Electronic Supplementary Material (ESI) for Journal of Materials Chemistry A.

Sustainable Electric Power Generation from Live Anaerobic Digestion of Sugar Industry Effluents Using Microbial Fuel Cells

Bhuvanendran Revamma Sreelekshmy,[a] Rubina Basheer,[a] Savithri Sivaraman[b], Vipinlal Vasudevan,[c] Liju Elias[c] and Sheik Muhammadhu Aboobakar Shibli*[c,d]

[a]Department of Biotechnology, University of Kerala, Kariavattom Campus, Thiruvananthapuram, Kerala 695 581, India

[b]Materials Science & Technology Division, CSIR – National Institute for Interdisciplinary Science and Technology, Thiruvananthapuram, Kerala, 695 019, India

[c]Department of Chemistry, University of Kerala, Kariavattom Campus, Thiruvananthapuram, Kerala 695 581, India

[d]Centre for Renewable Energy and Materials, University of Kerala, Kariavattom Campus, Thiruvananthapuram, Kerala 695 581, India

Supplementry Information

Experimental Section

Table S1 Input parameters of each operating condition, where BL - biomass loading, IL - inoculum loading and EC - electrical conductivity of anolyte

No	Input parameters	Unit	Output parameters
1	pH	-	OCV (mV)/current density (mA m^{-2})/ power density (mW m^{-2})
2	BL	%	OCV (mV)/current density (mA m^{-2})/ power density (mW m^{-2})
3	IL	%	OCV (mV)/current density (mA m^{-2})/ power density (mW m^{-2})
4	EC	mS cm^{-1}	OCV (mV)/current density (mA m^{-2})/ power density (mW m^{-2})

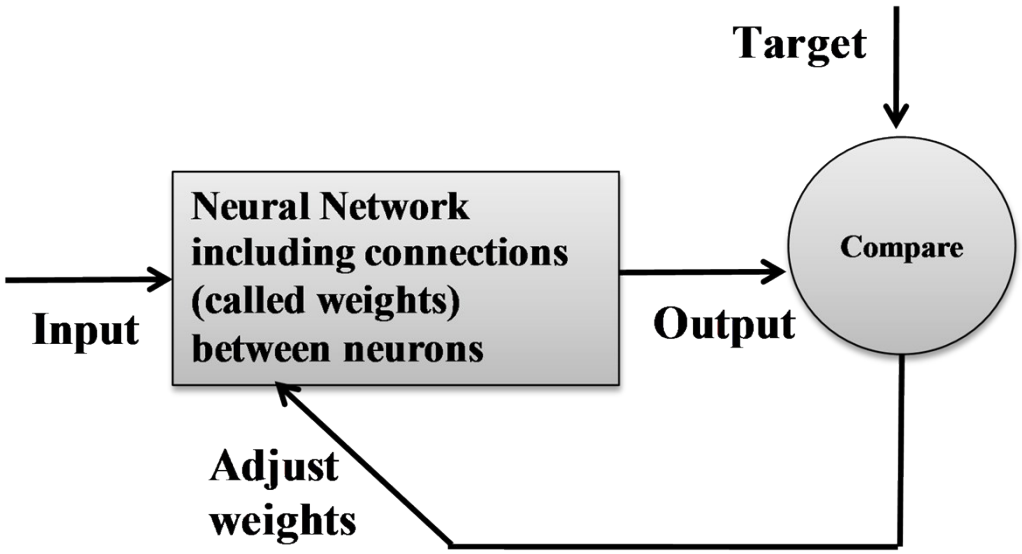

Fig. S1. The schematic representation of situation explaining the neural networks are adjusted, or trained, so that a particular input leads to a specific target output taken from MATLAB manual

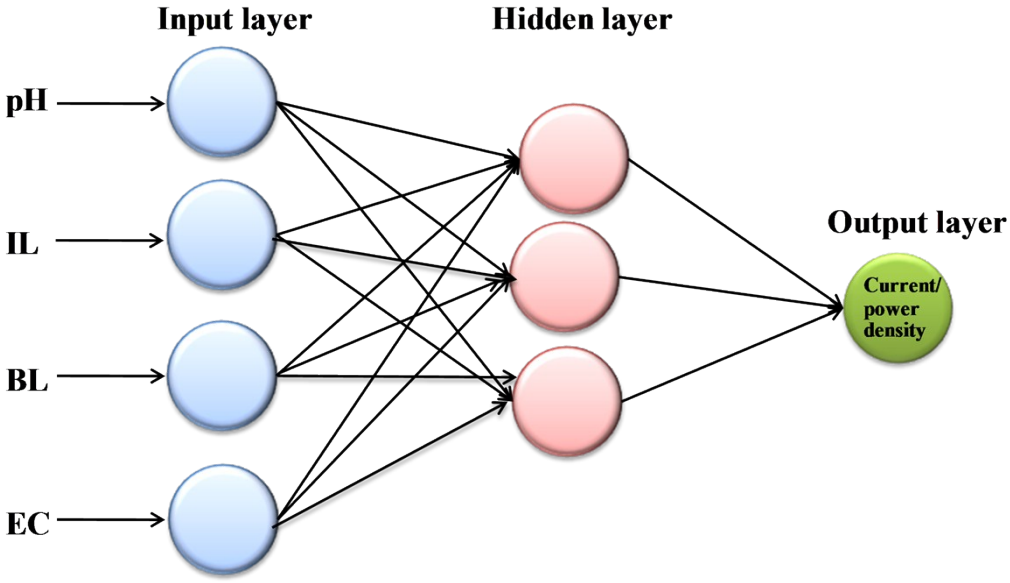

Fig. S2. A schematic layout of the ANN network containing inputs as, **IL** - percentage of inoculum loading, **BL** - percentage of biomass loading, and **EC** - electrical conductivity

16S V3 - V4 metagenome sequencing and analysis

Sequencing methodology

25 ng of DNA was used to amplify 16S rRNA hyper variable region V3 - V4. The reaction includes KAPA HiFi HotStart Ready Mix and 100 nm final concentrations of modified 341F and 785R primers.[a] The PCR involved an initial denaturation of 95 °C for 5 min followed by 25

cycles of 95 °C for 30 s, 55 °C for 45 s and 72 °C for 30 s and a final extension at 72 °C for 7 min. The amplicons were purified using Ampure beads to remove unused primers. Additional 8 cycles of PCR were performed using Illumina barcoded adapters to prepare the sequencing libraries. V3 - V4 primers used for the sequencing are shown in **Table S2**.

Table S2 V3 - V4 primers used for sequencing

Sl. no.	Primer name	Primer sequence 5' - 3'
1	V3V4F	CCTACGGGNGGCWGCAG
2	V3V4R	GACTACHVGGGTATCTAATCC

Sequence data QC

The sequence data was generated using Illumina MiSeq. Data quality was checked using FastQC [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/] and MultiQC software. The data was checked for base call quality distribution, % bases above Q20, Q30, %GC, and sequencing adapter contamination (**Fig. S3**). All the samples have passed QC threshold (Q20 > 95%).

