Appendix B

Supplementary information for Chapter 3:
Structure and selection in an engineered symbiotic biofilm consortium

B.1 Details of construction of strains MGd- and MGfwc-

Strains MGd- and MGfwc- were constructed by recombination with the lambda red recombinase plasmid pKD46, as outlined in [150]. The chromosomal inserts to replace *dapD, csgG–csgC*, and *wza–wcaL* were all constructed by PCR with template plasmid pKD4.

The primers used were:

dapD-P1-fwd: 5'-
ATGCAGCAGTTACAGAACATTATTGAAACCGCTTTTGAACGCCGGTGTAGGC
TGGAGCTGCTTC
and dapD-P2-rev: 5'-
TTAGTCGATGGTACGCAGCAGTTCGTTAATGCCGACTTTGCCGCATATGAATA
TCCTCCTTA;
csgG-P1-fwd: 5'-
TCAGGATTCCGGTGGAACCGACATATGGCGGTATTTCACCAGAATGTCATGT
GTAGGCTGGAGCTGCTTC
and csgC-P2-rev: 5'-
TTAAGACTTTTCTGAAGAGGGCGGCCATTGTTGTGATAAATGAAGTGACTGC
ATATGAATATCCTCCTTA;

wza-P1-fwd: 5'-
ATGATGAAATCCAAAATGAAATTGATGCCATTATTGGTGTCAGTAACCTTGTG
TAGGCTGGAGCTGCTTC
and wcaL-P2-rev: 5'-
CTATAAAGCCTGCAGCAAGCTGGCGAGTTCTCGATTGATCACCTGCTGGTCAT
ATGAATATCCTCCTTA

Recombinant clones were selected with 50 $\mu g ml^{-1}$ kanamycin, cured at 42°C, and tested with colony PCR reactions using internal primers to confirm the presence of the kanamycin resistance gene and absence of the target genes. Plasmid pCP20, containing the Flp recombinase, was transformed into cells containing successful kanamycin inserts to remove the inserts [162]. Finally, clones were again cured at 42°C to remove pCP20, and the same colony PCR reactions with internal primers were repeated to confirm the deletions of target genes and of the kanamycin resistance gene insert.

B.2 Stability of the engineered symbiotic biofilm

The blue and yellow populations coexisted for periods of up to 288 days in the biofilm environment (experiments were terminated after that, and imaging data was not collected past 178 hours). Biofilm sloughing and growth phases occurred at repeatable times from experiment to experiment, even when experiments were conducted months apart (Figure B1). Repeatability over the first 120 hour time period—the length of time for which we obtained the most independent replicates—was confirmed with analysis of variance (ANOVA), $F_{0.05}(3,28) = 0.47 < F_{crit} = 2.947$, $P = 0.70$. This demonstrates that the biofilm environment we created provides repeatable results over at least a period of 120 hours, which is longer than necessary to validate all quantitative comparisons that we make between generations. However, some weeks all biofilms accumulated more total biomass than other weeks (see November versus December in Figure B1) which was probably due to changes in incubation temperature, which was difficult to precisely control. The relationships between relevant biofilms were repeatable from week to week even if total biomass accumulation was not equivalent.

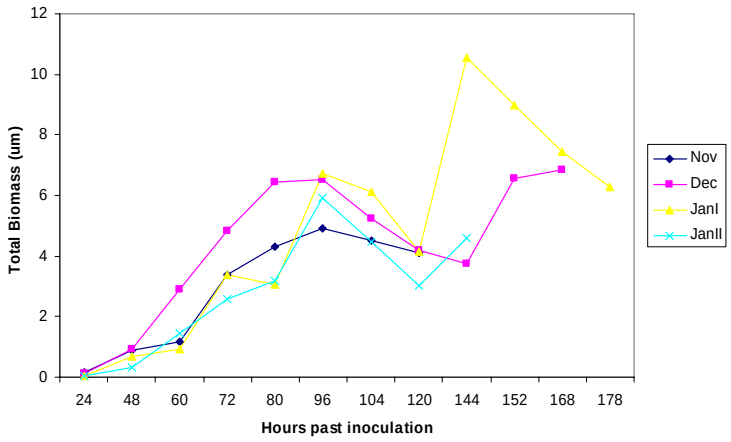

Figure B1 Mean total biomass in µm, with respect to hours past inoculation, for independent symbiotic biofilms cultured months apart. These data reveal that biomass accumulation, including timing of growth and dispersal phases, of the symbiotic ecosystem is very repeatable between independent biological replicates ($P = 0.7$).

