# Electronic Supplementary Information (ESI)

# Exploiting the real wastewater potential for resource recovery – n-caproate production from acid whey

Anna Duber, Lukasz Jaroszynski, Roman Zagrodnik, Joanna Chwialkowska, Wojciech Juzwa, Slawomir Ciesielski, Piotr Oleskowicz-Popiel

**Table S1. The characteristics of seed sludge and feedstock after delivery.**

| **Parameter** | **Seed sludge** | **Acid whey wastewater** |
|---|---|---|
| pH | 7.48 ± 0.30 | 4.70 ± 0.16 |
| Total solids (TS) [g/L] | 30.86 ± 3.85 | 58.15 ± 1.58 |
| Volatile solids (VS) [g/L] | 19.046 ± 3.69 | 50.99 ± 1.14 |
| Total chemical oxygen demand (TCOD) [g/L] | 50.86 ± 0.21 | 71.92 ± 0.21 |
| Soluble chemical oxygen demand (SCOD) [g/L] | 1.54 ± 0.04 | 70.71 ± 0.12 |
| VFAs [g/L] | 0.12 ± 0.04 | 0.75 ± 0.13 |
| VFAs [mM C] | 0.01 ± 0.00 | 27.73 ± 6.15 |
| Lactic acid [g/L] | - | 25.16 ± 1.94 |
| Lactic acid [mM C] | - | 837.93 ± 64.77 |
| Lactose [g/L] | - | 18.63 ± 1.87 |
| Lactose [mM C] | - | 653.05 ± 65.50 |
| Ethanol [g/L] | - | 0.78 ± 0.38 |
| Ethanol [mM C] | - | 33.91 ± 16.71 |

**Table S2. The operating and performance parameters for each phase of the fermentation process.**
Specificity was calculated as product-to-carboxylates production ratio in % mol C. The table contain average production rates for each HRT of each phase.

| | | Stage I | Stage II | | Stage III |
|---|---|---|---|---|---|
| **HRT [d]** | | 20 | 10 | 5 | 2.5 |
| **Days of the process [d]** | | 0 - 177 | 178 - 230 | 231 - 350 | 351 - 410 |
| **SCCAs (C2, C3, C4, C5 without lactic acid)** | **Production rate [mmol C/L/d]** | 11.25 ± 5.21 | 22.23 ± 15.60 | 52.55 ± 11.53 | 56.67 ± 31.48 |
| | **SCCA specificity [% mol C]** | 47 ± 20 | 35 ± 13 | 44 ± 7 | 31 ± 14 |
| **Caproic acid (C6)** | **Production rate [mmol C/L/d]** | 12.65 ± 5.67 | 34.03 ± 7.72 | 65.52 ± 15.76 | 119.95 ± 30.68 |
| | **n-caproate specificity [% mol C]** | 50 ± 20 | 63 ± 12 | 55 ± 7 | 70 ± 14 |
| **MCCAs (C6, C7, C8)** | **Production rate [mmol C/L/d]** | 13.16 ± 5.60 | 35.03 ± 7.72 | 66.72 ± 15.38 | 122.40 ± 30.21 |
| | **MCCA specificity [% mol C]** | 53 ± 20 | 65 ± 13 | 56 ± 7 | 69 ± 14 |

