Amelie Ott, Timothy J. Martin, Kishor Acharya, Delina Y. Lyon, Nik Robinson, Bob Rowles, Jason R. Snape, Ian Still, Graham F. Whale, Vurtice C. Albright, Petra Bäverbäck, Nicola Best, Ruth Commander, Curtis Eickhoff, Sarah Finn, Björn Hidding, Heiko Maischak, Katherine A. Sowders, Masanori Taruki, Helen E. Walton, Aina Charlotte Wennberg, Russell J. Davenport. 2020. Multi-laboratory Validation of a New Marine Biodegradation Screening Test for Chemical Persistence Assessment. Environmental Science & Technology. 54 (7): 4210-4220.

Multi-laboratory validation of a new marine biodegradation screening test for chemical persistence assessment

Amelie Ott [a], Timothy J. Martin [a,1], Kishor Acharya [a], Delina Y. Lyon [b], Nik Robinson [c], Bob Rowles [d], Jason R. Snape [a,e,f], Ian Still [c,2], Graham F. Whale [g,3], Vurtice C. Albright III [h], Petra Bäverbäck [i], Nicola Best [j], Ruth Commander [k], Curtis Eickhoff [l], Sarah Finn [m], Björn Hidding [n], Heiko Maischak [o], Katherine A. Sowders [p], Masanori Taruki [q], Helen E. Walton [d], Aina Charlotte Wennberg [r], Russell J. Davenport [a]*

[a] Newcastle University, School of Engineering, Newcastle upon Tyne, NE1 7RU, United Kingdom

[b] Shell Oil Company, 150 N. Dairy Ashford Rd, Houston, TX 77079, United States

[c] European Oilfield Specialty Chemicals Association (EOSCA), Aberdeen, AB11 6YQ, United Kingdom

[d] Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, NR33 0HT, United Kingdom

[e] AstraZeneca Global Environment, Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TF, United Kingdom

[f] University of Warwick, School of Life Sciences, Gibbet Hill Campus, Coventry CV4 7AL, United Kingdom

[g] Shell International Ltd, Risk Science Team, 4 York road, London, SE1 7NA, United Kingdom

[h] The Dow Chemical Company, Toxicology & Environmental Research & Consulting, Midland, MI 48674, United States

[i] Schlumberger, Sandslikroken 140, 5254 Sandsli, Norway

[j] Covance CRS Research Limited, Shardlow Business Park, London Road, Derby, DE72 2GD, United Kingdom

[k] Scymaris Ltd., Brixham Laboratory, Brixham, TQ5 8BA, United Kingdom

[l] Nautilus Environmental Company Inc., Burnaby, BC, V5A 4N7, Canada

[m] National Oilwell Varco (NOV), Flotta, Stromness, Orkney, KW16 3NP, United Kingdom

[n] BASF SE, Carl-Bosch-Straße 38, 67056 Ludwigshafen am Rhein, Germany

[o] Noack Laboratorien GmbH, Käthe-Paulus-Straße 1, 31157 Sarstedt, Germany

[p] Baker Hughes – Environmental Services Group, 369 Marshall Ave., Webster Groves, Missouri 63119, United States

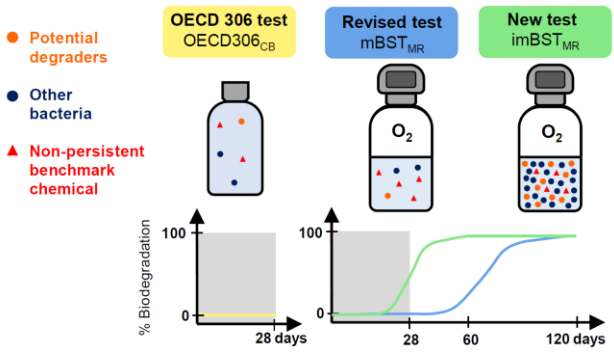
[q] Chemicals Evaluation and Research Institute, Japan, Kurume (CERI Kurume), 3-2-7 Miyanojin, Kurume-shi, Fukuoka 839-0801, Japan

[r] Norwegian Institute for Water Research (NIVA), Gaustadalléen 21, 0349 Oslo, Norway

ABSTRACT

Current biodegradation screening tests are not specifically designed for persistence assessment of chemicals, often show high inter- and intra-test variability, and often give false negative biodegradation results. Based on previous studies and recommendations, an international ring test involving 13 laboratories validated a new test method for marine biodegradation with a focus on improving the reliability of screening to determine the environmental degradation potential of chemicals. The new method incorporated increased bacterial cell concentrations to better represent the microbial diversity a chemical is likely to be exposed to in the sampled environments and ran beyond 60 days, which is the half-life threshold for chemical persistence in the marine environment. The new test provided a more reliable and less variable characterization of the biodegradation behavior of five reference chemicals (sodium benzoate, triethanolamine, 4-nitrophenol, anionic polyacrylamide, pentachlorophenol), with respect to REACH and OSPAR persistence thresholds, than the current OECD 306 test. The proposed new method provides a cost effective screening test for non-persistence that could streamline chemical regulation and reduce the cost and animal welfare implications of further higher tier testing.

GRAPHICAL ABSTRACT



INTRODUCTION

Regulatory frameworks (REACH [1], OSPAR [2]) combined with standardised test guidelines (OECD [3], ISO [4]) help to protect the environment and human health from the risks and hazards posed by globally manufactured chemicals. Within chemical risk assessment there has been a philosophical shift towards prioritising chemicals based on the hazard of potential environmental persistence, but regulatory tests have not reflected this change.[5] Biodegradation screening tests (BSTs) have not changed for over 30 years and are not effective at prioritising potential environmental persistence; they are laboratory based short-term test whose duration is much less than international half-life thresholds for persistence (60 days for seawater [6]), they are variable [1,7–9] and frequently report false negative outcomes.[10,11] These outcomes can result in additional costly biodegradation tests and potentially unnecessary bioaccumulation and toxicity tests of non-persistent chemicals. It is estimated that effective persistence assessments may save upwards of 600 fish and $75K per chemical reliably screened out earlier in the risk assessment process.[12]

BSTs' reliability can be increased by improving the representation of the environmental microbial community in the test vessel through increasing microbial numbers and diversity in the BSTs to more environmentally-relevant levels. It is hypothesised that this increases the likelihood of including competent degraders in the test vessel; in comparison to previous tests that were described as a "biodegradation lottery"; where small sample sizes can lead to variable test outcomes.[7,13–16] Intra-laboratory studies validated this concept for activated sludge and seawater BSTs in a modified OECD 301B setup.[7,17,18] Here, BSTs with more environmentally relevant cell numbers improved the reliability and accuracy of identifying the relative biodegradation classification of five radiolabeled benchmark chemicals. In addition, extended test durations beyond 28 days resulted in a more reliable identification of non-persistent chemicals.

Following the findings from this research and a workshop on improvements to the marine BST OECD 306 [16], an international ring test was performed to gather further scientific evidence towards validating the impact of increased bacterial cell concentrations and prolonged test durations of a new marine BST for chemical persistence assessment. The findings and recommendations of this multi-laboratory study are presented here.

MATERIAL AND METHODS

Ring Test Organization.

The new marine BST was validated following the OECD guidance document 34 which highlights fundamental aspects to consider when designing new test methods for regulatory acceptance. [19] The pre-validation [7,17,18] and inter-laboratory ring test validation conformed to all key factors recommended in chapter IV except that chemicals were not coded and sent blind to the contract research organizations (CROs). [19] Such coding was not used since only a restricted set of reference chemicals were sent to each CRO, and to enable correct handling and use of the chemicals in a way that conformed to health and safety policies.

For the ring test, the biodegradability of a group of reference chemicals was compared in three different test setups (see below) at 13 CROs in Canada, Germany, Italy, Japan, Norway, UK and USA (Supporting Information, Figure S1). The CROs (12/13 GLP accredited) conducted the tests at their own expense under GLP, or GLP-like, conditions to a ring test design and protocols developed by Newcastle University (NU) in collaboration with industry and regulatory bodies. The ring test setups were as described below:

a. **OECD306$_{CB}$:** Standard OECD 306 Closed Bottle Method with non-concentrated, aged seawater over 28 days [20] as a benchmark, against which to compare the revised and new test, plus one single measurement at day 60 to assess biodegradation potential and the previously reported oxygen limitation in this test beyond 28 days [20]

b. **mBST$_{MR}$:** Revised marine biodegradation screening test measuring biodegradation with manometric respirometers (MRs) for 120 days with non-concentrated seawater to validate use of MRs for marine BSTs and biodegradation potential beyond 60 days

c. **imBST$_{MR}$:** A new ("improved") marine biodegradation screening test measuring biodegradation with MRs for 120 days with 100-fold nominally increased bacterial concentrations in seawater to validate the effect of cell numbers and biodegradation potential beyond 60 days [18]

Sampling and Seawater Preparation.

Seawater collection for all three tests followed the OECD 306 Closed Bottle Method protocol [20] with subsequent pretreatments varying according to the test. The OECD306$_{CB}$ followed the original protocol that allows filtration or sedimentation and ageing of the seawater to remove coarse particles and reduce the content of dissolved organic material, respectively (Supplementary Information, Table S1).[20] For both MR methods, raw seawater was pre-filtered through a 10-µm polypropylene filter bag (Cole-Parmer, Vernon Hills, USA), but not aged. Marine bacterial cell numbers were increased 100-fold nominally by tangential flow filtration (TFF) for the new test (imBST$_{MR}$) only. CROs were asked to measure pH, temperature (T), dissolved oxygen (DO), conductivity, salinity and heterotrophic plate counts (HPC) in raw seawater (sample S1), post 10 µm filtration (sample S2), post TFF bacteria concentration (sample S3) and post ageing (sample S4) (Supplementary Information, Table S2). For CROs conducting the MR methods, NU took samples to additionally measure total cell counts (TCC) in samples S1, S2 and S3. Additional analysis included DNA sequencing for microbial community profiling from seawater samples collected prior test setup (samples S1, S2 and S3) and post 120 day incubation, but this data is not included here.

Tangential Flow Filtration.

Bacterial cell concentration was performed with a Pellicon 2 Mini TFF system (Merck, Darmstadt, Germany), operated with five 0.1 m^2 surface 0.22 µm pore-size polyvinylidene fluoride filters (Merck, Darmstadt, Germany), 3/8 in Tygon tubing (Merck, Darmstadt, Germany) and two peristaltic pumps (Watson Marlow, Falnmouth, UK) (Figure 1, Supplementary Information, Figure S2).[17]

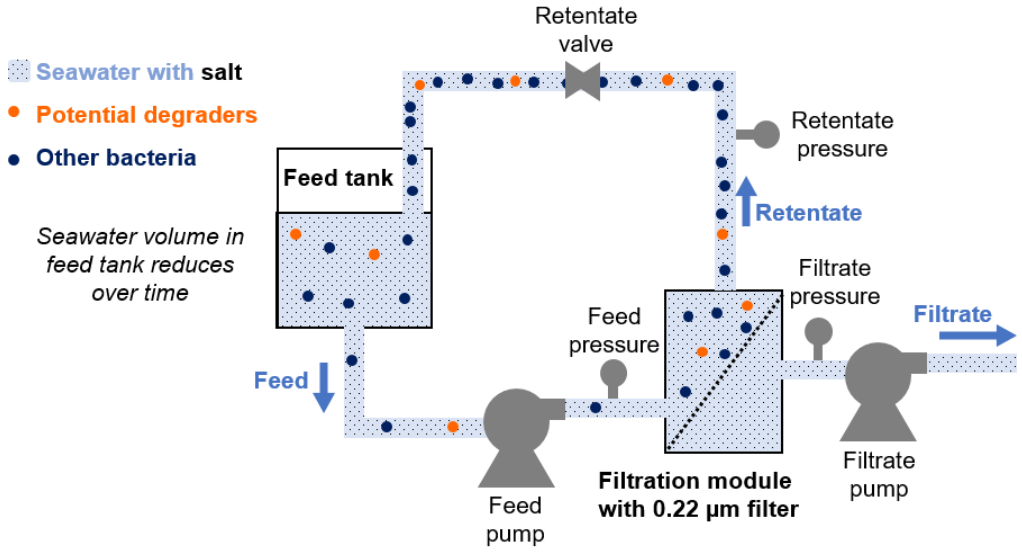


Figure 1. Schematic tangential flow filtration setup to increase bacterial cell numbers in seawater (based on [21]).