B.3 Metrics and COMSTAT image processing

Image stacks were pre-processed with the COMSTAT script LOOK, the background cutoff threshold was set at 15 for all images, and any images still containing background auto-fluorescence of the substrate were deleted from the stack (thus, any errors in the measured values will be too low, rather than too high). Options chosen for processing in COMSTAT were 1 (biomass calculation) and 24 (no connected volume filtration, because yellow and blue populations were intermingled in the biofilm but were captured in parallel image stacks, so not all clusters captured in each single-channel stack were connected to the substrate). We also edited the code for option 1—the biomass calculation—to generate the biomass median as well as biomass totals for each layer in the stack (allowing us to confirm total biomass and biomass median calculations). The COMSTAT code that we used to generate the biomass totals and medians is below.

```
function [y,bmp,hs]=biomass_func(bb,xyarea,voxel)
count=0; hs=0; layersum=0; bmpcount=0; halfsum=0;

for side=1:size(bb,3)
   loc1=bb(:,:,side)>0;
   count=count+sum(sum(loc1));
end

y=count*voxel/(xyarea*size(bb,1)*size(bb,2));
halfcount=count/2;

for side=1:size(bb,3)
   loc1=bb(:,:,side)>0;
   layersum=sum(sum(loc1));
   bmp(side)=(layersum/count)*100;
end

for side=1:size(bb,3)
   loc1=bb(:,:,side)>0;
   halfsum=halfsum+sum(sum(loc1));
   if halfsum>=halfcount
       hs=side;
       break;
   end
end
```

B.4 Repeatability of growth advantage in untreated second-generation biofilm

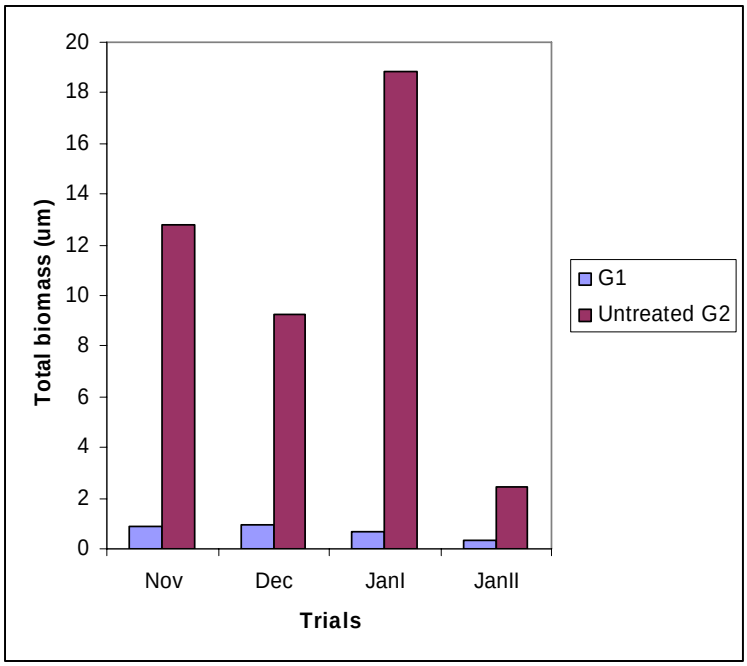

Figure B2 The untreated second-generation biofilms (Untreated G2) always exhibited a growth advantage over the first-generation biofilms (G1). The second-generation consortia formed biofilms far more quickly: on average, after 48 hours of growth, second-generation consortia accumulated 15 times the biomass accumulated by first-generation consortia in their first 48 hours (ANOVA $F_{0.05}(3,4) = 0.358 < F_{crit} = 6.591$, $P = 0.78$). In one example (JanI) the untreated second generation accumulated 30 times the biomass of the first generation.