**Table S3. The microbial population biodiversity.**

| Seqs/ Sample | observed_ otus Ave, | observed_ otus Err, | observed _species Ave, | Observed _species Err, | Shannon Ave, | Shannon Err, | chao1 Ave, | chao1 Err, | Gini _index Ave, | Gini _index Err, |
|---|---|---|---|---|---|---|---|---|---|---|
| **10** | 6.264 | 1.498 | 6.264 | 1.498 | 2.331 | 0.545 | 14.24 | 10.455 | 0.895 | 0.03 |
| **8269** | 354.05 | 122.585 | 354.05 | 122.585 | 4.263 | 1.237 | 546.08 | 148.966 | 0.981 | 0.017 |
| **16528** | 463.364 | 141.717 | 463.364 | 141.717 | 4.289 | 1.24 | 649.46 | 167.224 | 0.981 | 0.016 |
| **24787** | 532.221 | 151.283 | 532.221 | 151.283 | 4.299 | 1.245 | 709.485 | 173.731 | 0.981 | 0.016 |
| **33046** | 581.743 | 158.461 | 581.743 | 158.461 | 4.305 | 1.245 | 757.595 | 183.112 | 0.981 | 0.016 |
| **41305** | 620.114 | 164.489 | 620.114 | 164.489 | 4.307 | 1.246 | 792.933 | 189.09 | 0.981 | 0.016 |
| **49564** | 651.893 | 168.682 | 651.893 | 168.682 | 4.31 | 1.246 | 820.388 | 189.146 | 0.981 | 0.016 |
| **57823** | 678.357 | 171.628 | 678.357 | 171.628 | 4.312 | 1.246 | 844.205 | 192.206 | 0.98 | 0.016 |
| **66082** | 700.964 | 175.13 | 700.964 | 175.13 | 4.314 | 1.247 | 866.28 | 197.39 | 0.98 | 0.016 |
| **74341** | nan | nan | nan | nan | nan | nan | nan | nan | nan | nan |
| **82600** | nan | nan | nan | nan | nan | nan | nan | nan | nan | nan |