In TFF, water is pumped tangentially across the filter surface to reduce the chance of filter cake formation. Seawater including salts passes the 0.22 µm filter membrane as a partial flow and is removed as filtrate while bacteria remain in the retentate and are enriched in the feed tank. Using relatively "open" membranes with a pore size of 0.22 µm , a filtrate pump reduces filtrate flow and ensures a robust TFF process with reduced membrane wall concentrations and membrane fouling.[21]

NU provided the CROs with the TFF equipment and a NU representative performed the concentration and provided knowledge transfer of technical expertise to the host CRO. After the

first two test setups (CRO C and L), the permeate flow was reduced (2.6 L min^{-1} m^{-2} to 2.2 L min^{-1} m^{-2}) at same feed flow (6 L min^{-1} m^{-2}) to operate the TFF more stably across laboratories under varying seawater characteristics. Additionally, two recirculation steps with each 1 L of collected filtrate were included to flush any microorganisms sticking to the membrane in the retentate. The filtrate was flushed through the system at maximum feed pump speed (6.7 L min^{-1} m^{-2} feed flow) and clamped filtrate tubing (0 L min^{-1} m^{-2} permeate flow) for a cycle of 2 min run, 1 min break and 2 min run.

The same TFF filters were used throughout the ring test. Prior to concentration, filters were sanitized with 300 ppm sodium hypochlorite (NaOCl, pH 9, pH adjusted with 1 M hypochloric acid) up to 30 min and permeability tested with the normalized water permeability (NWP) test to assure filter cleanness and integrity. Following the manufacturers manual [22], NWP was calculated recording feed, retentate and permeate pressure under a set flow rate with high quality water. The initial NWP of the new membrane was used as the basis to determine membrane recovery, i.e. how effectively the membrane was cleaned back to its original state. After concentration, filters were cleaned with 300 ppm NaOCl, pH 9 for up to one hour and filter integrity was reassessed with the NWP test, before storing the filters in a bacteriostatic solution of 0.1M H_3PO_4, pH 2 at 4°C until next usage.

Flow Cytometry for TCC.

TCC were measured by fluorescence staining of nucleic acids combined with quantitative flow cytometry (FC) [23,24], using a FACScan flow cytometer (Becton Dickinson, Franklin Lakes, USA) with a 15 mW 488 nm air-cooled argon-ion laser. Seawater samples were collected and fixed in absolute ethanol (1:1 v/v) at the CROs, transported at 4°C to NU within 3 days and then stored at -20 °C until use. Microbial cells in 1 mL of sample were stained with 10 µL mL^{-1} SYBR Green I

working solution (10,000 x concentrated SYBR Green I in DMSO, Sigma Aldrich, St. Louis, USA, diluted 100 times in 10 mM Tris-HCL 1 mM EDTA, pH 8, Sigma Aldrich, St. Louis, USA) and incubated in the dark at 38 °C for 13 min before measurement.[25] Where necessary, seawater samples were diluted with filtered TE-buffer (0.22 µm; polyethersulfone membrane, Merck, Darmstadt, Germany) before staining to achieve an event (defined as a single particle detected by the instrument) rate between 200 and 800 bacteria/ s to avoid coincidence (i.e., two or more bacteria being at the same time within the sensing zone).[26] Readings were collected in logarithmic mode and analysed with Flowing Software 2.0, using electric gating to separate signals from background.[23,27]

Test Chemicals.

The following five test chemicals were selected to evaluate the limits of the tests (Supplementary Information, Table S3 and S4): a positive (sodium benzoate: SB) and negative (pentachlorophenol: PCP) reference chemical and three chemicals previously having shown variable degradation (triethanolamine: TEA, 4-nitrophenol: 4NP and anionic polyacrylamide: APAM). Based on the ECHA database and further literature (Supplementary Information, Table S4), chemicals were assigned following reference persistence and biodegradation categories: non-persistent and rapidly biodegradable (SB, TEA), non-persistent and inherently biodegradable (4-NP); or potentially persistent (PCP). APAM was chosen as a representative chemical used in the marine environment. As polymers are currently exempt from REACH regulation [28], its biodegradability behaviour is not classified in the ECHA database. Due to a lack of published reference biodegradation data for APAM, it was not possible to assign an expected biodegradation classification for this test chemical. Consequently, APAM results were reported separately to summaries of the SB, TEA, 4NP and PCP data.

Test System Setup.

The test setups were based on the capacity and ability of each CRO to perform either the $OECD306_{CB}$ or the MR methods, or both. In general, each CRO tested the positive and negative reference chemical; for the three variable chemicals, 4NP and TEA were tested more often than APAM, due to a greater volume of existing data for 4NP and TEA (Supporting Information, Table S4-S5). CRO L also conducted a toxicity control for PCP as part of their $imBST_{MR}$ setup (PCP + SB).

$OECD306_{CB}$ was prepared according to the original protocol, which uses natural seawater as the sole source of microorganisms.[20] Briefly, sacrificial 300 mL biological oxygen demand (BOD) bottles were filled with no headspace in triplicate for the oxygen blank and reference chemicals (test concentration 2 mg L^{-1}) to measure biodegradation on day 0, 7, 14, 21, 28 and optional day 60 after incubation at 20°C in the dark.

An MR method, similar to OECD 301F [29], was selected for the revised ($mBST_{MR}$) and new test ($imBST_{MR}$), using natural seawater for which increased microbial cells are used in the later. The headspace in MRs provides more O_2, which is required for prolonged test durations and thus renders ageing of seawater unnecessary to reduce background dissolved carbon content. At least 34-times more O_2 was available in MRs than in the $OECD306_{CB}$ (Supplementary Information, Table S6). Other advantages of MRs are that they require less seawater than sacrificial bottles, continuous biodegradation curves can be monitored and that they are already accepted by regulators.[29] However, it must be noted that MRs have a lower sensitivity compared to DO analysis and require higher chemical test concentrations.[29] In the ring test, CROs used OxiTop Control/ IS (WTW, Weilheim, Germany), CES (Coordinated Environmental Services, Kent, UK) and Micro-Oxymax (Columbus Instruments, Columbus, USA) respirometers (Supplementary Information,

Table S1). Media were prepared following OECD 301F guidelines[29], with the only difference that mineral and chemical stock solutions were diluted with filtered seawater ($mBST_{MR}$) or filtered, 100-fold concentrated seawater ($imBST_{MR}$) instead of water. The mineral medium was aerated with clean compressed air for 20-60 min at 20 °C.[20] Triplicate MR units were filled with 250 mL for the oxygen blank and reference chemicals (test concentration 75 mg $ThOD_{NH3}$ L^{-1}) to measure biodegradation continuously under stirred conditions over 120 days at 20°C in the dark. At the first CRO (C), PCP was tested at its water solubility limit of 14 mg L^{-1} (7.6 mg $ThOD_{NH3}$ L^{-1}). However, at subsequent setups, PCP was also added at 75 mg $ThOD_{NH3}$ L^{-1} to overcome MR detection limits. OxiTop systems were backed up and reset at day 60 to allow data collection past the system's memory capacity limit. Incubator temperatures were measured throughout the study and media temperatures, dissolved oxygen and pH were also recorded in all MR units after test termination.

Biodegradation Determination and Interpretation.

In all three tests, biodegradation of a chemical was measured indirectly as a function of O_2 consumption. While the $OECD306_{CB}$ monitors DO in the liquid phase [20], MRs measure O_2 consumption either from the change in volume or pressure in the apparatus (OxiTop), or by monitoring the quantity of O_2 produced electrolytically required to maintain constant gas volume in the flask (CES), or by measuring the O_2 and CO_2 concentrations in the headspace via closed-loop method (Micro-Oxymax). A solution of potassium hydroxide or another suitable absorbent adsorbed the evolved CO_2 in the OxiTop and CES system.[29,30] For all tests, biodegradation calculation was based on theoretical oxygen demand (ThOD).[20] Briefly, net O_2 consumption was calculated by subtracting the blank respiration from the O_2 depletion recorded in the test chemical bottles. Percentage biodegradation was then determined by accounting for chemical test

concentration and $ThOD_{NH3/NO3}$ ($ThOD_{NO3}$ for nitrogen containing TEA, 4NP, APAM). The Micro-Oxymax MR measures O_2 consumption as well as CO_2 production. Consequently, biodegradation was also calculated based on measured CO_2 with mineralization yield and $ThCO_2$.

For the $OECD306_{CB}$, 7/9 CROs included a single measurement at day 60 to assess biodegradation potential and previously reported O_2 limitation occurring in this test beyond 28 days.[20] Depending on the test setup and weekends, 4/7 CROs measured DO directly on day 60, with the other three CROs conducting the measurement on day 62, 59 and 63. For the purpose of comparing the tests with each other, all measurements were treated as if they took place on day 60. Following OECD 306 paragraph 4 and 15, blank BOD values on day 60 needed to be under 30% of that of the reference substances for the degradation measurements to be included in the analysis.[20] For the MR methods, blank respiration was evaluated against the OECD 301F threshold defined in paragraph 22 of 60 mg L^{-1} in 28 days.[29]

Biodegradation outcome was assessed both on the marine REACH and OSPAR threshold for persistence assessments. In REACH's integrated assessment and testing strategy (ITS), chemicals are classified as non-persistent if they show ≥ 60% biodegradation measured as ThOD over 60 days in an enhanced biodegradation screening tests.[31] Biodegradation under 60% ThOD in 60 days indicates potential persistence.[31] OSPAR (§2.2, 57) considers a substance to be persistent if "biodegradation is < 20% in OECD 306, Marine BODIS or any other accepted marine protocols or < 20% in 28 days freshwater (ready test)".[2] Continuous biodegradation recording in MR systems allowed the calculation of additional descriptors to assess the impact of increased bacterial cell concentrations on degradation. For each test chemical in the $mBST_{MR}$ and $imBST_{MR}$, time to reach 10% degradation i.e. lag time (t_L), time to reach 50% degradation (t_{50}, this descriptor is different to the t_{50} descriptor mentioned in the OECD 306 that excludes the lag phase – see below) and dt_{50}

256 (t_{50}-t_L, this descriptor is equivalent to t_{50} as mentioned in OECD 306) were determined. The values
257 t_L, t_{50} and dt_{50} were only based on those replicates that showed degradation and excluded those
258 that did not degrade. The exclusion of such zero values therefore influences the observed variance,
259 median and mean values.

260 For biodegradation results to be valid, at least two out of three replicates needed to show
261 degradation. Biodegradation values over 120% were classified as outliers and excluded from the
262 analysis. Negative biodegradation values were set to zero to calculate the coefficient of variation
263 (CV) based on mean degradation and standard deviation from the triplicate test setups. Data
264 analyses and visualisation was performed using R.[32]

265 The new test ($imBST_{MR}$) is based on the intra-laboratory validated marine environmentally
266 relevant BST (erBST).[18] For a detailed description of test protocol modifications from the erBST
267 to the $imBST_{MR}$, see Supplementary Information, Methods M1.

268

RESULTS AND DISCUSSIONS

Seawater Pretreatment.

Following the OECD 306 guideline [20], seawater was collected from 0.5 - 60 m below the surface and 40 - 5000 m offshore from March-August 2017 (Supplementary Information, Table S7) depending on the CROs normal practices and sampling locations. For the $OECD306_{CB}$, CROs followed their standard operating procedure (SOP) to pretreat seawater, providing an interesting insight in test variation within the OECD 306. After 8/9 CROs removed coarse particles by filtration or sedimentation, all CROs aged the seawater in the dark for 6-10 days with varying aeration conditions at 18-21 °C (Supplementary Information, Table S1). With ageing not being required, the MR tests were set up sooner after seawater collection (mean 2.8 ± 1.4 days, range 1 – 5 days) than the $OECD306_{CB}$ tests (mean 8.2 ± 2.3 days, range 6 – 13 days).

7/9 CROs determined HPC for the $OECD306_{CB}$ [20] (Supplementary Information, Table S2), with NU measuring TCC at all MR test setups. Culture-dependent HPC only measures a small fraction of TCC (0.01-1% [33]), but a moderate positive correlation between both HPC and flow cytometry methods (Supplementary Information Figure S3) allows comparison of the impact of ageing and TFF on cell numbers in the $OECD306_{CB}$ and $imBST_{MR}$, respectively. It should be noted that different CROs used different media and methods for HPC culturing (Supplementary Information, Table S2), which can affect the number and types of microorganisms recovered.[34] Therefore, greater value can be placed on concentration changes within one CRO, rather than comparisons across laboratories. The variation of bacterial concentrations in raw seawater collected from different sites varied by an order of magnitude for both enumeration methods (HPC and TCC), even if HPC on average only accounted for 7% of the TCC (Figure 2).