Not only did the second-generation biofilms accumulate biomass more quickly than the first generation, as seen in Figure B2, but second-generation biofilms also accumulated more total biomass. We first averaged the highest ever recorded total biomasses from four separate first-generation biofilms (regardless of the time point). We then averaged total biomass after 48 hours of second-generation growth from the same four separate trials. By comparing these two averages we found that, on average, the second generation accumulated two times more biomass than the first generation ever accumulated in its lifespan.

B.5 Images of treated and untreated biofilms at the substrate

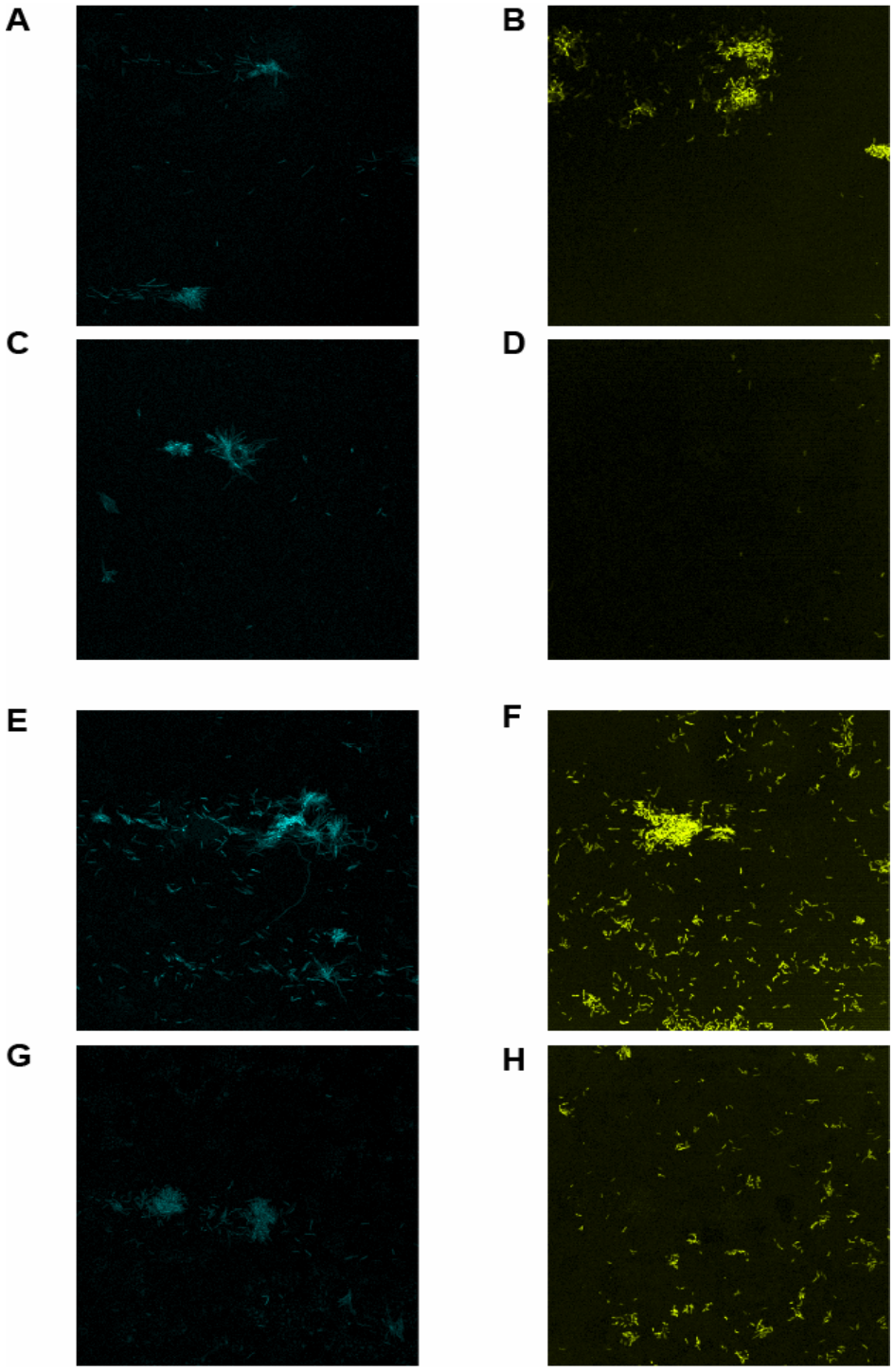

Figure B3 (A) and (B) Blue and yellow biomass is co-localized at the substrate in untreated 12-hour second-generation (G2) biofilms. (C) and (D) Single and very small clusters of yellow cells are distributed across the substrate in 12-hour treated G2 biofilms, but blues are still found in clusters. (E) and (F) In the untreated G2 case, blue and yellow continue to be co-localized and to grow well at 24 hours. (G) and (H) In the treated G2 case, blues appear less healthy and yellows are distributed across the substrate in small clusters rather than being co-localized with blues after 24 hours.