**Table S4. The flow cytometry measurement of non-active, mid-active and active microbial subpopulations.**

| | | | non-active | | | mid-active | | | active | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Day** | **Sample name** | **Sample type** | **P4 #Events** | **P4 %Parent** | **P4 FITC-A Median** | **P5 #Events** | **P5 %Parent** | **P5 FITC-A Median** | **P6 #Events** | **P6 %Parent** | **P6 FITC-A Median** |
| **121** | **1** | **effluent** | 1178 | 12.9 | 410 | 2295 | 25.1 | 1267 | 4607 | 50.5 | 4580 |
| | **1_001** | | 1161 | 12.8 | 409 | 2226 | 24.5 | 1425 | 4581 | 50.4 | 4823 |
| | **1_002** | | 1136 | 12.5 | 394 | 2275 | 25.1 | 1478 | 4551 | 50.1 | 4836 |
| | **2** | **flocks** | 361 | 4.4 | 585 | 4529 | 55.6 | 1415 | 2055 | 25.2 | 10302 |
| | **2_001** | | 302 | 3.9 | 572 | 4290 | 54.7 | 1550 | 2112 | 26.9 | 10481 |
| | **2_002** | | 290 | 3.7 | 588 | 4299 | 55.2 | 1603 | 2117 | 27.2 | 10154 |
| **155** | **C** | **effluent** | 347 | 3.9 | 493 | 899 | 10.1 | 995 | 6087 | 68.1 | 9182 |
| | **C_001** | | 321 | 3.6 | 439 | 899 | 10.1 | 1373 | 5918 | 66.6 | 8379 |
| | **C_002** | | 316 | 3.4 | 411 | 1022 | 11.1 | 1495 | 5183 | 56.5 | 9241 |
| | **O** | **flocks** | 229 | 2.9 | 468 | 839 | 10.5 | 995 | 5152 | 64.3 | 9968 |
| | **O_001** | | 207 | 2.6 | 469 | 896 | 11.1 | 1034 | 5137 | 63.9 | 10155 |
| | **O_002** | | 232 | 2.8 | 428 | 935 | 11.1 | 1022 | 5436 | 64.7 | 10397 |
| **175** | **C** | **effluent** | 167 | 2.1 | 461 | 563 | 7 | 1133 | 5906 | 73.5 | 17160 |
| | **C_001** | | 220 | 2.7 | 501 | 625 | 7.8 | 1201 | 5688 | 70.9 | 16950 |
| | **C_002** | | 282 | 3.2 | 515 | 614 | 6.9 | 1233 | 6080 | 67.9 | 16613 |
| | **O** | **flocks** | 427 | 5.4 | 825 | 1185 | 14.8 | 2245 | 4330 | 54.3 | 19842 |
| | **O_001** | | 391 | 4.9 | 797 | 1156 | 14.5 | 2495 | 4386 | 54.9 | 20778 |
| | **O_002** | | 362 | 4.6 | 751 | 1224 | 15.4 | 2643 | 4266 | 53.7 | 20516 |
| **230** | **C** | **effluent** | 418 | 4.6 | 667 | 324 | 3.6 | 3254 | 6596 | 72.7 | 21591 |
| | **C_001** | | 384 | 4.2 | 650 | 330 | 3.6 | 3851 | 6744 | 74.4 | 20826 |
| | **C_002** | | 396 | 4.4 | 605 | 340 | 3.8 | 4185 | 6672 | 73.8 | 20055 |
| | **O** | **flocks** | 173 | 2 | 895 | 724 | 8.4 | 3065 | 5932 | 68.9 | 24137 |
| | **O_001** | | 155 | 1.8 | 866 | 795 | 9.3 | 3285 | 5795 | 67.5 | 24334 |
| | **O_002** | | 189 | 2.2 | 878 | 780 | 9 | 3159 | 5810 | 67 | 24091 |
| **237** | **C** | **effluent** | 128 | 1.5 | 350 | 167 | 1.9 | 3303 | 7673 | 87.6 | 11752 |
| | **C_001** | | 135 | 1.5 | 358 | 143 | 1.6 | 3195 | 7595 | 86.7 | 11672 |
| | **C_002** | | 106 | 1.2 | 368 | 153 | 1.7 | 3057 | 7579 | 86.5 | 11686 |
| | **C_003** | | 93 | 1.1 | 340 | 162 | 1.9 | 3054 | 7509 | 86.9 | 11644 |
| | **O** | **flocks** | 89 | 1 | 370 | 177 | 2.1 | 2777 | 7606 | 88.2 | 13212 |
| | **O_001** | | 96 | 1.1 | 353 | 170 | 1.9 | 2860 | 7785 | 88.9 | 13354 |
| | **O_002** | | 100 | 1.2 | 378 | 184 | 2.1 | 2988 | 7672 | 88.3 | 13328 |
| | **O_003** | | 85 | 1 | 363 | 174 | 2 | 3261 | 7660 | 88.4 | 13165 |
| **292** | **C** | **effluent** | 290 | 3.4 | 271 | 612 | 7.3 | 423 | 5785 | 68.8 | 5165 |
| | **C_001** | | 295 | 3.5 | 266 | 655 | 7.7 | 432 | 5792 | 68 | 4997 |
| | **C_002** | | 128 | 1.7 | 401 | 259 | 3.4 | 604 | 4318 | 56.8 | 8933 |
| | **C_003** | | 157 | 2.1 | 325 | 300 | 4 | 537 | 4366 | 58.2 | 8063 |
| | **C_004** | | 130 | 1.7 | 405 | 217 | 2.8 | 689 | 4530 | 57.6 | 8504 |
| | **C_005** | | 181 | 2.5 | 267 | 400 | 5.5 | 453 | 4655 | 63.9 | 6630 |
| | **O** | **flocks** | 217 | 2.7 | 262 | 618 | 7.7 | 366 | 5962 | 73.8 | 6411 |
| | **O_001** | | 280 | 3.3 | 279 | 715 | 8.3 | 359 | 6422 | 74.8 | 6008 |
| | **O_002** | | 251 | 2.9 | 265 | 710 | 8.3 | 368 | 6460 | 75.2 | 5692 |
| | **O_003** | | 256 | 2.9 | 272 | 734 | 8.4 | 371 | 6517 | 74.4 | 5908 |