Ageing with preceding filtration/ sedimentation had a variable impact on bacterial numbers (Figure 2). Depending on the CRO, OECD 306 pretreatment increased (up to 142-fold) or reduced (80% lower) cell concentrations from raw seawater (based on HPC from raw (sample S1) to filtered/ sedimented and aged seawater (sample S4), Supplementary Information, Table S8). This variable change in cell numbers is not solely explained by the variation in pretreatment methods, but probably also depends on the initial microbial composition of the seawater. Ageing has previously shown to impose a selective pressure on the microbial community and change its composition from the sampled environment.[35] When the test chemical is then added, the bacterial community may have become atypical of the environment. This may lead to a higher or lower biodegradation potential to be observed and consequently increases the uncertainty and inaccuracy of extrapolating laboratory biodegradation data to the environment.[35]

Based on TCC, raw and 10 µm filtered seawater across CROs conducting the new ($imBST_{MR}$) and revised ($mBST_{MR}$) tests contained on average 10^5 bacterial cells mL^{-1} (ranging from 10^4 to 10^5 bacterial cells mL^{-1}) with increased average bacteria concentrations after TFF processing of 10^7 bacterial cells mL^{-1} (ranging from 10^6 to 10^8 bacterial cells mL^{-1}). TFF increased the concentration of bacteria at all CROs on average 107-fold, ranging from a 14-fold to 222-fold increase (based on TCC from 10 µm filtered (sample S2) to concentrated (sample S3) seawater, Supplementary Information Table S8). Due to time and logistical constraints, TFF was optimized at NU and those conditions applied at each CRO. The process could be improved towards achieving the intended 100-fold increase at all locations through optimizing the flow rates for each seawater source. As expected, TFF did not increase salinity (Supplementary Information Table S7). Martin *et al.* (2018) previously showed that TFF does not significantly change the relative microbial community composition, with concentrated marine bacteria communities being a good

representation of the sampled environments.[17] A chemical in the sea encounters a vast amount of microbes in a short amount of time with cell concentrations in the range of 10^{10}-10^{11} TCC m^{-3} (as determined here), seawater turnover times in the order of 10^{5}-10^{6} m^{3} s^{-1} [36] and typical velocities in coastal oceans of 0.1-1 m s^{-1}.[37] For the $imBST_{MR}$, the test chemical is therefore introduced to a more environmentally relevant wider microbial community by increasing the bacterial numbers used in the test.

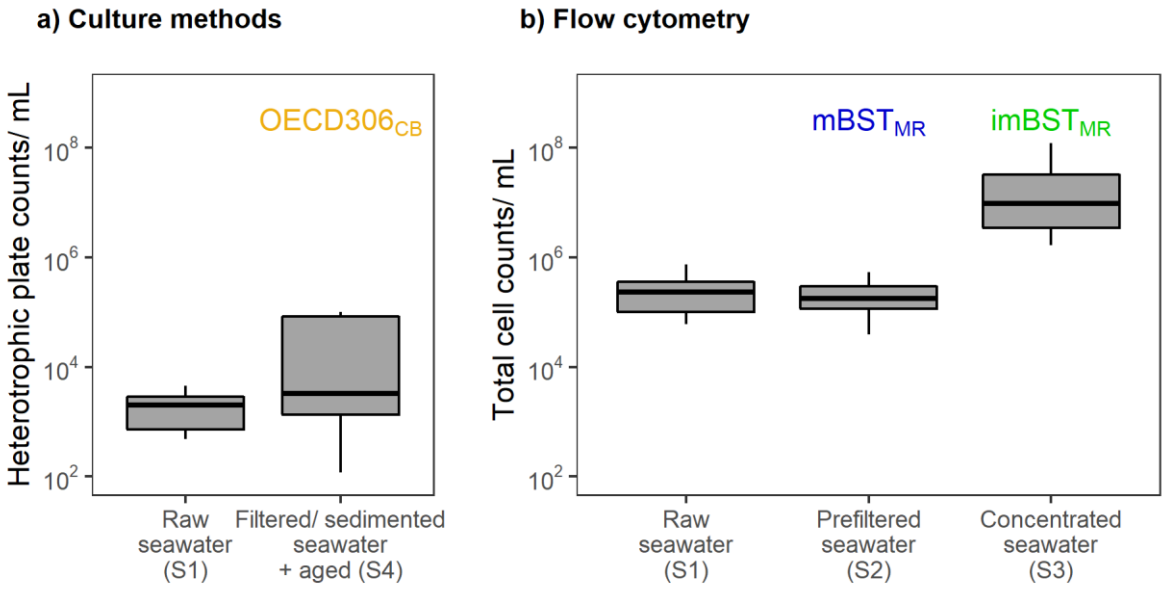


Figure 2. Boxplot showing the effect of pretreatment on bacterial concentrations for all three test setups. (a) Heterotrophic plate counts (determined by different culture methods) for $OECD306_{CB}$ setups. (a) Total cell concentrations (determined by flow cytometry) for $mBST_{MR}$ and $imBST_{MR}$ setups.

Chemical Classification.

The new test ($imBST_{MR}$) was more accurate and less variable than the comparator-screening tests, the $mBST_{MR}$ or $OECD306_{CB}$ (Table 1, Supplementary Information, Figures S4-13, Tables

328 S9-16). According to the REACH biodegradability criterion in marine water, the $imBST_{MR}$
329 correctly classified 70% of the reference chemicals to their respective persistence category (non-
330 persistent or potentially persistent) and had a coefficient of variation of 30% between tests. In
331 contrast, the $OECD306_{CB}$ only correctly classified 48% of the test chemicals and had a coefficient
332 of variation of 48%. Thus, the new test method has a much lower rate of false negatives according
333 to the REACH criterion compared to the current test method; 41% and 62%, respectively (Table
334 1). Within the non-persistent chemicals, SB degraded in almost all replicates with more variable
335 biodegradation results for TEA and 4NP (Supplementary Information, Table S15). While the new
336 test increased the correct classification of SB, TEA and 4NP as non-persistent (55% ± 43% based
337 on replicates) in comparison to the revised (36% ± 55%) and current OECD 306 test (37% ± 55%),
338 it shows that some non-persistent chemicals are still going to fail this new test. For instance, 4NP
339 degraded in 11% of the replicates in the new test according to the REACH criterion, but in no
340 replicates in the revised or OECD 306 test (Supplementary Information, Table S15). While 4NP
341 has been observed to fully degrade in activated sludge BSTs [7], its biodegradation in marine BSTs
342 has been found to be more variable.[9,18] This appears to be related to previous exposure to 4NP
343 where rapid biodegradation is observed with pre-adapted inocula.[9]
344 The variability in biodegradation results differed across CROs, test chemicals and test setups with
345 the lowest coefficient of variation value for SB in the OECD 306 test (5%) and the highest
346 coefficient of variation value for 4NP, also in the OECD 306 test (75%) (Supplementary
347 Information, Table S16).

348 Some erratic degradation behavior was observed in all three test setups for the negative control
349 (Supplementary Information, Figures S12-13). For the $mBST_{MR}$ and $imBST_{MR}$, these anomalous
350 replicates may relate to solubility and toxicity issues associated with PCP at the test concentrations

351 employed to overcome MR detection limits. The toxicity control performed at one CRO showed
352 an inhibitory effect of PCP at the concentration employed in the MR tests (139 mg L^{-1})
353 (Supplementary Information, Figure S14). PCP was the best possible choice out of over 30
354 chemicals investigated to find a negative reference chemical [38,39], but a measure of caution should
355 be taken when interpreting these biodegradation results. Neither, the new ($imBST_{MR}$) or revised
356 ($mBST_{MR}$) test showed any false positives under the criteria chosen for evaluation. This is
357 consistent with the intra-laboratory validation, where radiolabeled PCP was employed at test
358 concentrations below the solubility and toxicity threshold at 10 mg L^{-1}.[18]

359 False positives (33%) were only reported for the $OECD306_{CB}$ method across CROs when using
360 the OSPAR persistence criterion, though based on some unusual biodegradation curves, since
361 often the value was a spike in all replicates at a single time point (Table 1, Supplementary
362 Information, Figure S13). It is unclear if this was due to the low test concentration of PCP applied,
363 the general increased variability of the $OECD306_{CB}$, and/or that the OSPAR criterion for
364 persistence is different than that used by REACH.

365 For all three tests, the REACH non-persistence criterion, with its higher biodegradation
366 threshold of 60% over 60 days, appeared to characterize the reference chemicals more accurately
367 and reliably than the OSPAR persistence criterion of <20% over 28 days (Table 1). Assessing the
368 biodegradation data based on the REACH threshold resulted not only in no false positives, but also
369 reduced false negative rates across all three tests in comparison to the OSPAR criterion (Table 1).
370 It is also worthwhile noting that within REACH a "result of >20% ThOD or DOC removal is
371 indicative of a potential for primary biodegradation in the marine environment".[1]

Table 1. Correct persistence assessment, false negatives and false positives in the three test setups across CROs as evaluated against two current regulatory thresholds for the reference chemicals according to their expected classification (Supplementary Information, Table S4). Test variation across three tests is described by the coefficient of variation (CV) including and excluding the negative control (for CVs per chemical, see Supplementary Information, Table S16).

	According to:	**Current test** $OECD306_{CB}$	**Revised test** $mBST_{MR}$	**New test** $imBST_{MR}$
Correct persistence assessment: SB, TEA, 4NP are non-persistent and PCP is potentially persistent	OSPAR [a]	42%	55%	63%
	REACH [b]	48% [c]	59%	70%
False negatives: incorrect assessment of SB, TEA, 4NP as potentially persistent	OSPAR [a]	63%	62%	50%
	REACH [b]	62% [c]	57%	41%
False positives: incorrect assessment of PCP as non-persistent	OSPAR [a]	33%	0%	0%
	REACH [b]	0% [c]	0%	0%
Coefficient of variation including negative control		49%	42%	35%
Coefficient of variation excluding negative control		48%	47%	30%

[a] OSPAR: Biodegradation ≥ 20% over 28 days = non-persistent; biodegradation < 20% over 28 days = persistent [2]
[b] REACH: Biodegradation ≥ 60% over 60 days = non-persistent; biodegradation < 60% over 60 days = potentially persistent [31]
[c] Test extended to 60 days in accordance with OECD 306 Closed Bottle Method § 4 and 15 [20]

The synthetic polymer APAM was tested as polyacrylamides (PAMs) are highly relevant to the marine environment. PAMs are widely used in several industrial fields such as for water treatment, agriculture and oil recovery.[40] As its biodegradability behavior is not classified in the ECHA database, and peer-reviewed scientific reference data is lacking, reference values for the comparison in Table 1 are not available and its degradation results are mentioned separately. For the revised and new test, APAM did not show any degradation under the OSPAR and REACH criteria. However, APAM was classified as non-persistent in 25% of CROs in the $OECD306_{CB}$ according to the OSPAR persistence criterion, but not according to the REACH biodegradability

criterion (Supplementary Information, Figure S11, Table S9). These results should be evaluated carefully considering the false positive PCP characterizations under the same assessment conditions ($OECD306_{CB}$ and OSPAR criterion). Additionally, APAM previously showed no degradation in BSTs and studies found PAM macromolecules resistant to microbial attack, requiring an initial physical-chemical break-down.[41,42]

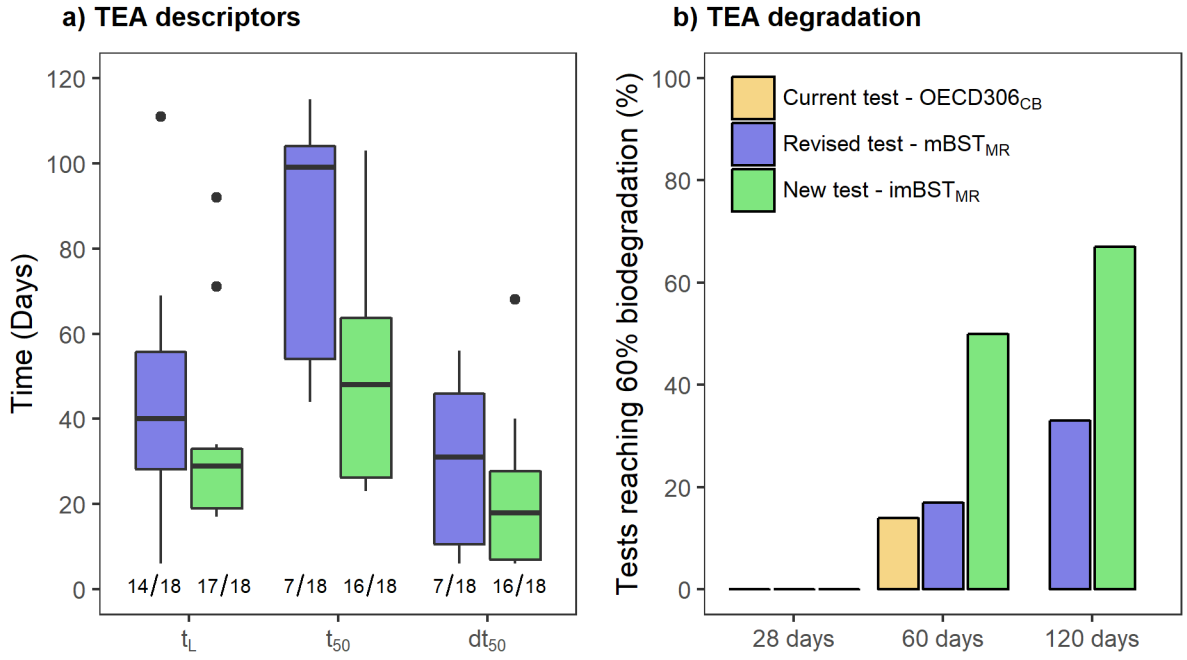


Figure 3. Example plots for triethanolamine (TEA). a) Increased cell numbers in the new test reduce t_L (time to 10% degradation), t_{50} (time to 50% degradation) and dt_{50} ($t_{50} - t_L$). Boxplot based on $mBST_{MR}$ and $imBST_{MR}$ replicates where descriptor values could be determined (indicated by values under each boxplot) within the 120 day test period. In Figure S15, t_L, t_{50} and dt_{50} were set to 121 days for non-degrading $mBST_{MR}$ and $imBST_{MR}$ replicates. **b)** Correct non-persistence assessment increases with longer test durations.

