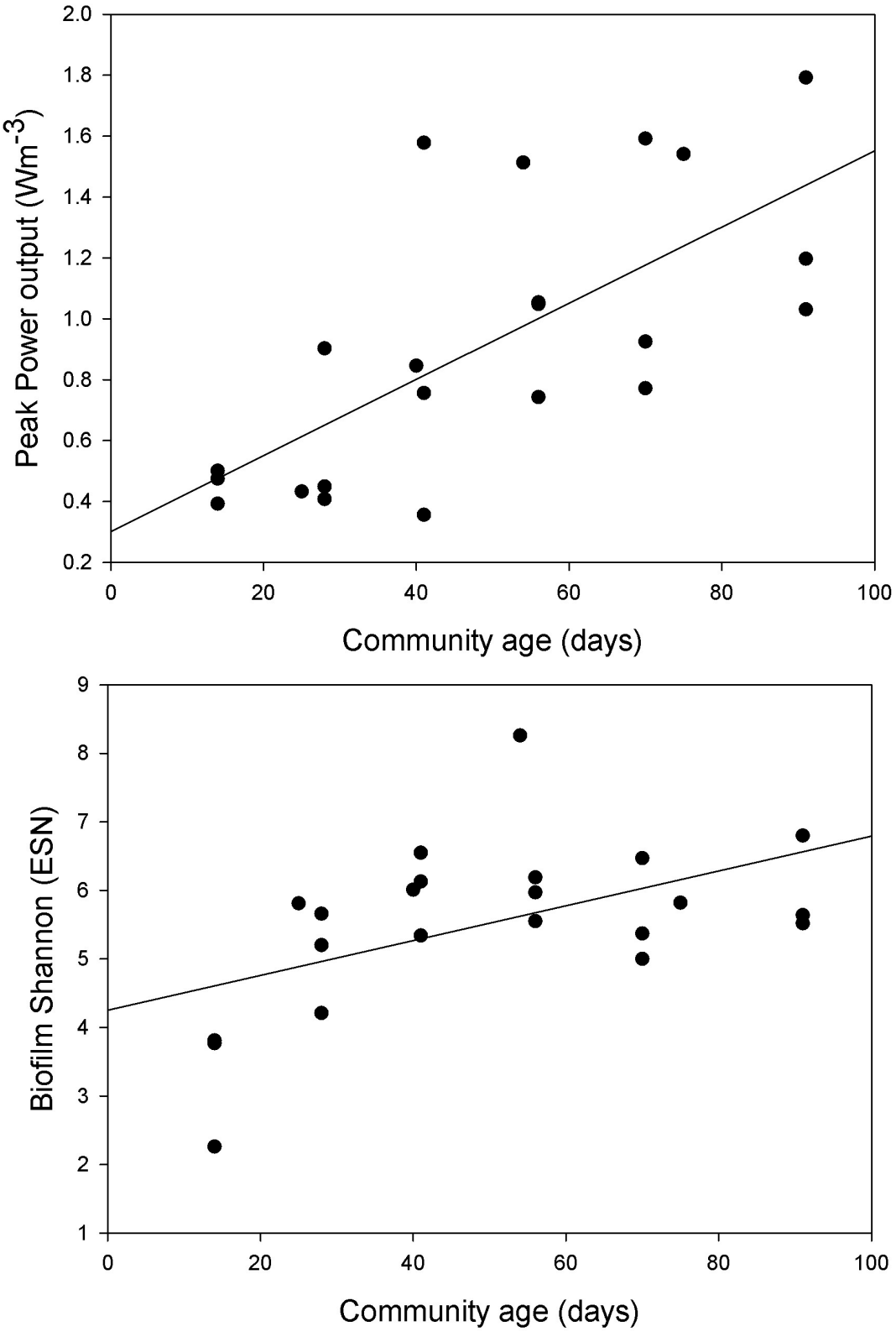

Figure 2. Linear relationship between the Shannon diversity (ESN) of the MFC anodic biofilm community and the anodic planktonic community.

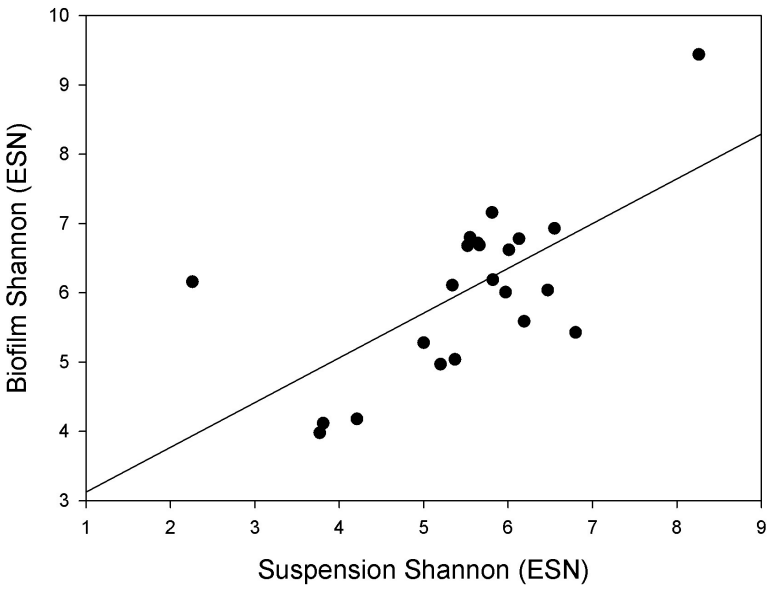

Figure 3. Linear relationships between each of the diversity indices and the other two diversity indices (left) and the peak power output of the microbial fuel cells (right). The diversity indices (Shannon entropy, Richness and Simpson Index) were calculated from DGGE data and expressed as ESN.