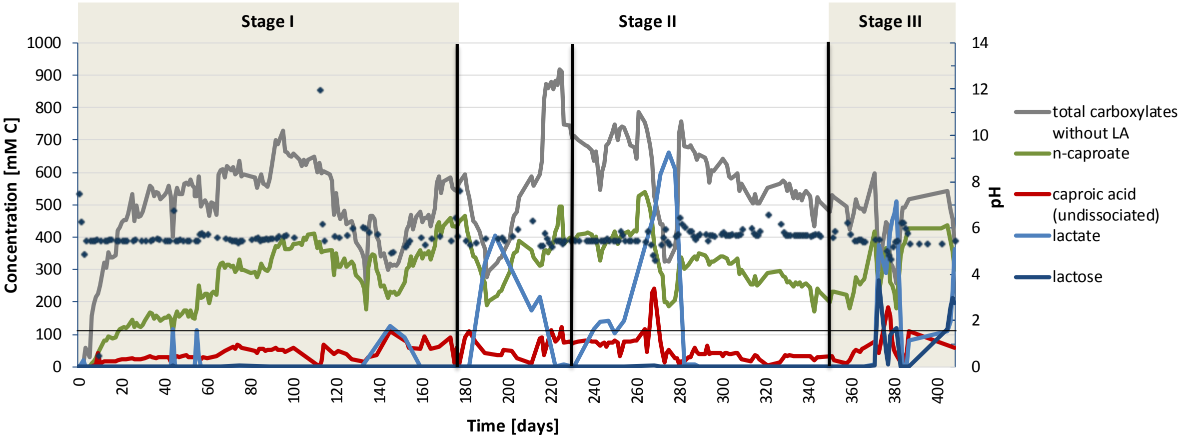

**Fig. S1 Concentration of total carboxylates, n-caproate (dissociated form), caproic acid (undissociated form) lactate and lactose as well as pH along the fermentation process.**