Extended test durations in the $mBST_{MR}$ and $imBST_{MR}$ allowed the analysis of lag phases, which extended beyond the standard 28-day test duration. These lag phases, particularly for TEA and

4NP, were often followed by fast and complete degradation of the test chemical (Figure 3a, Supplementary Information, Figure S15) indicating the presence of an acclimated viable degrading community. Acclimation is a common but poorly understood phenomenon that requires further investigation in the context of regulatory biodegradation testing.[43–45] In general, increased cell numbers in the new test resulted in shorter and more consistent lag periods (Supplementary Information, Tables S10-S13).

The probability of observing degradation increased with time so that for those chemicals that have previously shown variable biodegradation results, 4NP and TEA, the 120-day duration gave a more reliable characterization of the persistence category for a given chemical than the 60 and 28-day duration test (Figure 3b, Supplementary Information, Tables S11-S12). There was a positive relationship between TCC and biodegradation potential (Pearson correlation, P(4NP) and TCC 0.83, P(TEA) and TCC: 0.65 with $p<0.01$). However, the greatest rates and extents of degradation were not necessarily observed at the CROs with the highest cell concentrations, suggesting that cell concentration is not the only factor influencing the degradation potential of an environmental sample. Indeed further research is needed to investigate how microbial diversities at different sampling locations affect biodegradation test outcome. Microbial community analysis of seawater samples collected prior test setup and post 120 day incubation in the ring test will be subject of a separate publication.

Test system performance, anomalies, caveats and data quality checks.

For the $OECD306_{CB}$, seawater ageing allowed the test to run past 28 days without oxygen limitation occurring (blank BOD under 30% [20]; Supplementary Information, Figure S16). It should be noted that these results are not based on a time series but on only one sacrificial triplicate measurement past 28 days at day 60. Blank readings for the $mBST_{MR}$ and $imBST_{MR}$ were within

the 60 mg L^{-1} 28 days threshold defined in OECD 301F (Supplementary Information, Figure S17). [29] In closed system MRs, blank respiration remained under 60 mg L^{-1} over 120 days for all CROs except for CRO A in the $imBST_{MR}$ setup. Interestingly, higher blank oxygen consumptions were recorded for the oxygen replenishing MRs (CRO K and M) than in for the closed system MRs.

Incubator temperatures for all tests were within 20 ± 2°C.[20] However, at 7/9 CROs, temperature increases over 22 °C were detected in MR bottle contents after 120 days, probably caused by residual heat from the stirring motion in the MRs or from the stirring platforms on which they sit (Supplementary Information, Figure S18). The water bath operated CES respirometer showed the lowest temperature increase (mean 20.1 ± 0.2 °C). The use of water baths instead of incubators, reducing stirrer speed, or incubation temperatures may help to mitigate such variation.

Out of 528 sacrificial $OECD306_{CB}$ bottles, 18 bottles were excluded from the analysis at CRO F with biodegradation values >120% and systematic anomalous results in all batches on day 7 (Supplementary Information, Figure S19). Out of 205 started MR units ($mBST_{MR}$ 100, $imBST_{MR}$ 105), eleven units were not included in the analysis (Supplementary Information, Figure S19). At CRO I, two OxiTop units stopped working as batteries ran out of power within the first week. At the first CRO C, all PCP units were excluded as the chemical was added at concentrations under the detection limit. Three units in oxygen replenishing MR systems were excluded with biodegradation values over 120% (Supplementary Information, Figure S20). To reliably assess the $mBST_{MR}$ and $imBST_{MR}$ in the Micro-Oxymax at CRO M, the more robust CO_2 production data instead of O_2 consumption biodegradation data was included in the analysis (Supplementary Information, Figure S21). MRs proved suitable for monitoring biodegradation in seawater, but reliability varied depending on the system used. In general, biodegradation values over 120% in all three tests may have been caused by bottle contamination or calibration errors and negative

biodegradation values by test chemical inhibition or disproportionally high blank respiration (e.g. contamination with organic debris/ protozoa).

Practical aspects of tangential flow filtration.

Concentration of bacterial cells in the ring test was performed with the previously tested and optimized Pellicon 2 Mini TFF system (Merck, Darmstadt, Germany) [17], but other filtration systems could also be employed for the new test as long as they do not alter the microbial composition of the raw seawater. TFF costs vary depending on the required sample throughput and manufacturer. The compact TFF setup as employed for the ring test costs around $15K (including holder, tubing, fitting kit, pressure gauges, filters) with additional costs of approximately $6K for the two peristaltic pumps. The time to increase bacterial cell concentrations by a nominal 100-fold in seawater using TFF depends on following aspects: seawater volume to filter (defined by test setup e.g. number of test chemicals, replicates and test volume), filter surface and seawater characteristics (e.g. pollution status, particle content). For instance, performing the new test ($imBST_{MR}$) with triplicate blank, positive control and test chemical would require bacterial cells present in 300 L to be concentrated to 3 L. Filtering this water would take 5 h with the compact "travel-friendly" ring test TFF setup (filter surface 0.5 m^2, conservative permeate flow 2.2 L min^{-1} m^{-2}), but only 20 min with a bigger system at same permeate flow (e.g. Pellicon Cassette Acrylic Holder, filter surface 5 m^2). For less viscous (clearer) seawater, permeate flow can be increased to further reduce filtration time while maintaining conditions of minimal fouling and operating a steady process.

Regulatory Implications.

The purpose of regulatory BSTs is to screen out those chemicals that degrade rapidly from those that are potentially persistent. This ring test demonstrated that the new test ($imBST_{MR}$) provides a more robust prioritization on potential persistence than the current OECD 306, improving the reliability of BSTs by increasing bacterial cell numbers and extending test durations (Table 2), as suggested by previous studies.[7,13–15,18] This new test would provide more robust data and increase confidence in biodegradation conclusions.

The findings of the ring test together with other research [7,17,18] demonstrate that increasing bacterial concentrations is a suitable modification to improve persistence assessment for "enhanced screening tests", despite its recent exclusion as an accepted approach in the REACH endpoint specific guidance.[1,18,46] While the new test better represents the microbiome of the sampled environment by capturing 100-fold more bacteria in the test vessel [17], it is still a conservative screening test, being based on growth-linked biodegradation using unrealistically high test chemical concentrations to overcome analytical constraints. In the ring test, standard OECD 306 seawater pretreatment had a variable effect on bacterial concentrations, sometimes increasing them by two orders of magnitude. This increase is comparable to cell concentrations in the $imBST_{MR}$. However, in contrast to TFF [17], the incubation conditions during ageing have been documented to apply an unnatural selection pressure and alter the microbial community composition from that in the original seawater sample.[47] The ratio of bacterial cells to test chemical in the standard OECD 306 method were comparable to the new test given the one to two orders of magnitude higher test chemical concentrations employed in the latter and the variable bacterial cell concentration effects of ageing in the former. Previous studies have also shown that kinetics in BSTs with increased bacterial cell concentrations can be indistinguishable from those in current BSTs.[7] In general, it should be highlighted that bacteria to test chemical ratios can vary greatly in

existing OECD BSTs with cell concentrations varying by five orders of magnitude and chemical concentrations varying by two orders of magnitude.[20,29]

Table 2. Advantages and disadvantages of using the new test ($imBST_{MR}$) to screen for non-persistent chemicals in seawater (in comparison to the $OECD306_{CB}$).

Advantages	Disadvantages
• **Increased reliability:** New test is more reliable and less variable in screening for non-persistent or potentially persistent chemicals than OECD 306_{CB} method (based on tested reference chemicals); effective persistence assessment saves costs and reduces potentially unnecessary animal testing [12]; • **Regulatory acceptance:** MRs are already accepted by regulators to monitor biodegradation for the OECD 301F [29]; some CROs already have MRs available and are familiar with their use; • **Extended test durations:** Headspace (and oxygen replenishing mechanisms) in MRs reduce oxygen limitation, render seawater ageing unnecessary and allow to extend test durations beyond 28 days; • **Reduced maintenance:** Once MRs are setup, biodegradation measurements can be recorded continuously and automatically; • **Increased environmental relevance:** While both, ageing ($OECD306_{CB}$) and TFF (new test) increased cell concentrations up to two orders of magnitude in the ring test, TFF has been shown previously to not alter the microbial community significantly [17], in comparison to ageing;[47]	• **Higher test chemical concentrations:** MRs are less sensitive than dissolved oxygen measurements in $OECD306_{CB}$;[a] • **More seawater required:** CROs have to collect 100-fold more seawater for the cell concentration step (note however, that less seawater is required to run non-destructive MR units than sacrificial $OECD306_{CB}$ bottles);[b] • **Investment:** CROs need to invest in a filtration system (and potentially MR units) and familiarize themselves with the equipment;[b] • **Testing poorly soluble and/or volatile chemicals:** To expand on the scope of chemicals tested in the new test, modifications for poorly soluble chemicals as described in OECD 301 Annex III and by other methods [48,49] might be necessary; some MR systems with plastic components might not be suitable to test volatile hydrocarbons due to abiotic losses[50];

[a] Radiolabeling could allow testing at lower test chemical concentrations; [b] Seawater concentration could be performed at specialized facilities located near to the sea

To avoid the cost and animal welfare implications of additional potentially unnecessary testing, it is crucial to reduce the variability, and thus number of false negatives in current first tier BSTs.[7] Within the integrated testing strategy for persistence assessment, the new test ($imBST_{MR}$) could sit at a tier lower than the more complex, costly and time-consuming simulation tests (OECD 307, 308 and 309 [31]).

Better guidance is required on interpreting prolonged lag phases followed by quick degradation observed in the ring test and other marine studies.[18,51,52] It should be investigated whether these long lag phases are likely to occur during the degradation of chemicals in the sea or whether they are artefacts of the stringent but less environmentally relevant physico-chemical conditions in BSTs. In the absence of such comparisons, the new test offers a practical and economical means to improve the screening of chemicals likely to end up in the marine environment as part of the current persistence assessment testing strategy.

AUTHOR INFORMATION

Corresponding Author

*(A.O.) E-mail: a.i.g.ott2@ncl.ac.uk

Present Addresses

[1] Department for Environment, Food and Rural Affairs, Lancaster House, Newcastle upon Tyne, NE4 7JB, United Kingdom

[2] Nalco Champion, WSW Building, Peterseat Drive, Aberdeen, AB12 3HT, United Kingdom

[3] Whale Environmental Consultancy Ltd., 55 Earlsway Curzon Park, Chester, CH4 8AZ, United Kingdom

Author Contributions

R.J.D., A.O., T.J.M. conceived and designed the study with substantial contributions from D.Y.L, N.R., B.R., J.R.S, I.S., G.F.W., V.C.A., P.B., N.B., R.C., C.E., S.F., B.H., H.M., K.A.S., M.T., H.E.W., A.C.W.; A.O., K.A., V.C.A., P.B., N.B., R.C., C.E., S.F., B.H., H.M., K.A.S., M.T., H.E.W., A.I.W. collected and contributed data; A.O., T.J.M performed data analysis; A.O. and R.J.D. wrote the paper with critical feedback from all co-authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was primarily funded by the European Chemical Industry Council (Cefic) Long-range Initiative (LRi; project ECO11.2-3). Additional funding support was provided by the Engineering and Physical Sciences Research Council [EPSRC] (Grant number: EP/1025782/1, Challenging Engineering Award). Biodegradation data supporting this publication is openly available under http://doi.org/10.25405/data.ncl.9938198.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors would like to acknowledge the useful advice and discussions of the ring test organizing committee and CROs (Baker Hughes, BASF, Cefas, CERI, Covance, DOW, Lamberti, Nautilus Environmental Company, NIVA, Noack Laboratorien, NOV, Schlumberger, Scymaris). We thank the Norwegian Environmental Agency and Claire Phillips (Cefas) for their comments during the test protocol development and final reporting. We would also like to show our gratitude to Ed Schäffer (EAG Laboratories) and Claas Hiebenthal (Helmholtz Centre for Ocean Research Kiel) for facilitating the seawater concentration at their laboratories for some of the CROs. The authors thank Paul Beckett (Merck) for his advice on optimizing the TFF process and Equinor and SNF for the provision with the test chemical anionic polyacrylamide. We thank ECETOC for providing assistance in project management.