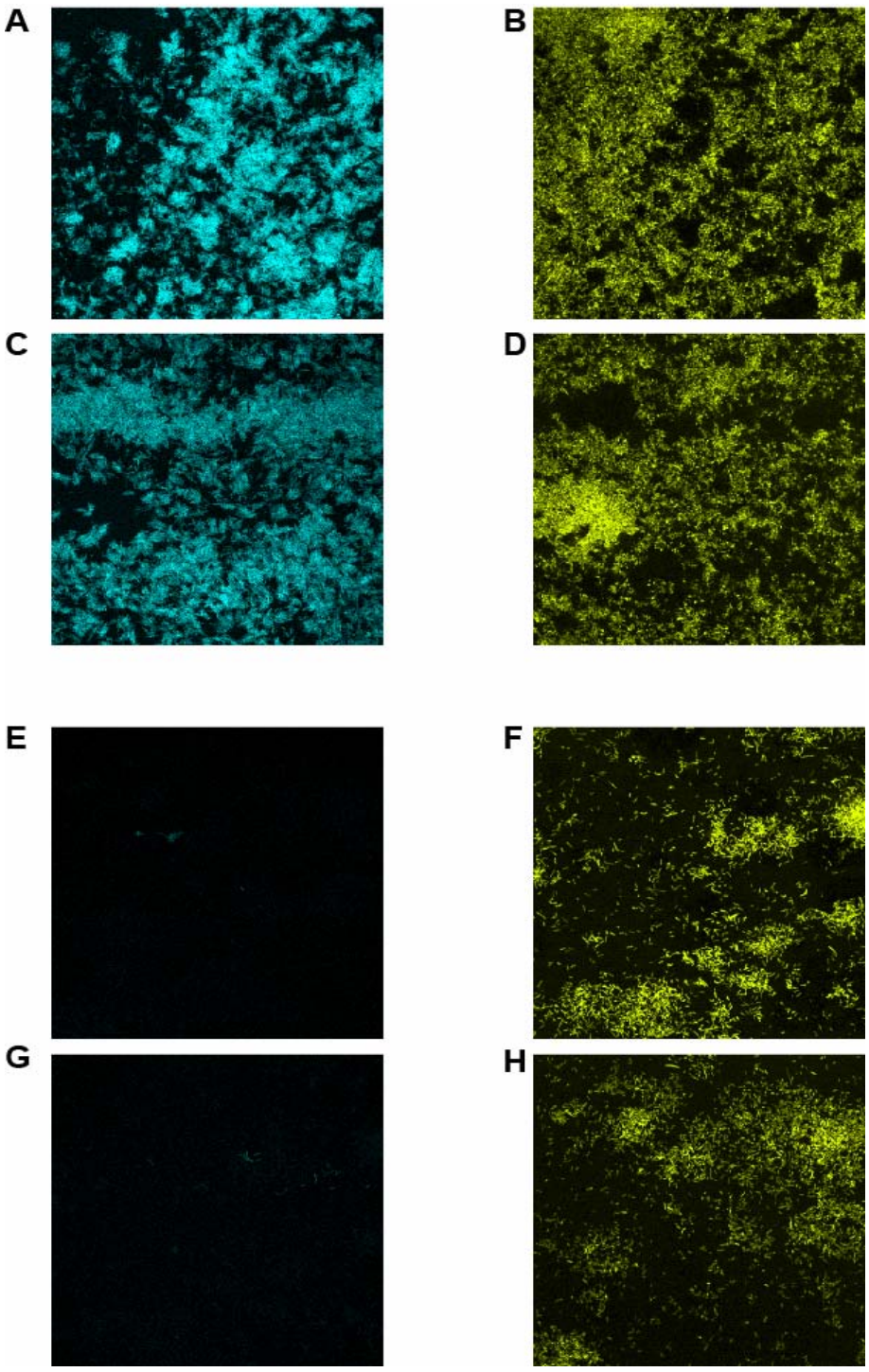

Figure B4 Note that this is a different layout of images from Figure B3. All images are taken after 72 hours of growth. (A)–(D) are untreated biofilms, (E)–(H) are treated biofilms. (A) and (B), (C) and (D) are paired blue and yellow images at the substrate from untreated second-generation (G2) biofilms. These show that blue and yellow are healthy, co-localized, and significant biomass has accumulated. (E) and (F), (G) and (H) are paired blue and yellow images at the substrate from treated G2 biofilms. These show that blue biomass is no longer healthy or has died, while yellow biomass is forming a scant biofilm. Overall, we