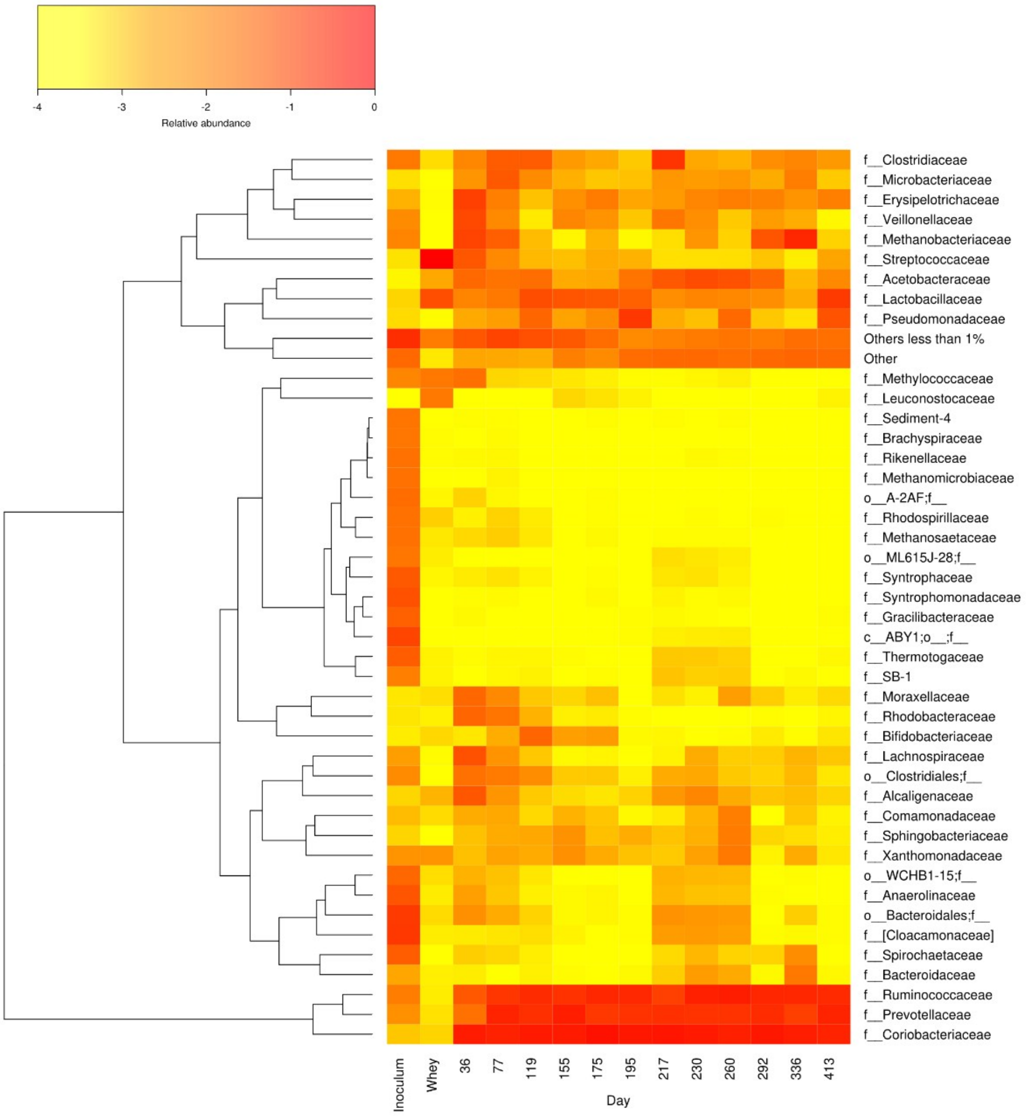

**Fig. S2 Heatmap of relative abundance at family level.** The relative abundance were calculated as log for better visualisation and clustered based on the Bray-Curtis dissimilarity index.

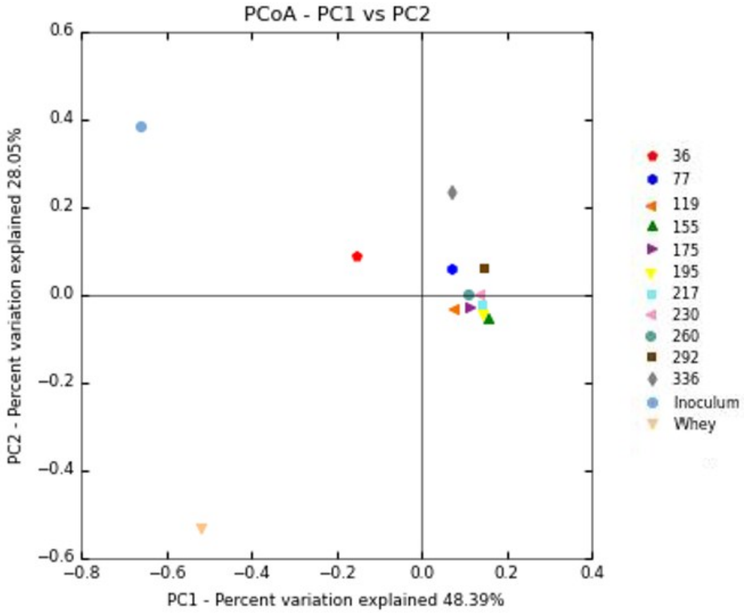

**Fig. S3 PCoA analysis of the microbial society of inoculum, feedstock and samples taken throughout bioreactor operation**. The PCoA analysis was performed based on the weighted UniFrac algorithm. The plot is based on the first two principal coordinates (PC1 and PC2) which explain 48.39% and 28.05% of overall phylogenetic variation, respectively.

PCoA analysis was used to determine the dissimilarity between the community of the inoculum, feedstock and samples taken throughout bioreactor operation based on the weighted UniFrac algorithm. Weighted UniFrac is a measure of the pairwise dissimilarity in community composition which takes into account the relative abundance of different members of the microbial community. The distances are visualised based on the first two principal coordinates (PC) that collectively explained 76.44% of overall phylogenetic variations between samples (PC1 explains 28% of the overall phylogenetic variation, while PC2 explains 48.4%). Inoculum and ACW samples were distanced from each other and from the reactor samples significantly. Sample from day 36 (beginning of the process) diverged from inoculum and feedstock, however it was slightly separated from the other reactor samples on the PCoA plot. Samples from days 77 to 292 were grouped together. Sample from day 336 was separated along the PC2 making it closer to the inoculum as at that time methanogens appeared in the reactor. The results showed similar bacterial community in the reactor microbiome along the process with small spacial and temporal variations indicating a quite stable microbial population.