ABBREVIATIONS

4NP	4-nitrophenol
APAM	anionic polyacrylamide
BOD	biological oxygen demand
BST	biodegradation screening test
CRO	contract research organization
CV	coefficient of variation
DO	dissolved oxygen
dt_{50}	$= t_{50}-t_L$
FC	flow cytometry
HPC	heterotrophic plate counts
ITS	integrated assessment and testing strategy
$imBST_{MR}$:	new "improved" marine biodegradation screening test measuring biodegradation with manometric respirometers
$mBST_{MR}$:	marine biodegradation screening test measuring biodegradation with manometric respirometers
MR	manometric respirometer
NaOCl	sodium hypochlorite
NU	Newcastle University
NWP	normalized water permeability
$OECD306_{CB}$:	OECD 306 Closed Bottle Method
PAM	polyacrylamide
PCP	pentachlorophenol
SB	sodium benzoate
SOP	standard operating procedure
t_{50}	time to reach 50% degradation
T	temperature
TCC	total cell counts

TEA	triethanolamine
TFF	tangential flow filtration
ThOD	theoretical oxygen demand
t_L	lag phase; time to reach 10% degradation

Supporting Information.

This information is available free of charge via the Internet at http://pubs.acs.org. Test protocol modifications in the $imBST_{MR}$ from the pre-validated marine erBST. Location map of CROs participating in the ring test. Photos of example TFF setup. Graph of HPC and TCC correlation in seawater samples. Biodegradation plots and overview of chemical degradation data. Blank oxygen uptake plots. Boxplots of temperatures measured in MR batches. Summary tables for water analysis methods and measurements. Details on reference chemical properties, selection and testing strategy. Calculations of available oxygen in different test setups. Table of pretreatment effects on bacteria concentrations.

REFERENCES

(1) ECHA. *Guidance on Information Requirements and Chemical Safety Assessment, Chapter R . 7b : Endpoint Specific Guidance (Version 4.0)*; 2017. https://doi.org/10.2823/84188.

(2) OSPAR. *Guidelines for Completing the Harmonised Offshore Chemical Notification Format (HOCNF) (OSPAR Agreement: 2012/05. Update 2015)*; 2015.

(3) OECD. *OECD Guidelines for the Testing of Chemicals*; Paris, 2006.

(4) ISO. *ISO 10707:1994. Water Quality; Evaluation in an Aqueous Medium of the "Ultimate" Aerobic Biodegradability of Organic Compounds; Method by Analysis of Biochemical Oxygen Demand (Closed Bottle Test).*; 1994.

(5) Cousins, I. T.; Ng, C. A.; Wang, Z.; Scheringer, M. Why Is High Persistence Alone a Major

Cause of Concern? *Environ. Sci. Process. Impacts* **2019**, *21* (5), 781–792. https://doi.org/10.1039/c8em00515j.

(6) Regulation (EC) No 1907. *Of the European Parliament and of the Council of 18 December 2006 Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), Establishing a European Chemicals Agency, Amending Directive 1999/4*; 2006.

(7) Martin, T. J.; Snape, J. R.; Bartram, A.; Robson, A.; Acharya, K.; Davenport, R. J. Environmentally Relevant Inoculum Concentrations Improve the Reliability of Persistent Assessments in Biodegradation Screening Tests. *Environ. Sci. Technol.* **2017**, *51* (5), 3065–3073. https://doi.org/10.1021/acs.est.6b05717.

(8) Gerike, P.; Fischer, W. K. A Correlation Study of Biodegradability Determinations with Various Chemicals in Various Tests. *Ecotoxicol. Environ. Saf.* **1979**, *3* (1), 159–173. https://doi.org/10.1016/0147-6513(81)90045-2.

(9) Nyholm, N.; Lindgaard-Jørgensen, P.; Hansen, N. Biodegradation of 4-Nitrophenol in Standardized Aquatic Degradation Tests. *Ecotoxicol. Environ. Saf.* **1984**, *8* (5), 451–470. https://doi.org/10.1016/0147-6513(84)90066-6.

(10) ECETOC. *Workshop on Biodegradation and Persistence No. 10*; Brussels, 2007.

(11) Gartiser, S.; Schneider, K.; Schwarz, M. A.; Junker, T. *Assessment of Environmental Persistence: Regulatory Requirements and Practical Possibilities - Available Test Systems, Indentification of Technical Constraints and Indication of Possible Solutions.*; 2017.

(12) Martin, T. J.; Goodhead, A. K.; Acharya, K.; Head, I. M.; Snape, J. R.; Davenport, R. J. High Throughput Biodegradation-Screening Test to Prioritize and Evaluate Chemical Biodegradability. *Environ. Sci. Technol.* **2017**, *51* (12), 7236–7244. https://doi.org/10.1021/acs.est.7b00806.

(13) Thouand, G.; Friant, P.; Frédéric, B.; Alain, C.; Maul, A.; Block, J. C. Bacterial Inoculum Densitiy and Probability of Para-Nitrophenol Biodegradability Test Response. *Ecotoxicol. Environ. Saf.* **1995**, *30* (3), 274–282. https://doi.org/10.1006/eesa.1995.1031.

(14) Ingerslev, F.; Toräng, L.; Nyholm, N. Importance of the Test Volume on the Lag Phase in Biodegradation Studies. *Environ. Toxicol. Chem.* **2000**, *19* (10), 2443–2447. https://doi.org/10.1897/1551-5028(2000)019<2443:IOTTVO>2.3.CO;2.

(15) Goodhead, A. K.; Head, I. M.; Snape, J. R.; Davenport, R. J. Standard Inocula Preparations Reduce the Bacterial Diversity and Reliability of Regulatory Biodegradation Tests. *Environ. Sci. Pollut. Res.* **2014**, *21* (16), 9511–9521. https://doi.org/10.1007/s11356-013-2064-4.

(16) Ott, A.; Martin, T. J.; Whale, G. F.; Snape, J. R.; Rowles, B.; Galay-Burgos, M.; Davenport, R. J. Improving the Biodegradability in Seawater Test (OECD 306). *Sci. Total Environ.* **2019**, *666*, 399–404. https://doi.org/10.1016/j.scitotenv.2019.02.167.

(17) Martin, T. J.; Goodhead, A. K.; Snape, J. R.; Davenport, R. J. Improving the Ecological Relevance of Aquatic Bacterial Communities in Biodegradability Screening Assessments. *Sci. Total Environ.* **2018**, *627*, 1552–1559. https://doi.org/10.1016/j.scitotenv.2018.01.264.

(18) Ott, A.; Martin, T. J.; Snape, J. R.; Davenport, R. J. Increased Cell Numbers Improve Marine Biodegradation Tests for Persistence Assessment. *Sci. Total Environ.* **2020**, *706*, 135621. https://doi.org/10.1016/j.scitotenv.2019.135621.

(19) OECD. *OECD 34 Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*; 2005; Vol. 34.

(20) OECD. *OECD Guideline for Testing of Chemicals (306): Biodegradability in Seawater*; Paris, 1992. https://doi.org/http://dx.doi.org/10.1787/9789264070486-en.

(21) Millipore. *Protein Concentration and Diafiltration by Tangential Flow Filtration*; 2003.

(22) Millipore. *Pellicon 2 Filters*; 2011.

(23) Hammes, F.; Berney, M.; Wang, Y.; Vital, M.; Köster, O.; Egli, T. Flow-Cytometric Total Bacterial Cell Counts as a Descriptive Microbiological Parameter for Drinking Water Treatment Processes. *Water Res.* **2008**, *42* (1–2), 269–277. https://doi.org/10.1016/j.watres.2007.07.009.

(24) Vignola, M.; Werner, D.; Wade, M. J.; Meynet, P.; Davenport, R. J. Medium Shapes the Microbial Community of Water Filters with Implications for Effluent Quality. *Water Res.* **2018**, *129*, 499–508. https://doi.org/10.1016/j.watres.2017.09.042.

(25) Vignola, M.; Werner, D.; Hammes, F.; King, L. C.; Davenport, R. J. Flow-Cytometric Quantification of Microbial Cells on Sand from Water Biofilters. *Water Res.* **2018**, *143*, 66–76. https://doi.org/10.1016/j.watres.2018.05.053.

(26) Brown, M. R.; Camézuli, S.; Davenport, R. J.; Petelenz-Kurdziel, E.; Øvreås, L.; Curtis, T. P. Flow Cytometric Quantification of Viruses in Activated Sludge. *Water Res.* **2015**, *68*, 414–422. https://doi.org/10.1016/j.watres.2014.10.018.

(27) Hammes, F.; Egli, T. Cytometric Methods for Measuring Bacteria in Water: Advantages, Pitfalls and Applications. *Anal. Bioanal. Chem.* **2010**, *397* (3), 1083–1095. https://doi.org/10.1007/s00216-010-3646-3.

(28) ECHA. Guidance for Monomers and Polymers - Guidance for the Implementation of REACH. **2012**, No. April, 1–25.

(29) OECD. *OECD Guideline for Testing of Chemicals (301): Ready Biodegradability*; Paris, 1992.

(30) Jouanneau, S.; Recoules, L.; Durand, M. J.; Boukabache, A.; Picot, V.; Primault, Y.; Lakel,

669 A.; Sengelin, M.; Barillon, B.; Thouand, G. Methods for Assessing Biochemical Oxygen
670 Demand (BOD): A Review. *Water Res.* **2014**, *49* (1), 62–82.
671 https://doi.org/10.1016/j.watres.2013.10.066.
672 (31) ECHA. *Guidance on Information Requirements and Chemical Safety Assessment, Chapter*
673 *R.11: PBT/VPvB Assessment (Version 3.0)*; 2017. https://doi.org/10.2823/128621.
674 (32) R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria
675 2015.
676 (33) WHO. *Heteorotrophic Plate Counts and Drinking-Water Safety*; IWA Publishing: London,
677 2003. https://doi.org/10.7717/peerj.1259.
678 (34) Gensberger, E. T.; Gössl, E.-M.; Antonielli, L.; Sessitsch, A.; Kostić, T. Effect of Different
679 Heterotrophic Plate Count Methods on the Estimation of the Composition of the Culturable
680 Microbial Community. *PeerJ* **2015**, *3*, e862. https://doi.org/10.7717/peerj.862.
681 (35) ECETOC. *Assessment of the Biodegradation of Chemicals in the Marine Environment*
682 *Technical Report No. 54*; Brussels, 1993.
683 (36) Franić, Z. Estimation of the Adriatic Sea Water Turnover Time Using Fallout 90Sr as a
684 Radioactive Tracer. *J. Mar. Syst.* **2005**, *57* (1–2), 1–12.
685 https://doi.org/10.1016/j.jmarsys.2004.11.005.
686 (37) Csanady, G. T. *Circulation in the Coastal Ocean*; Dr. Reidel Publishing Company:
687 Dordrecht, Holland, 1982.
688 (38) Comber, M.; Holt, M. *Developing a Set of Reference Chemicals for Use in Biodegradability*
689 *Tests for Assessing the Persistency of Chemicals*; 2010.
690 (39) ECETOC. *Refined Approaches for Risk Assessment of PBT/VPvB Chemicals Technical*
691 *Report No.112*; Brussels, 2011.

(40) Caulfield, M. J.; Qiao, G. G.; Solomon, D. H. Some Aspects of the Properties and Degradation of Polyacrylamides. *Chem. Rev.* **2002**, *102* (9), 3067–3084. https://doi.org/10.1021/cr010439p.

(41) El-Mamouni, R.; Frigon, J. C.; Hawari, J.; Marroni, D.; Guiot, S. R. Combining Photolysis and Bioprocesses for Mineralization of High Molecular Weight Polyacrylamides. *Biodegradation* **2002**, *13* (4), 221–227. https://doi.org/10.1023/A:1021272519714.

(42) Seybold, C. A. Polyacrylamide Review: Soil Conditioning and Environmental Fate. *Commun. Soil Sci. Plant Anal.* **1994**, *25* (11–12), 2171–2185.

(43) Poursat, B. A. J.; van Spanning, R. J. M.; de Voogt, P.; Parsons, J. R. Implications of Microbial Adaptation for the Assessment of Environmental Persistence of Chemicals. *Crit. Rev. Environ. Sci. Technol.* **2019**, *49* (23), 2220–2255. https://doi.org/10.1080/10643389.2019.1607687.

(44) Poursat, B. A. J.; van Spanning, R. J. M.; Braster, M.; Helmus, R.; de Voogt, P.; Parsons, J. R. Long-Term Exposure of Activated Sludge in Chemostats Leads to Changes in Microbial Communities Composition and Enhanced Biodegradation of 4-Chloroaniline and N-Methylpiperazine. *Chemosphere* **2020**, *242*, 125102. https://doi.org/10.1016/j.chemosphere.2019.125102.

(45) Poursat, B. A. J.; van Spanning, R. J. M.; Braster, M.; Helmus, R.; de Voogt, P.; Parsons, J. R. Biodegradation of Metformin and Its Transformation Product, Guanylurea, by Natural and Exposed Microbial Communities. *Ecotoxicol. Environ. Saf.* **2019**, *182* (May), 109414. https://doi.org/10.1016/j.ecoenv.2019.109414.