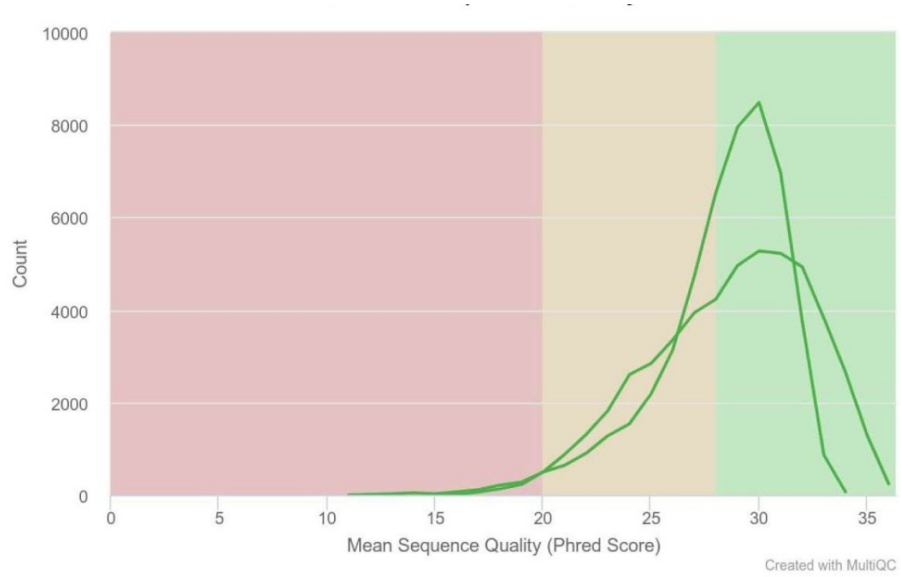

Fig. S3. Histogram of reads with average sequence quality scores

Data analysis

The reads were trimmed (20 bp) from 5' end to remove the degenerate primers. The trimmed reads were processed to remove adapter sequences and low quality bases using Trimgalore. The QC passed reads were imported into mothur and the pairs were aligned with each other to form contigs.[b] The contigs were screened for errors and only those between 300 bp and 532 bp were retained. Any contig with ambiguous base calls is rejected. The high quality contigs were checked for identical sequences and duplicates were merged. After this process, the gaps and the overhang at the ends from the contigs were removed and processed for chimera removal which may have formed due to PCR errors. UCHIME algorithm was used to flag contigs with chimeric regions.[c] A known reference of all the chimeric sequences was used to identify and remove possible chimeric sequences. The filtered contigs were processed and classified into taxonomical outlines based on the GREENGENES v.13.8 - 99 database.[d] The contigs were then clustered into OTUs (Operational Taxonomic Units). After the classification, OTU abundance was estimated. PICRUSt was used to predict gene family abundance.[e] PICRUSt program was designed to estimate the gene families contributing to a metagenome by bacteria or archaea identified using 16S rRNA sequencing. The 16S rRNA copy numbers were normalized by PICRUSt's precalculated files. The metagenomes were predicted using predict_metagenomes.py script. The predicted pathways were collapsed into higher categories. OTU contributions for the particular functions were estimated by metagenome_contributions.py script. The 16S rRNA metagenome work flow is shown in **Fig. S4**.

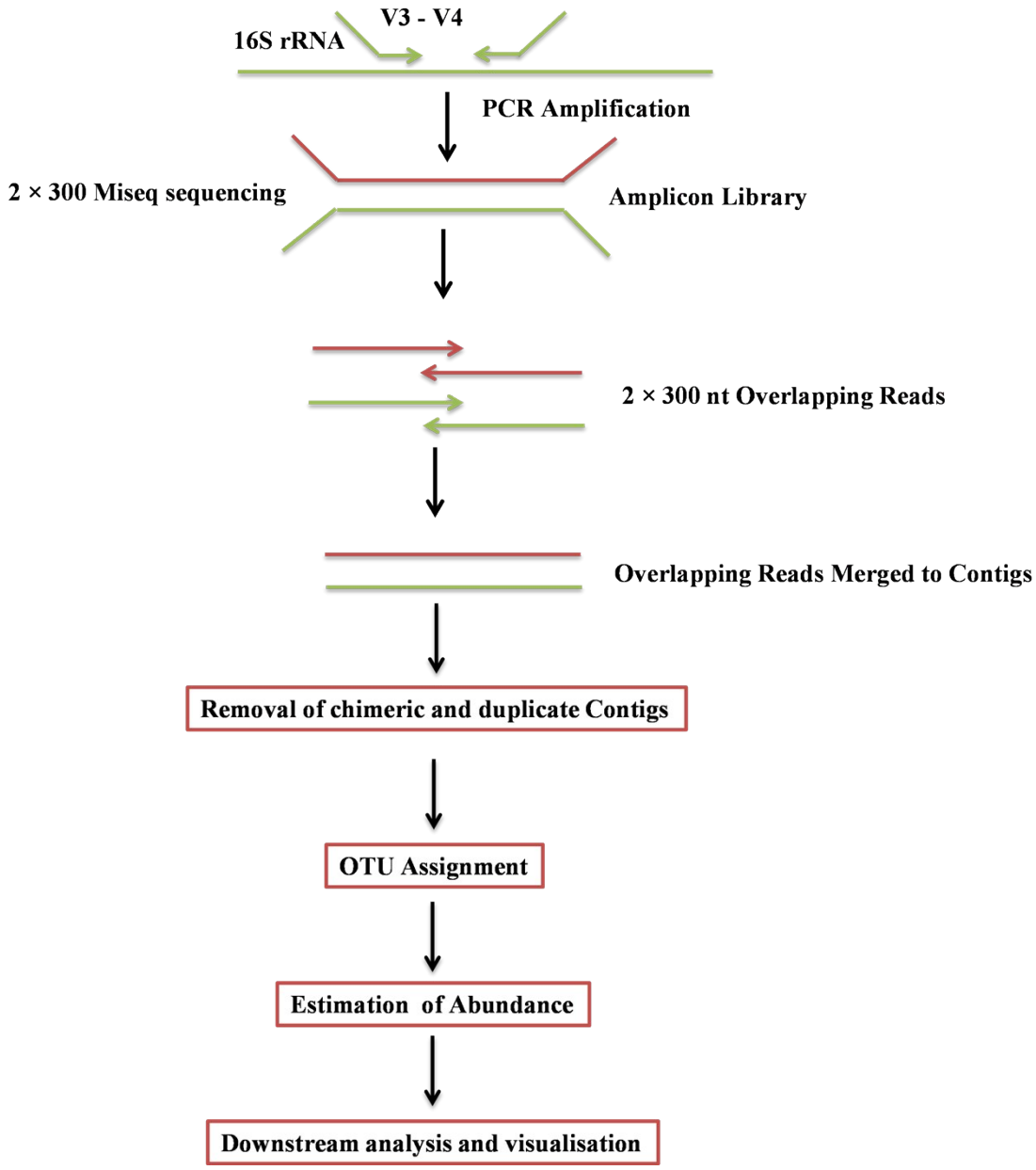

Fig. S4. The workflow of 16S rRNA metagenome

Results

Table S3 Physico-chemical characterization of sugarcane effluent before pretreatment

Parameters	Sugarcane Effluent (SE)
pH	4.6
Temperature	32.8 °C
Colour	Dark greenish brown
COD	4538 mg L^{-1}
TSS	348 mg L^{-1}
TDS	882 mg L^{-1}
TVS	829 mg L^{-1}
Total carbohydrates	18 µg mL^{-1}
Total proteins	80 µg mL^{-1}
C:N	16:3
Conductivity	10 mS cm^{-1}