can see that there is a significant difference in substrate colonization between the treated and untreated second-generation biofilms.

B.6 Emergent structure does not reappear in the third-generation control biofilm

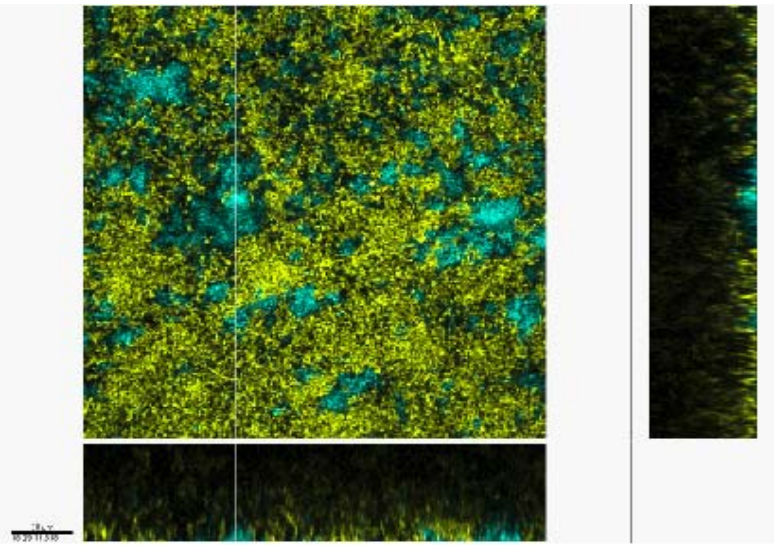

Figure B5 A projection taken at the substrate of the control third-generation biofilm, which was started from untreated effluent of the treated second-generation biofilm. It shows that emergent structure is no longer present in the third-generation biofilm. Although there are regions of co-localization of blue and yellow, the blue biomass and yellow biomass are both primarily localized near the substrate.