(46) ECHA. *Guidance on Information Requirements and Chemical Safety Assessment, Chapter R . 7b : Endpoint Specific Guidance (Version 1.0)*; 2008.

(47) ECETOC. Persistence of Chemicals in the Environment. ECETOC Technical Report. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels. **2003**, No. 90.

(48) Birch, H.; Hammershøj, R.; Comber, M.; Mayer, P. Biodegradation of Hydrocarbon Mixtures in Surface Waters at Environmentally Relevant Levels - Effect of Inoculum Origin on Kinetics and Sequence of Degradation. *Chemosphere* **2017**, *184*, 400–407. https://doi.org/10.1016/j.chemosphere.2017.05.169.

(49) ISO. *ISO 10634:1995. Water Quality; Guidance for the Preparation and Treatment of Poorly Water-Soluble Organic Compounds for the Subsequent Evaluation of Their Biodegradability in an Aqueous Medium*; 1995.

(50) Brown, D. M.; Hughes, C. B.; Spence, M.; Bonte, M.; Whale, G. Assessing the Suitability of a Manometric Test System for Determining the Biodegradability of Volatile Hydrocarbons. *Chemosphere* **2018**, *195*, 381–389. https://doi.org/10.1016/j.chemosphere.2017.11.169.

(51) Elf Akvamiljo. *Biodegradability of Chemical Substances in Seawater- Results of the Four OSPARCOM Ring Tests- Final Report*; 1996.

(52) Toräng, L.; Nyholm, N. Biodegradation Rates in Adapted Surface Water Can Be Assessed Following a Preadaptation Period with Semi-Continuous Operation. *Chemosphere* **2005**, *61* (1), 1–10. https://doi.org/10.1016/j.chemosphere.2005.03.040.

Multi-laboratory validation of a new marine biodegradation screening test for chemical persistence assessment

Amelie Ott [a], Timothy J. Martin [a,1], Kishor Acharya [a], Delina Y. Lyon [b], Nik Robinson [c], Bob Rowles [d], Jason R. Snape [a,e,f], Ian Still [c,2], Graham F. Whale [g,3], Vurtice C. Albright III [h], Petra Bäverbäck [i], Nicola Best [j], Ruth Commander [k], Curtis Eickhoff [l], Sarah Finn [m], Björn Hidding [n], Heiko Maischak [o], Katherine A. Sowders [p], Masanori Taruki [q], Helen E. Walton [d], Aina Charlotte Wennberg [r], Russell J. Davenport [a]*

[a] Newcastle University, School of Engineering, Newcastle upon Tyne, NE1 7RU, United Kingdom

[b] Shell Oil Company, 150 N. Dairy Ashford Rd, Houston, TX 77079, United States

[c] European Oilfield Specialty Chemicals Association (EOSCA), Aberdeen, AB11 6YQ, United Kingdom

[d] Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, NR33 0HT, United Kingdom

[e] AstraZeneca Global Environment, Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TF, United Kingdom

[f] University of Warwick, School of Life Sciences, Gibbet Hill Campus, Coventry CV4 7AL, United Kingdom

[g] Shell International Ltd, Risk Science Team, 4 York road, London, SE1 7NA, United Kingdom

[h] The Dow Chemical Company, Midland, Toxicology & Environmental Research & Consulting, MI 48674, United States

[i] Schlumberger, Sandslikroken 140, 5254 Sandsli, Norway

[j] Covance CRS Research Limited, Shardlow Business Park, London Road, Derby, DE72 2GD, United Kingdom

[k] Scymaris Ltd., Brixham Laboratory, Brixham, TQ5 8BA, United Kingdom

[l] Nautilus Environmental Company Inc., Burnaby, BC, V5A 4N7, Canada

[m] National Oilwell Varco (NOV), Flotta, Stromness, Orkney, KW16 3NP, United Kingdom

[n] BASF SE, Carl-Bosch-Straße 38, 67056 Ludwigshafen am Rhein, Germany

[o] Noack Laboratorien GmbH, Käthe-Paulus-Straße 1, 31157 Sarstedt, Germany

[p] Baker Hughes – Environmental Services Group, 369 Marshall Ave., Webster Groves, Missouri 63119, United States

[q] Chemicals Evaluation and Research Institute, Japan, Kurume (CERI Kurume), 3-2-7 Miyanojin, Kurume-shi, Fukuoka 839-0801, Japan

[r] Norwegian Institute for Water Research (NIVA), Gaustadalléen 21, 0349 Oslo, Norway

Present Addresses

[1] Department for Environment, Food and Rural Affairs, Lancaster House, Newcastle upon Tyne, NE4 7JB, United Kingdom

[2] Nalco Champion, WSW Building, Peterseat Drive, Aberdeen, AB12 3HT, United Kingdom

[3] Whale Environmental Consultancy Ltd., 55 Earlsway Curzon Park, Chester, CH4 8AZ, United Kingdom

SUPPORTING INFORMATION

Pages: 38

Figures: 21

Tables: 16

Methods M1: Test protocol modifications in the imBST$_{MR}$ from the pre-validated marine erBST.

The new test (imBST$_{MR}$) was based on a previous intra-laboratory validated marine environmentally relevant BST (erBST) [1], but differed in following aspects to incorporate recommendations from stakeholders and other studies:[1–3]

- Terminology: While the microbiome in the erBST and imBST$_{MR}$ aims to better represent the samples environment, other BST conditions still do not represent the environment well e.g. high test chemical concentrations and high incubation temperatures. Consequently, the terminology "environmentally relevant" was replaced with improved/new for the imBST$_{MR}$.
- Biodegradation measurement: To overcome potential biodegradation underestimations in OECD 301B tests [1,4–6], the imBST$_{MR}$ monitored biodegradation with MRs in a modified OECD 301F test.
- TFF: In the imBST$_{MR}$, the TFF protocol was optimized to incorporate an additional filtrate pump to reduce membrane wall pressures. No backflushing was performed to preserve membrane integrity.
- Test chemicals: Due to equipment and licensing limitations at CROs, test chemicals were not radiolabeled (^{14}C) in the ring test. Higher test chemical concentrations were employed in the new and revised MR test in comparison to the pre-validation study.[1] In MR tests, chemical stock solutions were prepared with seawater instead of OECD mineral medium to circumvent seawater dilution in the test vessel (of bacterial cell concentrations and salinity).[1] However, it should be noted that the high salt

concentrations in seawater can modify the solubility and related properties of some organic chemicals.[7]

- Test medium: Phosphate nutrient additions (OECD mineral medium solution a) in the MR tests followed the OECD 301F protocol [8] and were 10 × higher than in the pre-validation study which followed the OECD 306 recipe.[1,9] The OECD guidelines do not explain this difference, but the OECD 306 method probably requires less phosphate due to the natural buffering capacity of seawater [10] and lower test chemical concentrations employed. To account for increased test chemical levels, more phosphate was added in the MR tests. However, it should be noted that this alteration was expected to have little or no effect as phosphate is added to excess in all OECD BSTs and no adverse effects have been observed with increased phosphate levels in BSTs.[10,11]

Figure S1. Locations of laboratories participating in the ring test.

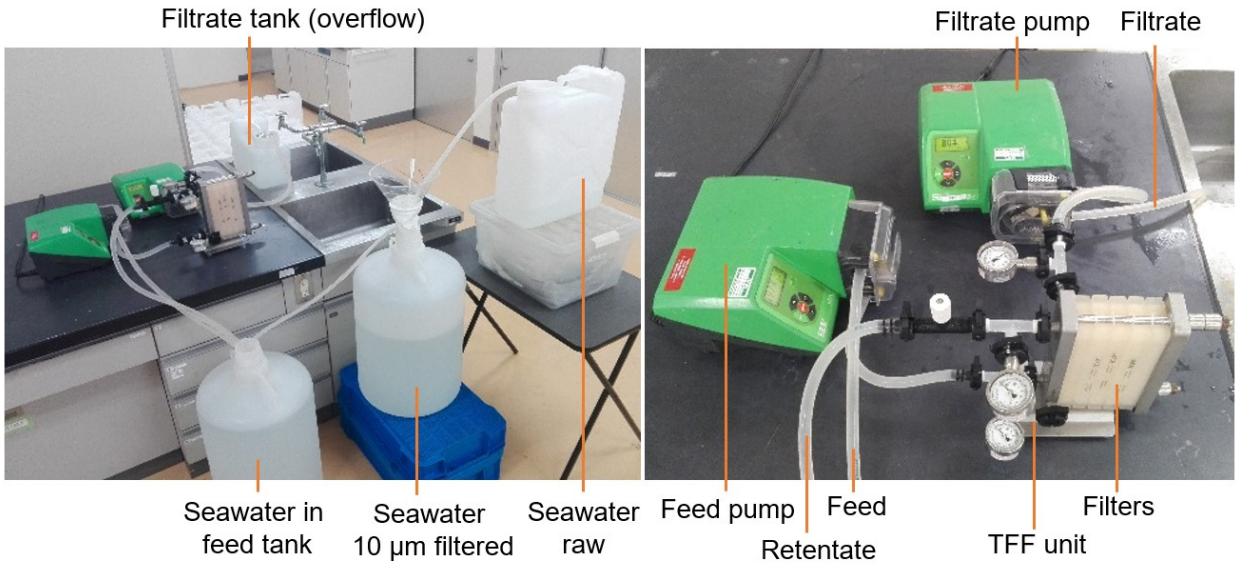


Figure S2. Example tangential flow filtration setup to increase bacterial cell numbers in seawater.

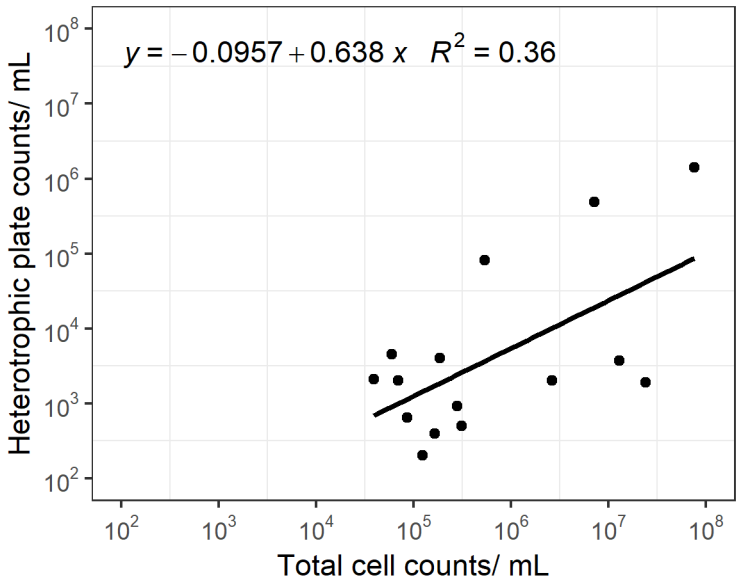
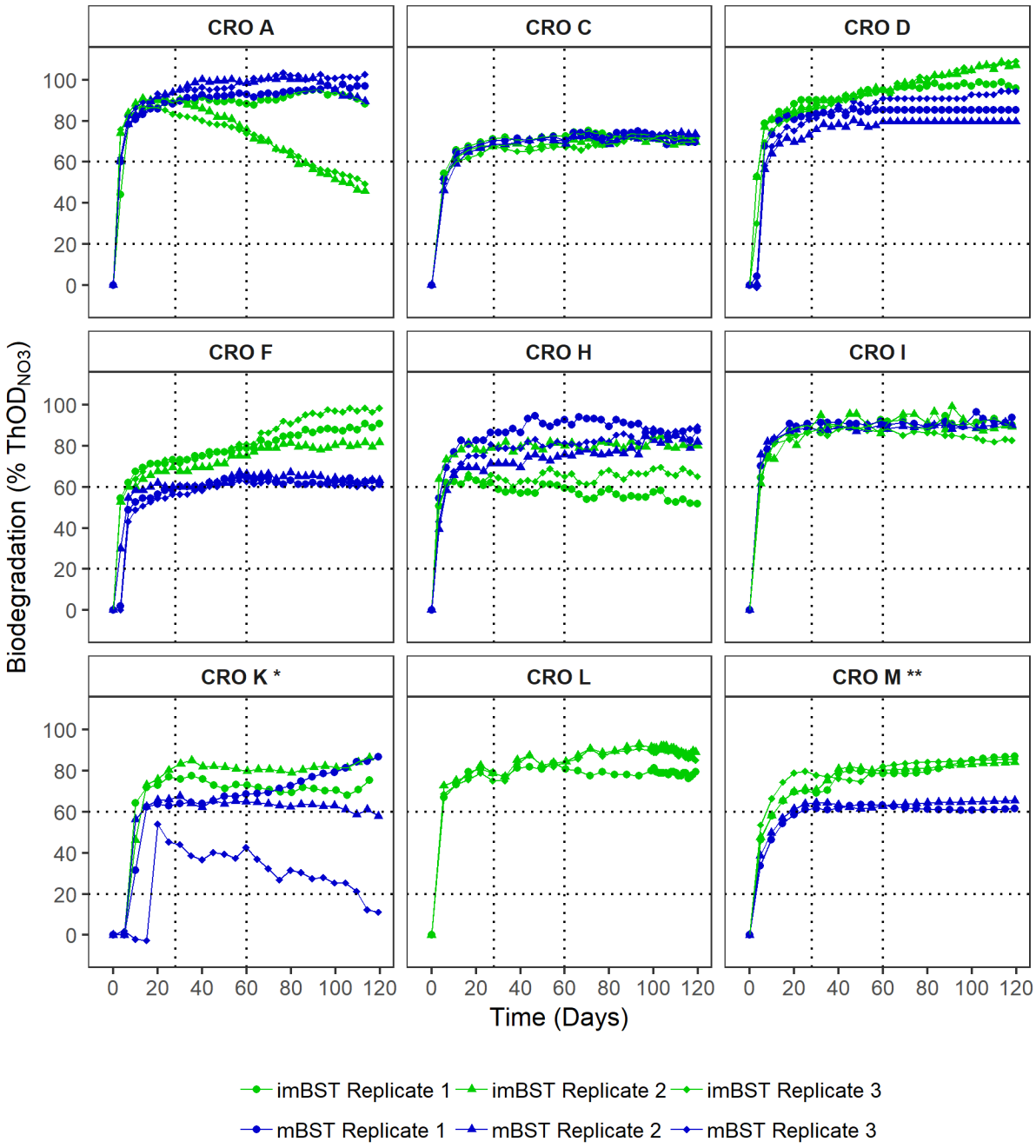