Table S4 Fitted Nyquist parameters and total internal resistance of the fabricated MFCs with different ionic conductivity

Systems	R_s (Ω cm^2)	R_c (Ω cm^2)	R_a (Ω cm^2)	IR (Ω cm^2)
SE (10 mS cm^{-1})	7.47×10^2	1.83×10^3	6.29×10^2	3.20×10^3
SE (20 mS cm^{-1})	6.30×10^1	1.03×10^3	4.86×10^2	1.58×10^3
SE (30 mS cm^{-1})	1.80×10^1	9.76×10^2	3.00×10^1	1.02×10^3
SE (40 mS cm^{-1})	7.30×10^1	2.52×10^3	3.70×10^1	2.63×10^3
SE (50 mS cm^{-1})	3.01×10^2	4.37×10^3	1.52×10^2	4.82×10^3

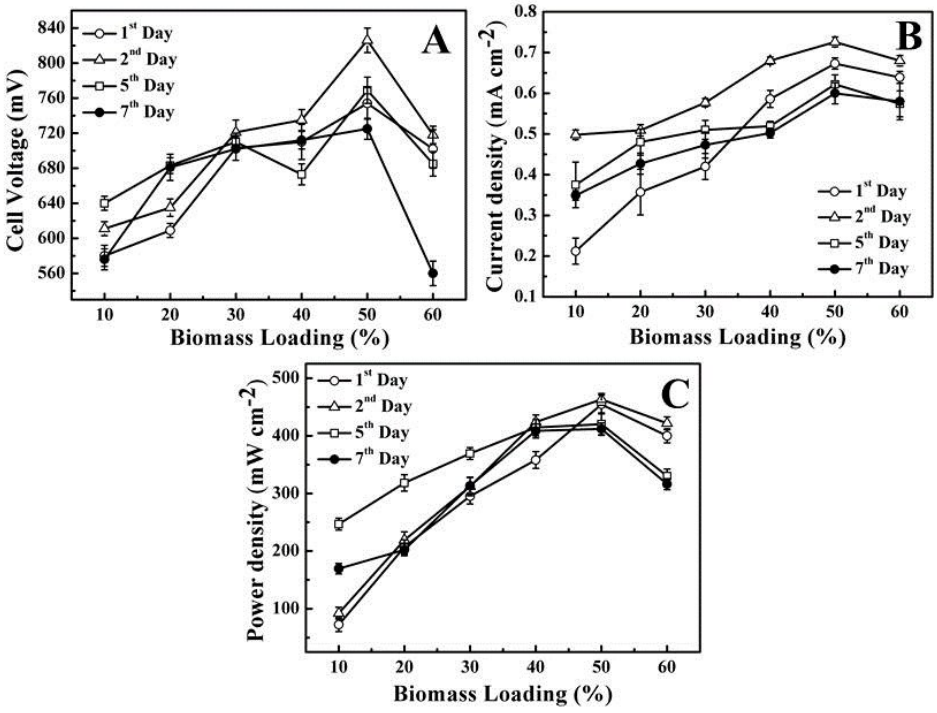

Fig. S5. (**A**) Cell voltage, (**B**) current density and (**C**) power density curve of MFCs fabricated with varying biomass loading with the course of time. Error bars represent standard deviation of four independent experiments.

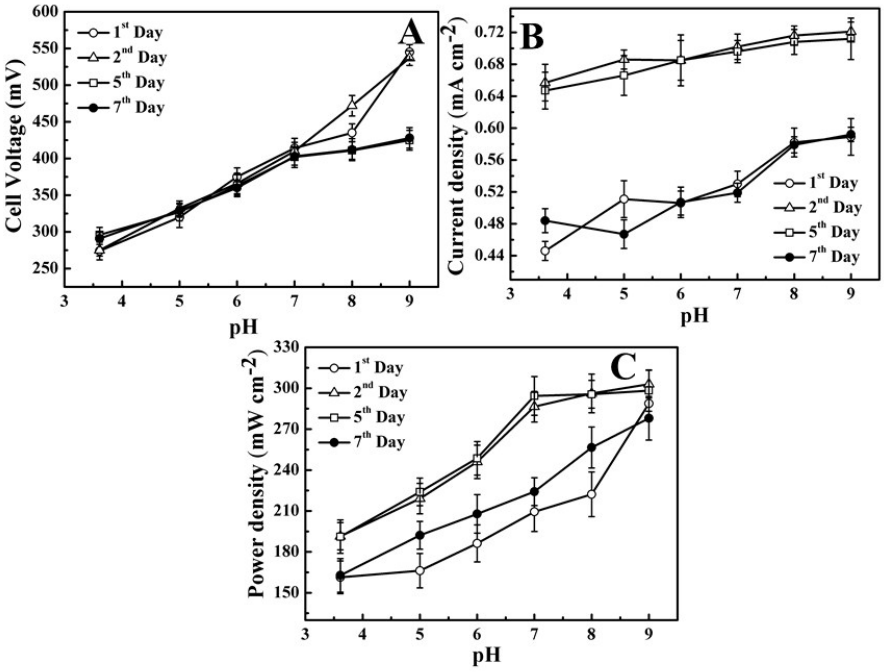

Fig. S6. (**A**) Cell voltage, (**B**) current density and (**C**) power density curve of MFCs fabricated with varying pH with the course of time. Error bars represent standard deviation of four independent experiments.

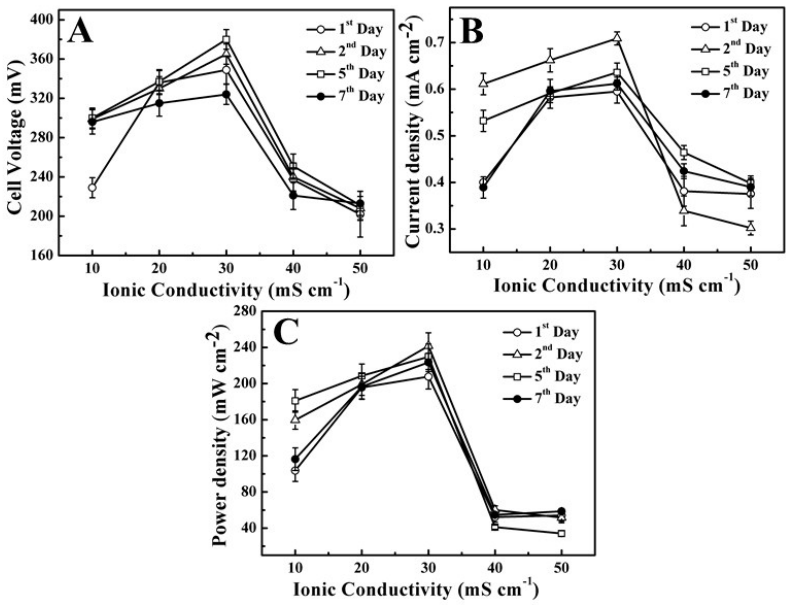

Fig. S7. (**A**) Cell voltage, (**B**) current density and (**C**) power density curve of MFCs fabricated with varying ionic conductivity with the course of time. Error bars represent standard deviation of four independent experiments.