B.7 FACS was used to separate aggregates and YFP+ single cells

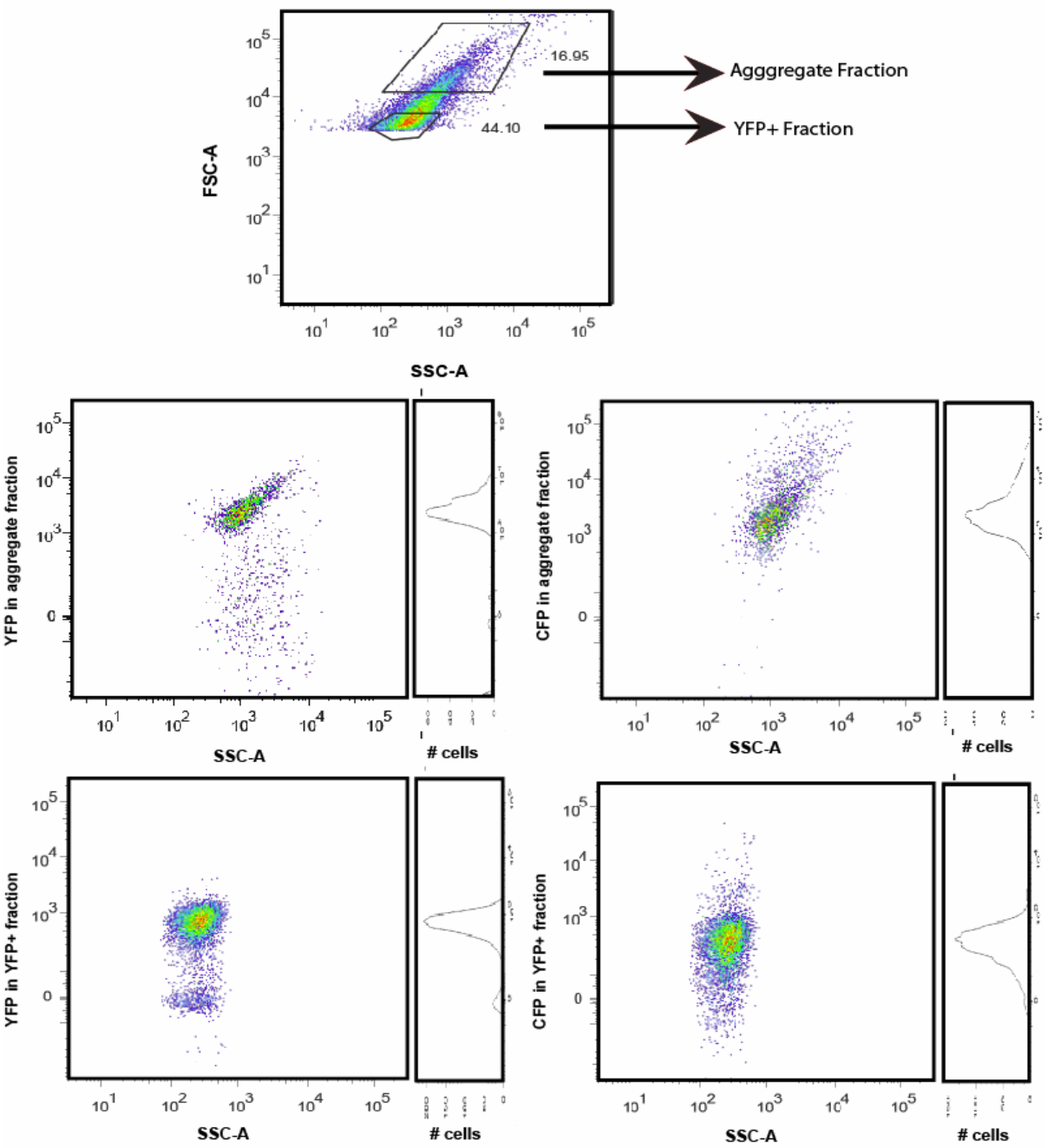

Figure B6 Representative data from FACS (this was repeated twice, this data is from one of the cases). The YFP+ fraction is conservatively gated to contain only single cells (based upon fluorescent and no-stain controls). The YFP+ fraction fluoresces in the yellow but not in the blue range, with a subpopulation that appears to lack fluorescence entirely (this could be single blue cells that are dead—they do not appear to adhere). The aggregate fraction contains groups of cells of unidentified shape and composition. It appears that most of these contain cells that fluoresce in the yellow range (although the smattering of dots below the primary group suggests the possibility of groups of just blue cells), and that most of the aggregates also exhibit fluorescence in the blue range (thus, most aggregates probably do contain at least some blue and some yellow biomass).