Figure S3. Correlation and linear regression between heterotrophic plate counts (measured using different culture methods) and total cell concentrations (measured by flow cytometry) in seawater samples (S1, S2, S3) where both measurement methods were conducted.

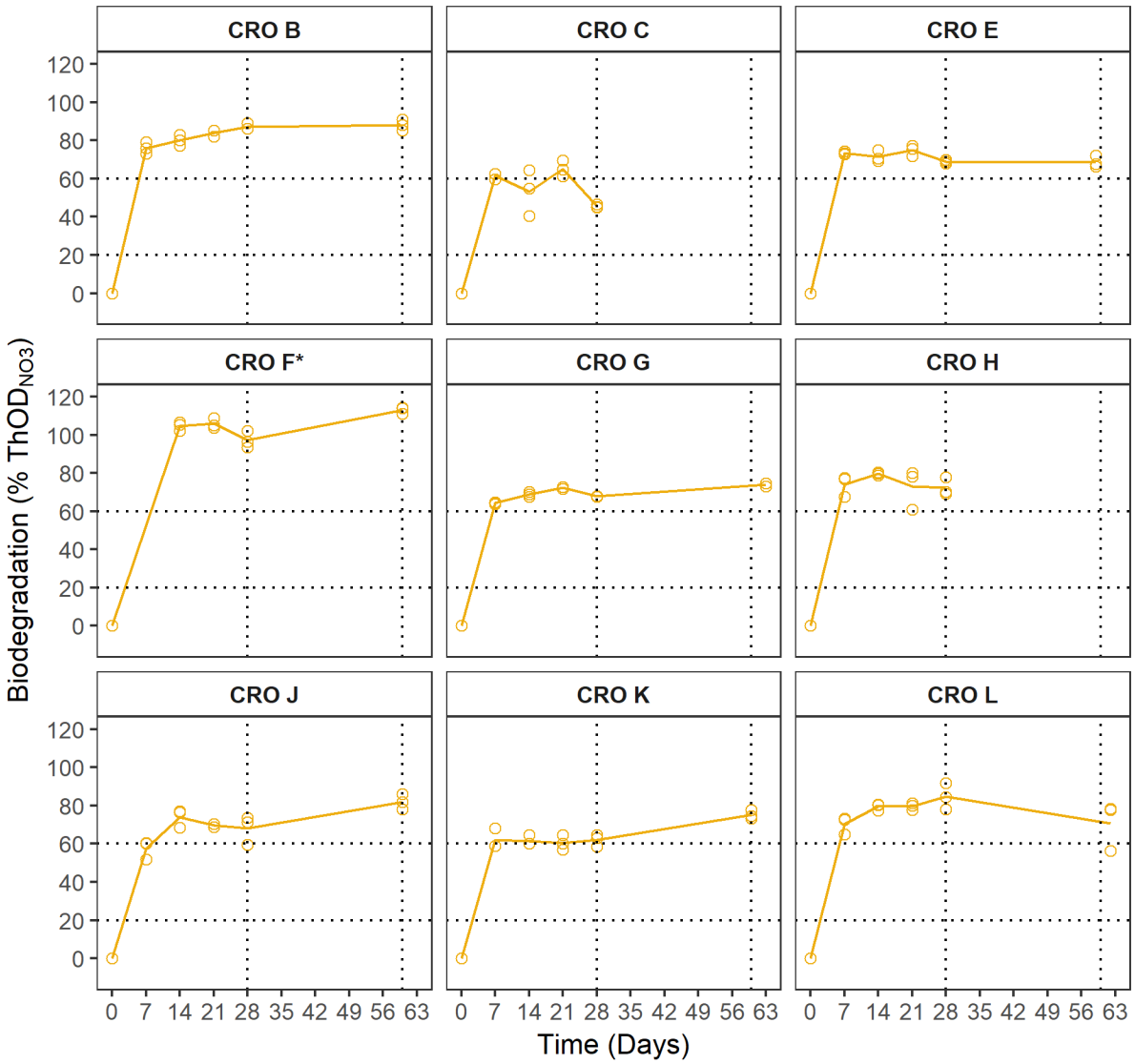
Note: For the following $imBST_{MR}$ and $mBST_{MR}$ biodegradation plots, every 20th data point was plotted for CRO A, C, D, F, H, K, L, M (automatic recordings every 4- 7 hours) and every 3rd data point for CRO I (manual daily recordings on weekdays). For the $OECD306_{CB}$ biodegradation plots, individual measurements of the sacrificial BOD bottles are plotted together with a line representing the arithmetic mean.



101

102 **Figure S4.** Biodegradation of sodium benzoate in the mBST$_{MR}$ and imBST$_{MR}$. * For removed outlier, see
103 Figure S20. ** Biodegradation based on CO_2 production instead of O_2 consumption.

104



105

106 **Figure S5.** Biodegradation of sodium benzoate in the OECD306$_{CB}$. * For removed outlier, see Figure S19.

107

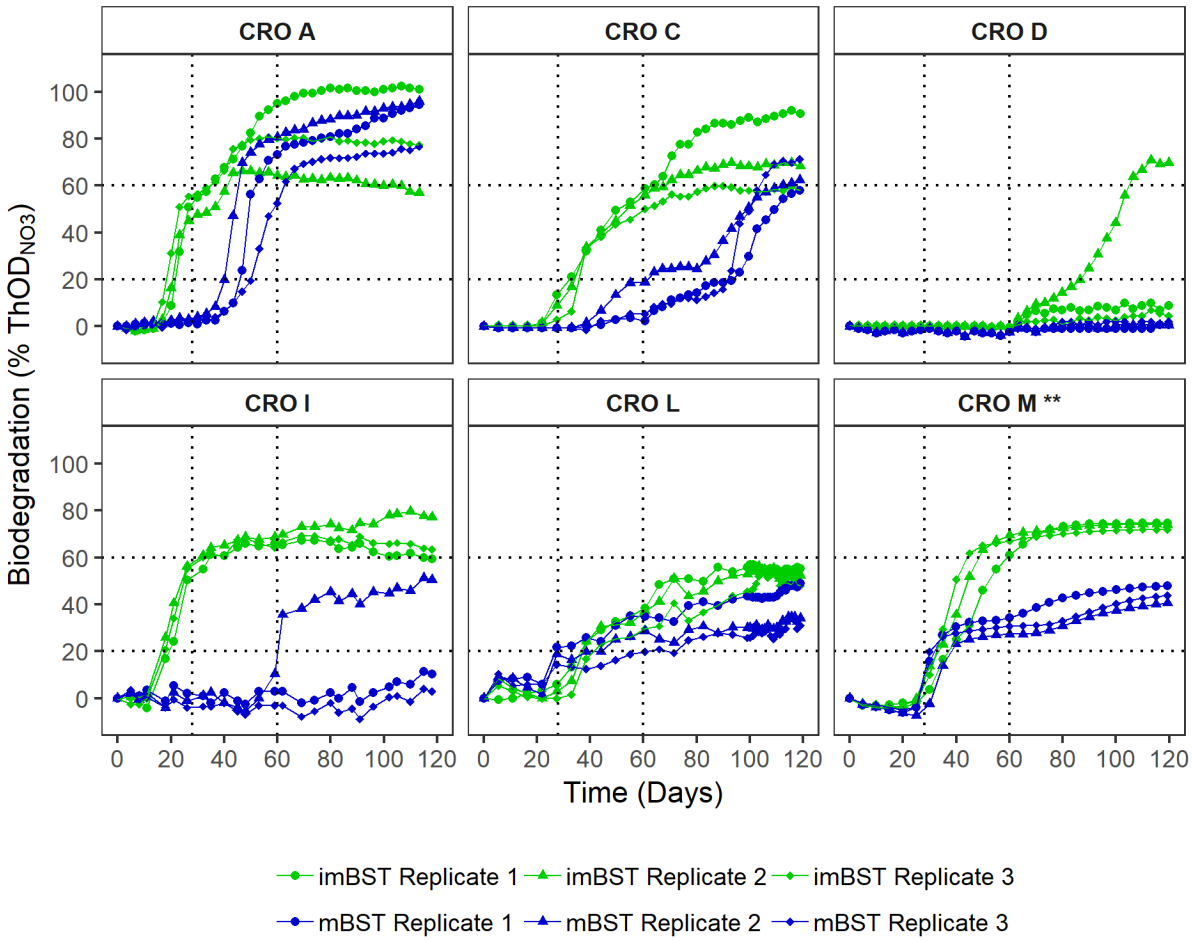
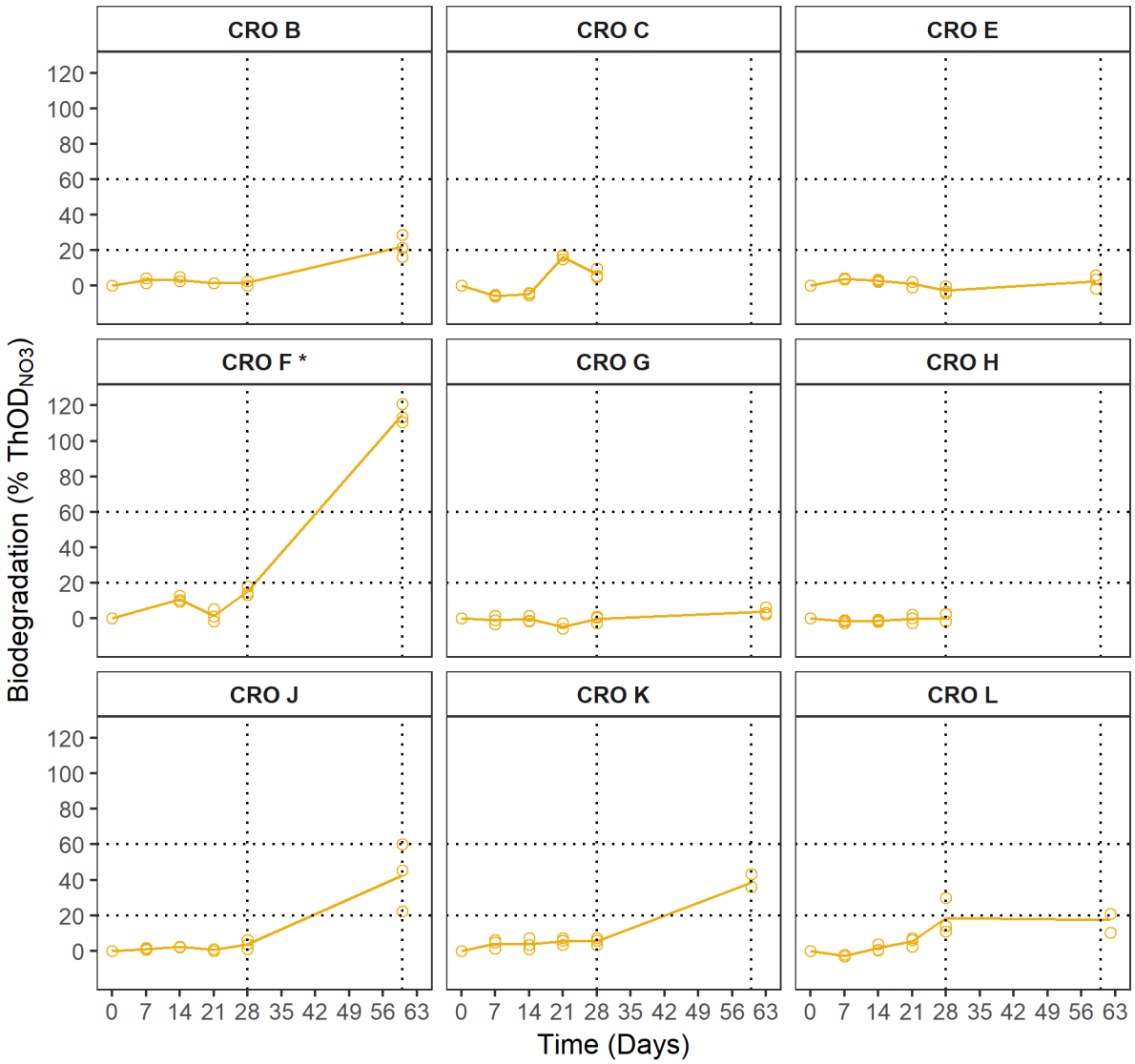