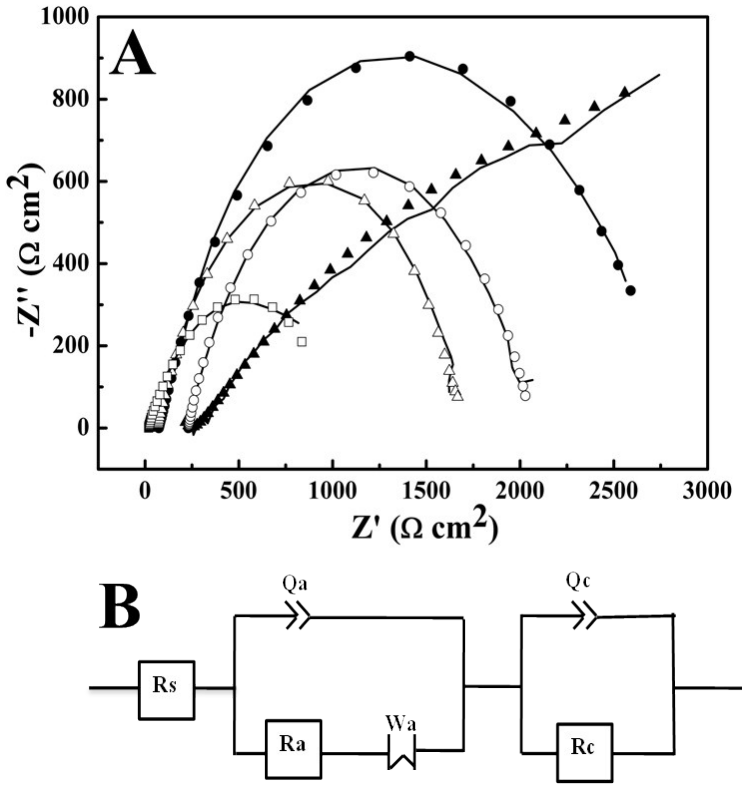

Fig. S8. (A) Fitted results of Nyquist plots of MFCs with varying ionic conductivity (-○-)10 mS cm^{-1}, (-Δ-) 20 mS cm^{-1} (-□-) 30 mS cm^{-1} (-●-) 40 mS cm^{-1} (-▲-) 50 mS cm^{-1} and Equivalent circuit used for fitting the nyquist plots of MFCs with varying ionic conductivity

ANN modelling and its validation:

The effect of interaction among the four different variables on the enhancement of MFC performance is studied using ANN. The Design of Experiments (DOE; Design expert, version 7.0.0, USA) is used for the statistical modelling of MFC.[f] The experimental design consists of 54 runs with two replicas each and is shown in **Table S5.**

Table S5 DOE for ANN modeling using four input variable and its outputs

Run	Input variables				Output variables		
	Variable 1	Variable 2	Variable 3	Variable 4	Response 1	Response 2	Response 3
	Biomass loading (%)	Inoculum loading (%)	pH	Conductivity (mS cm^{-1})	Cell voltage (mV)	Current density (mA cm^{-2})	Power density (mW m^{-2})
1	30	30	7	20	743	6.45	4792
2	30	30	8	20	728	6.30	4586
3	30	30	9	20	741	6.51	4824
4	30	30	7	25	759	6.53	4956
5	30	30	8	25	765	6.62	5064
6	30	30	9	25	745	6.44	4798
7	30	30	7	30	781	6.80	5312
8	30	30	8	30	772	6.58	5079
9	30	30	9	30	764	6.86	5241
10	30	35	7	20	707	6.22	4397
11	30	35	8	20	704	6.25	4400
12	30	35	9	20	698	6.23	4348
13	30	35	7	25	709	5.99	4247
14	30	35	8	25	707	6.31	4461
15	30	35	9	25	719	6.44	4630
16	30	35	7	30	718	6.32	4537
17	30	35	8	30	729	6.23	4541
18	30	35	9	30	747	6.52	4870

19	30	40	7	20	528	4.16	2196
20	30	40	8	20	555	4.78	2652
21	30	40	9	20	516	4.61	2378
22	30	40	7	25	606	4.64	2811
23	30	40	8	25	603	5.19	3120
24	30	40	9	25	647	5.77	3733
25	30	40	7	30	621	5.87	3645
26	30	40	8	30	660	6.10	4026
27	30	40	9	30	654	6.48	4237
28	50	30	7	20	666	4.21	2803
29	50	30	8	20	699	6.11	4270
30	50	30	9	20	729	6.41	4673
31	50	30	7	25	725	4.21	3052
32	50	30	8	25	717	6.30	4517
33	50	30	9	25	738	5.98	4413
34	50	30	7	30	732	4.29	3140
35	50	30	8	30	733	6.75	4947
36	**50**	**30**	**9**	**30**	**925**	**7.83**	**7242**
37	50	35	7	20	747	4.98	3720
38	50	35	8	20	727	6.82	4958
39	50	35	9	20	751	7.16	5377
40	50	35	7	25	731	5.67	4144
41	50	35	8	25	735	5.87	4314
42	50	35	9	25	759	6.88	5222
43	50	35	7	30	734	6.08	4462
44	50	35	8	30	745	5.16	3844
45	**50**	**35**	**9**	**30**	**776**	**7.84**	**6084**
46	50	40	7	20	502	4.53	2274
47	50	40	8	20	463	5.60	2592
48	50	40	9	20	451	6.93	3125
49	50	40	7	25	496	4.89	2425

50	50	40	8	25	510	5.86	2988
51	50	40	9	25	536	6.17	3308
52	50	40	7	30	522	4.98	2599
53	50	40	8	30	534	6.20	3310
54	50	40	9	30	547	6.00	3282

According to DOE matrix, a combination of 50% of biomass loading, 30% of inoculum loading, ionic conductivity of 30 mS cm^{-1} and pH of 9 is optimum for maximal MFC performance. Further, the concurrent effect of the variables on biofilm formation is determined using ANN and to correlate the power generation efficiency in the sugarcane effluent based MFC using anaerobic inoculum. Four input variables viz, Biomass loading, % of Inoculum added, pH, and ionic conductivity are used to feed forward neural network. Three output variables: cell voltage (mV), current density (mA cm^{-2}) and power density (mW m^{-2}) are used for validation. The MLP layer consists of an input layer of N neurons, a hidden layer of M neurons and an output layer of K neurons, the relationship between different layers can be described by equation (1),

$$g_k = F\left[\sum_{j=1}^{M} W_{kj}\, f\left(\sum_{i=1}^{N} w_{ji} x_i + \theta_j\right) + b_k\right] \quad \text{.....(1)}$$

where, j = 1, 2…M; i = 1,2….M and k = 1,2…K. In equation 1, w_{ji} is the weight connecting i^{th} neuron in the input layer to the j^{th} neuron in the hidden layer, θ_j is the bias of the j^{th} neuron in the hidden layer. Similarly, W_{kj} is the weight connecting the j^{th} neuron in the hidden layer to the k^{th} neuron in the output layer and b_k is the bias of the k^{th} neuron in the output layer.

Optimization of the topology of the ANN is the first crucial step. Of the various architectures investigated, one hidden layer with ten neurons is considered as the optimum configuration, which is based upon the regression values obtained. The dataset contains 216 input/output patterns for fitting. Out of these, 152 samples (70%) are taken for training and 32 samples (15%) each is used for validation and testing. The training data are the biggest set of data used by neural network to learn the pattern presented in the data by updating the network weights. Diagram of neural network used for the present study is shown in **Fig. S9**.

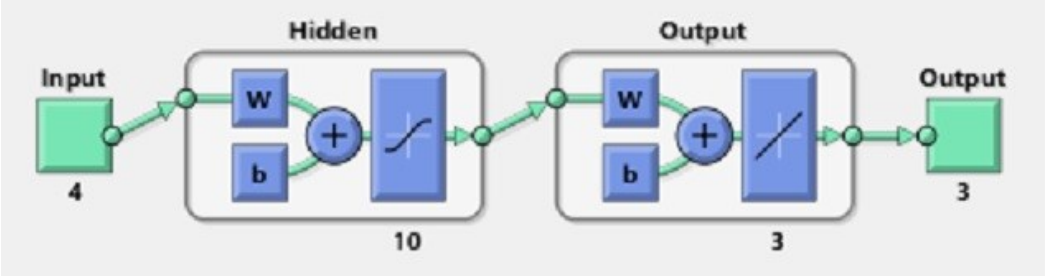

Fig. S9. Diagram of neural network used for the present study

The testing data is used for the evaluation of the quality of the network. The final performance and generalization ability of the trained network are carried out using validation data. To determine the optimum number of hidden nodes, various topologies are used, where the number of nodes in the hidden layer is varied from 2 to 12. The error function used in the present study is Root Mean Square Error (RMSE). Based on the minimum value of RMSE, we used 10 neurons in the hidden layer for the present study (**Fig. S10)**. The best regression curve obtained using 10 neurons in the hidden layer are shown in **Fig. S11**.

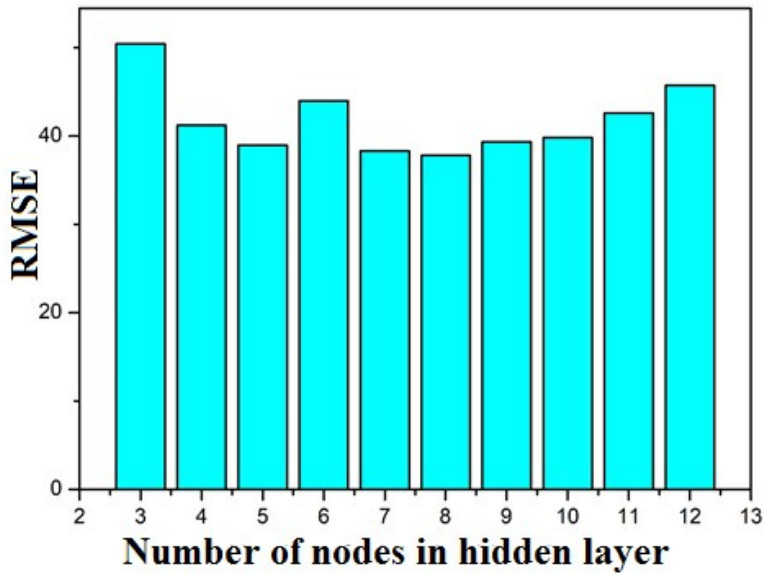

Fig. S10. The graph showing the relationship between number of neurons and root mean square error (RMSE)

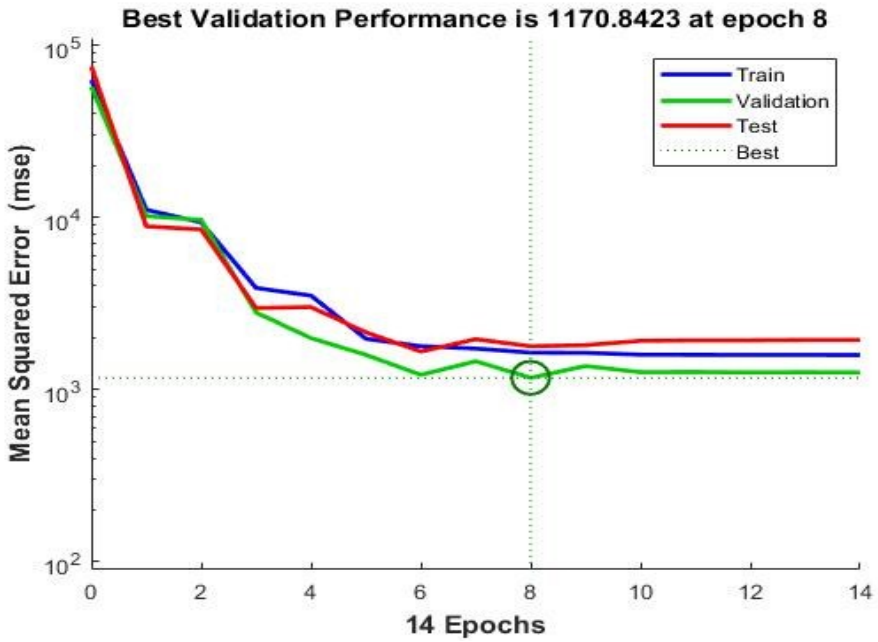

Fig. S11. Training, validation, and testing mean square error for the Levenberg – Marquardt algorithm with the best validation performance of 1170.8423 at epoch 8

The correlation coefficients for training, validation and testing for the data used in the present ANN model are 0.96063, 0.97251, and 0.96001, respectively are shown in **Fig. S12**. The training is stopped after 14 epochs as shown in **Fig. S11**, where the best performance is obtained for epoch 8 with a value of 1170.84. For the best fit, the data should fall along a 45 degree line, where the network outputs are equal to the experimental outputs. It evident from the figure that, almost all the data points fall on this line, except a few data points. **Fig. S13** shows the error histogram of the present work. The blue bars in the figure represent training data, the green bars represent validation data and the red bars represent the testing data. It is observed that, most of the errors fall between -21 and 21. But some of the training data points are outside this range. These further confirm that, neural network model can effectively reproduce the experimental results.

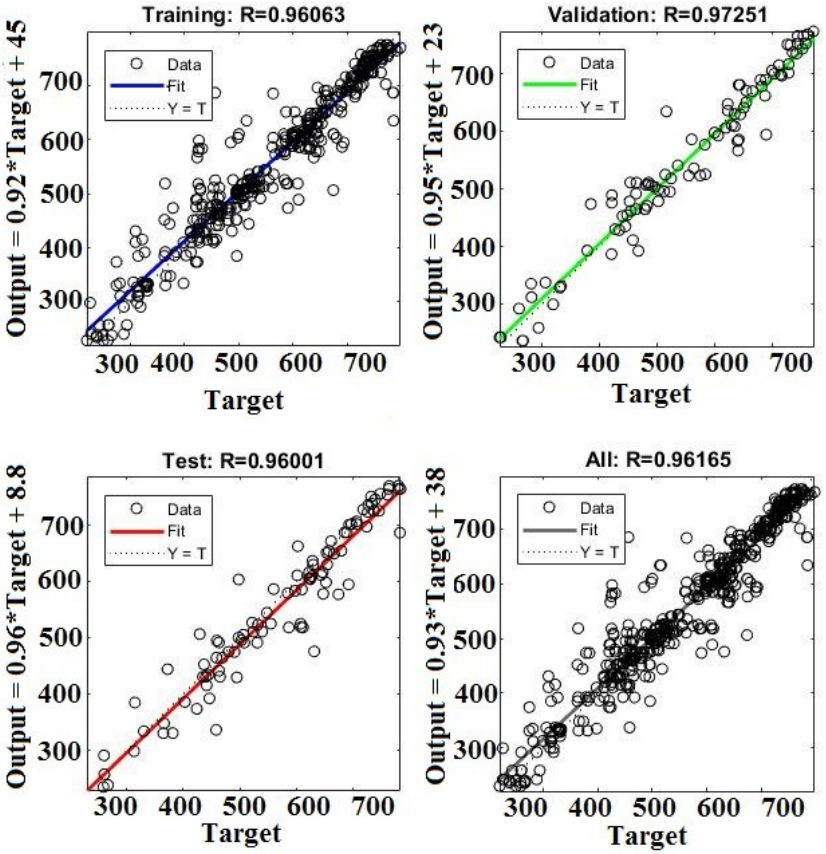

Fig. S12. Training, validation, and testing regression for the Levenberg–Marquardt algorithm

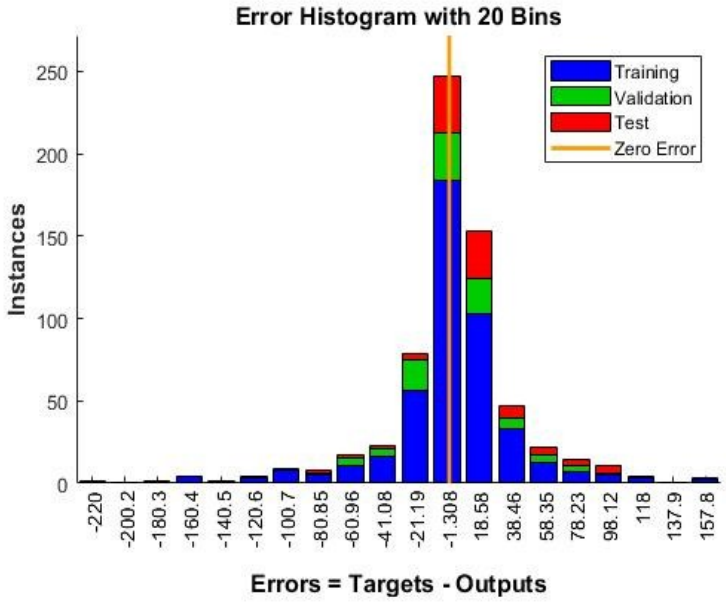

Fig. S13. The error histogram with 20 bins showing that most of the errors falls between - 21 and 21

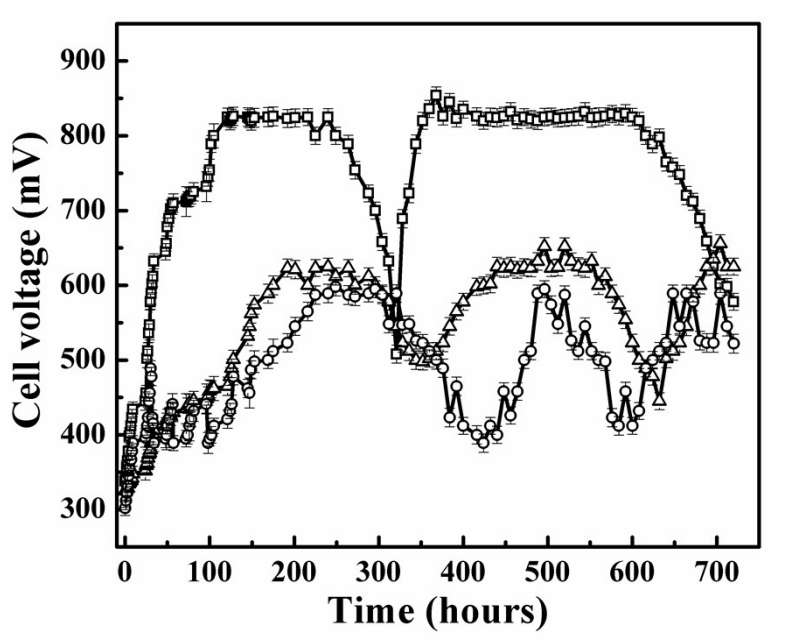

Fig. S14. Trend of open cell voltage with time of (-○-) MFC - 1, (-Δ-) MFC - 2 and (-□-) MFC–3. Error bars represent standard deviation of four independent experiments.

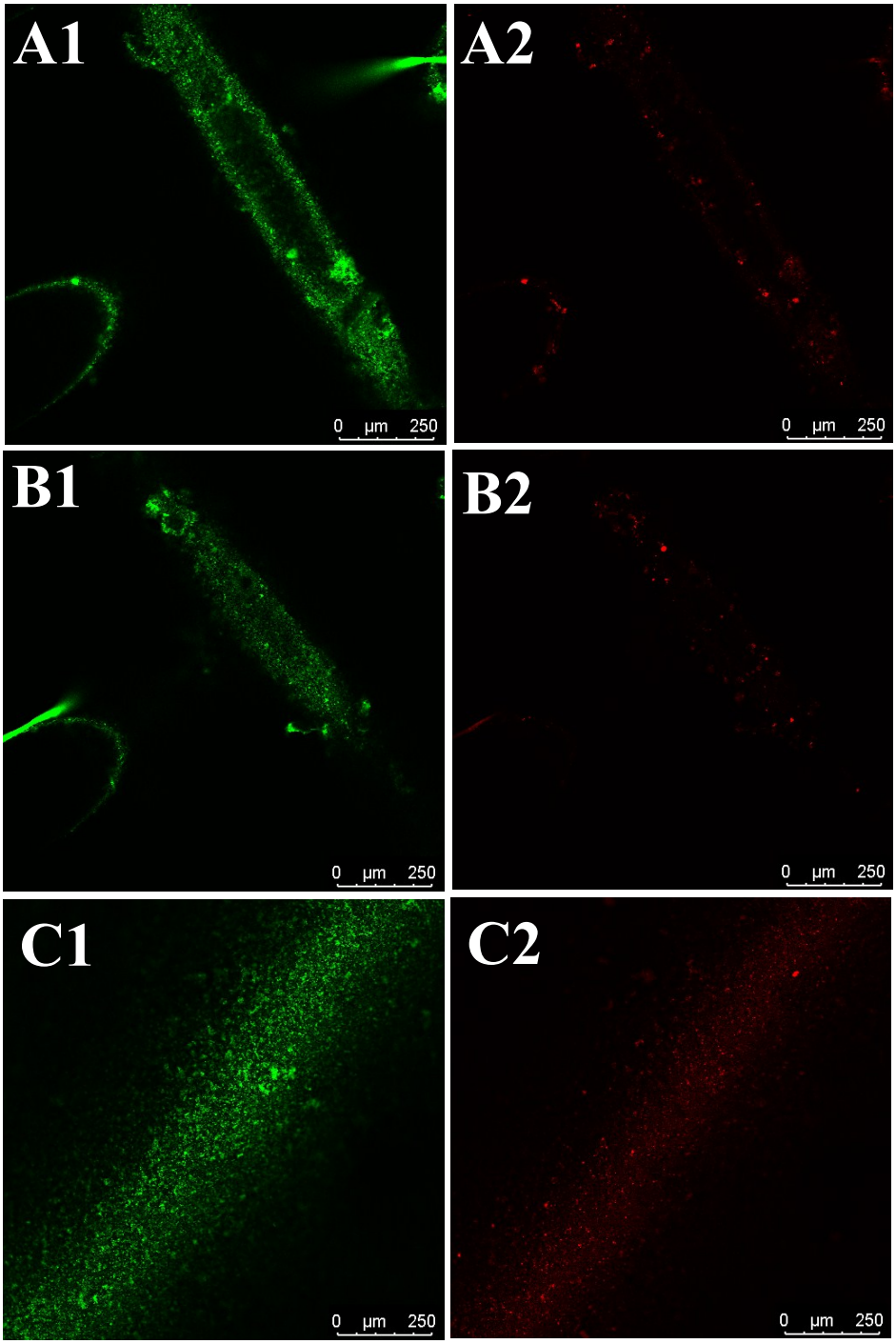

Fig. S15. CLSM images of anode in different anolyte of (**A1** and **A2**) - MFC - 1, (**B1** and **B2**) - MFC - 2 and (**C1** and **C2**) - MFC – 3

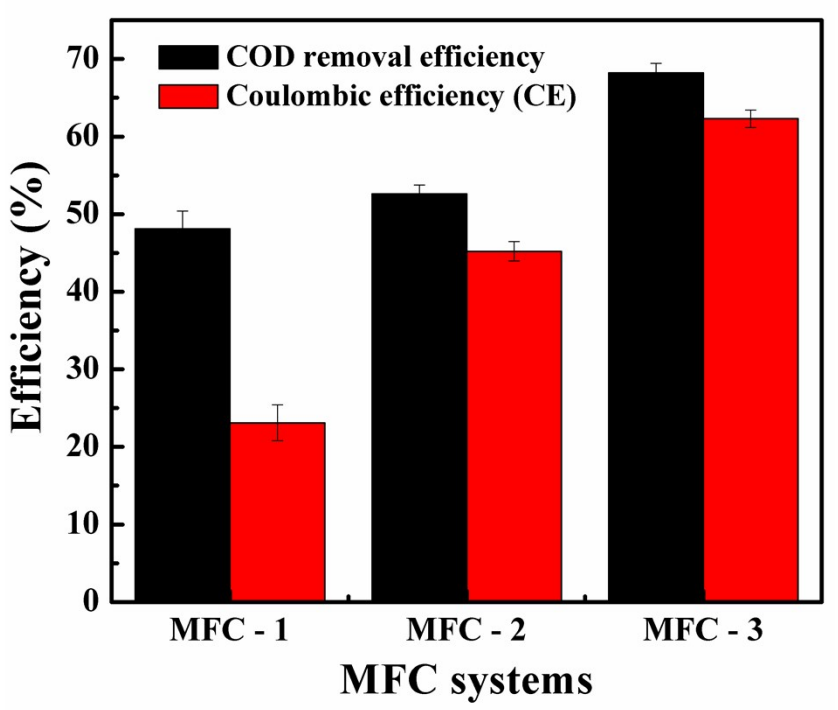

Fig. S16. Graphical representation of the COD removal efficiency and Columbic efficiency of fabricated MFCs

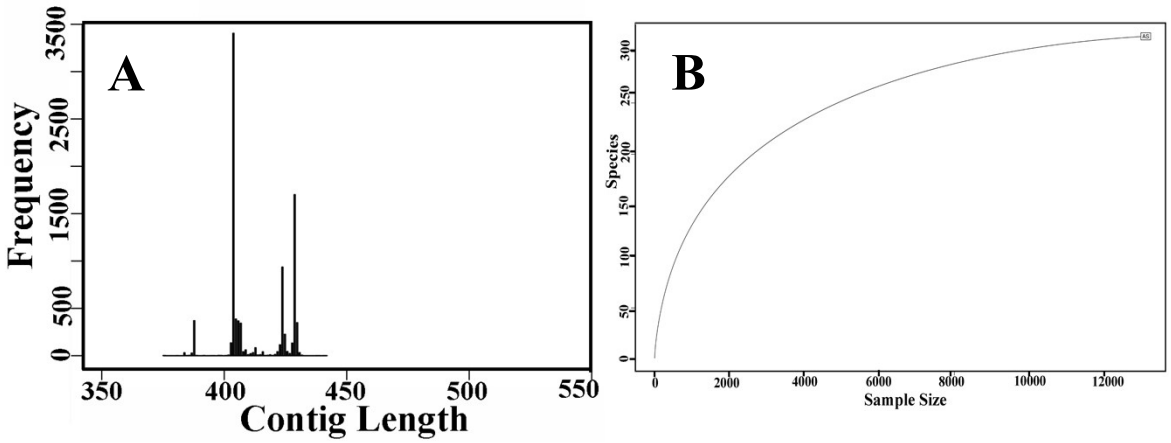

Fig. S17. (**A**) Histogram representing contig length distribution and (**B**) Rarefaction curve shows the measure of diversity that has been captured by a given number of reads in a sample

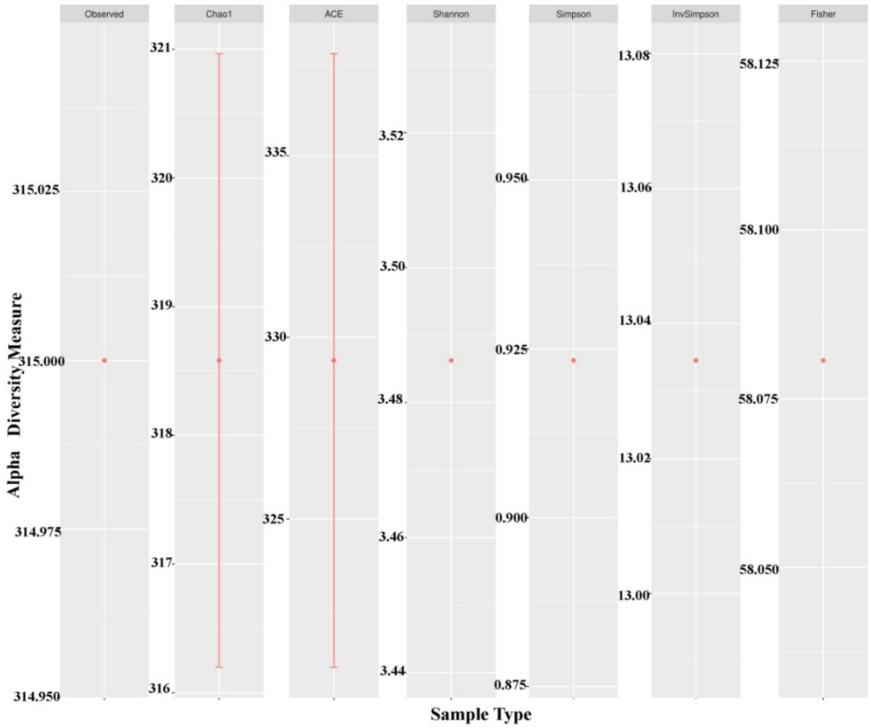

Fig. S18. Alpha diversity indices showing the richness in bacterial diversity with Chao 1 (318), ACE (329), Shannon (3.49), Simpson (0.92), InvSimpson (13.03) and Fisher (58.08)

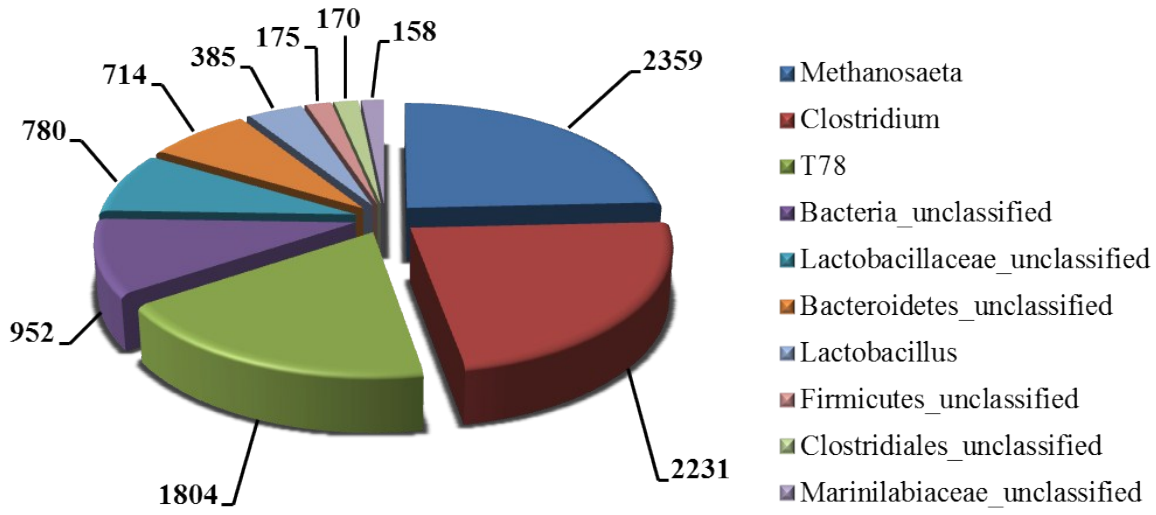

Fig. S19. Frequency of top 10 genus that present in the bacterial community of sugarcane effluent based Microbial Fuel Cell

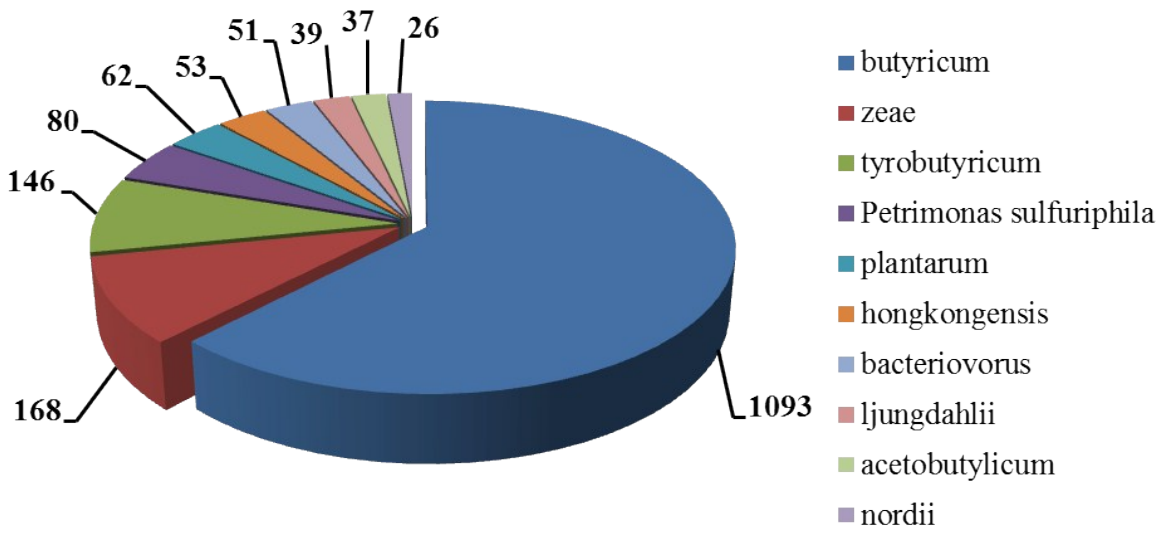

Fig. S20. Frequency of top 10 species that present in the bacterial community of sugarcane effluent based Microbial Fuel Cell

Table S6 Reagent cost for the development of anode materials

Reagents	Quantity needed	Annual cost (USD)
Fe_2TiO_5	2 g L^{-1}	0.01
NaOH	5%	0.04
HCl	3%	0.04
$SnCl_2$	10 g L^{-1}	0.09
$PdCl_2$	10 g L^{-1}	1.56
$NiSO_4$	30 g L^{-1}	0.18
$C_4H_6O_4$	25 g L^{-1}	0.14
$NaPO_2H_2$	25 g L^{-1}	0.03
Total		**2.01**

Table S7 Cost of operating and maintenance other than reagents

Items	Remark	Unit cost (USD)	Annual cost (USD)
Costs of routine maintenance	Once in a month (12 times a year)	7.02	84.28
Costs of water quality monitoring	Once in a month (12 times a year)	7.02	84.28
Total			**168.56**

Table S8 Reagent cost for catholyte

Reagents	Quantity needed	Annual Cost (USD)
NaH_2PO_4	4.4 g L^{-1}	0.04
Na_2HPO_4	3.4 g L^{-1}	0.04
$K_3[Fe(CN)_6]$	16.6 g L^{-1}	0.95
Total		**1.04**

Table S9 Cost of other accessories

Items	Cost (USD)
Double - chambered plexiglass (40.5 mL)	11.24
Mild steel (Rs. 660/m^2)	0.42
Nafion 117 (10 × 10 cm)	15.11
Gaskets (Rs. 1000/m^2)	0.70
Stainless steel	0.79
Total	**28.26**

References

a. A. Klindworth, E. Pruesse, T. Schweer, J. Peplies, C. Quast, M. Horn et al. *Nucleic Acids Res.*, 2013;41:e1.

b. P. D. Schloss, S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson,J. W. Sahl, B. Stres, G. G. Thallinger, D. J. V. Horn, C. F. Weber, *Appl. Environ. Microbiol.,* 2009, 75, 7537–41.

c. R. C. Edgar, B. J. Haas, J. C. Clemente, C. Quince, R. Knight, *Bioinformatics*, 2011;27:2194–200.

d. T. Z. DeSantis, P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, G. L. Andersen, et al., *Appl Environ Microbiol*., 2006;72(7):5069–5072

e. M. G. I. Langille, J. Zaneveld, J. G. Caporaso, D. McDonald, D. Knights, J. Reyes, J. C. Clemente, D. E. Burkepile, R. L. Vega Thurber, R. Knight, R. G. Beiko, C. Huttenhower, *Nature Biotechnology*, 2013;8:1-10.

f. Y. Sewsynker-Sukai, F. Faloye, & E. B. G. Kana, (2017), *Biotechnology & Biotechnological Equipment*, 2017;31(2): 221-235.