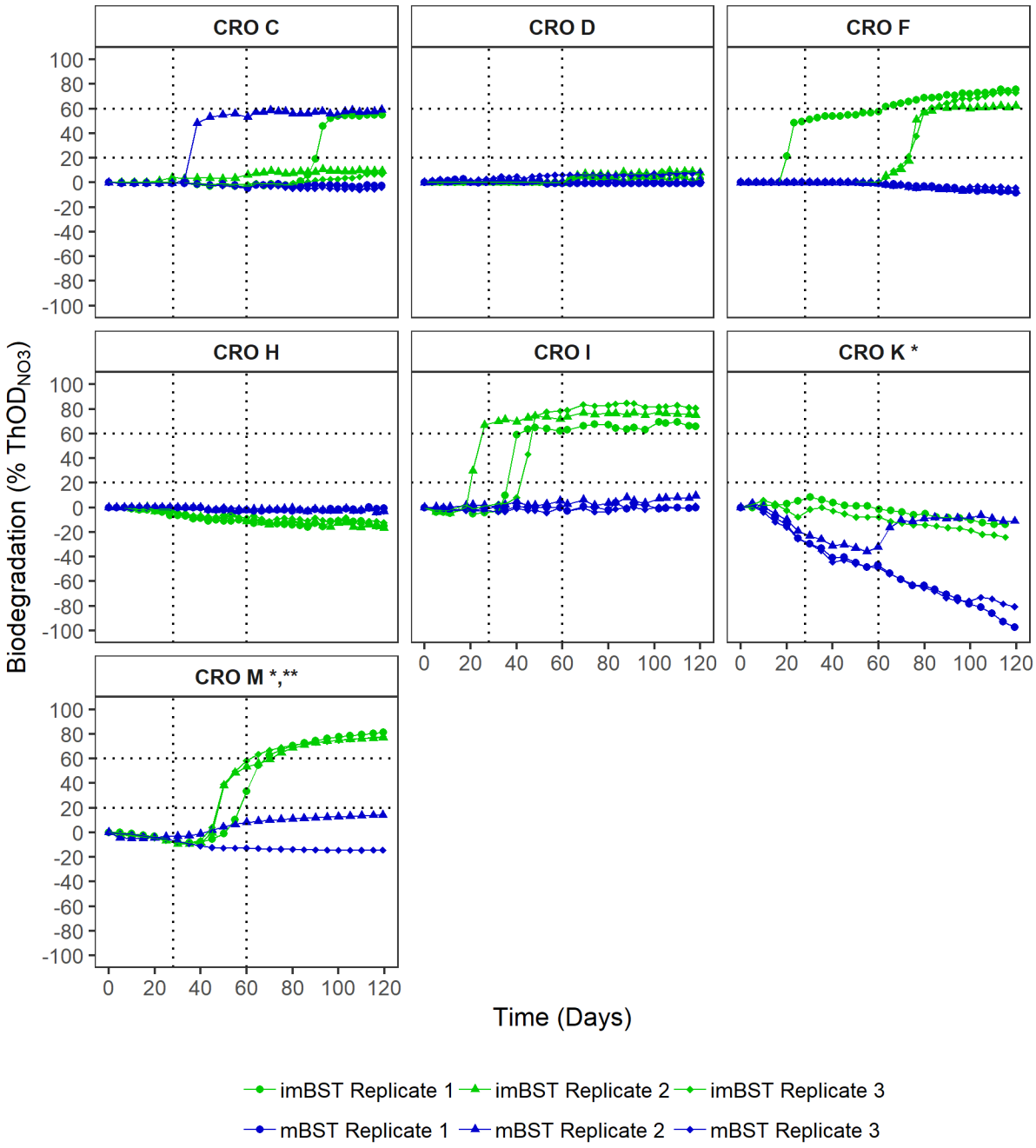
Figure S6. Biodegradation of triethanolamine in the mBST$_{MR}$ and imBST$_{MR}$. ** Biodegradation based on CO_2 production instead of O_2 consumption.



111

112 **Figure S7.** Biodegradation of triethanolamine in the OECD306$_{CB}$. * For removed outlier, see Figure S19.

113



114

115 **Figure S8.** Biodegradation of 4-nitrophenol in the mBST$_{MR}$ and imBST$_{MR}$. * For removed outlier, see
116 Figure S20. ** Biodegradation based on CO_2 production instead of O_2 consumption.

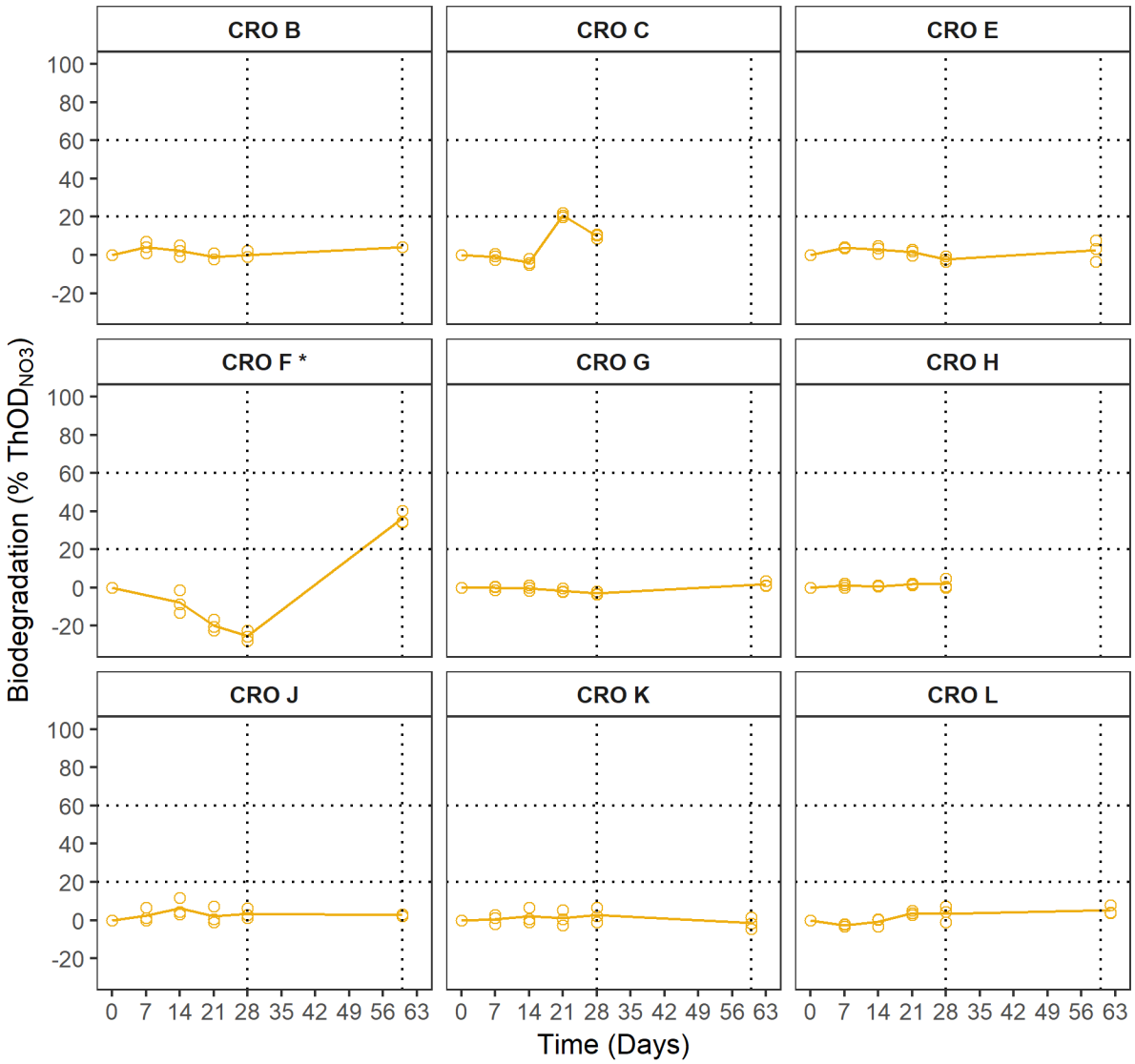


Figure S9. Biodegradation of 4-nitrophenol in the OECD306$_{CB}$. * For removed outlier, see Figure S19.

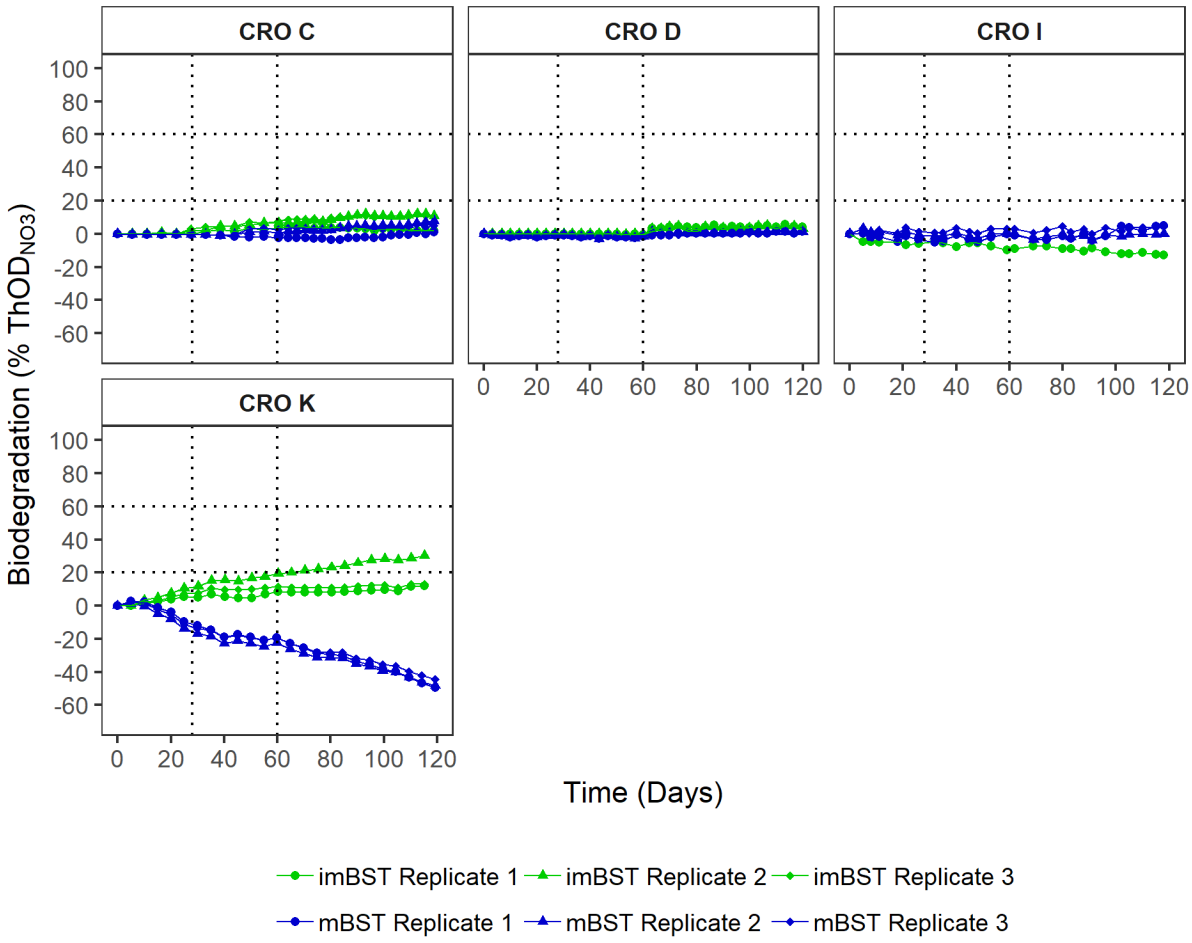
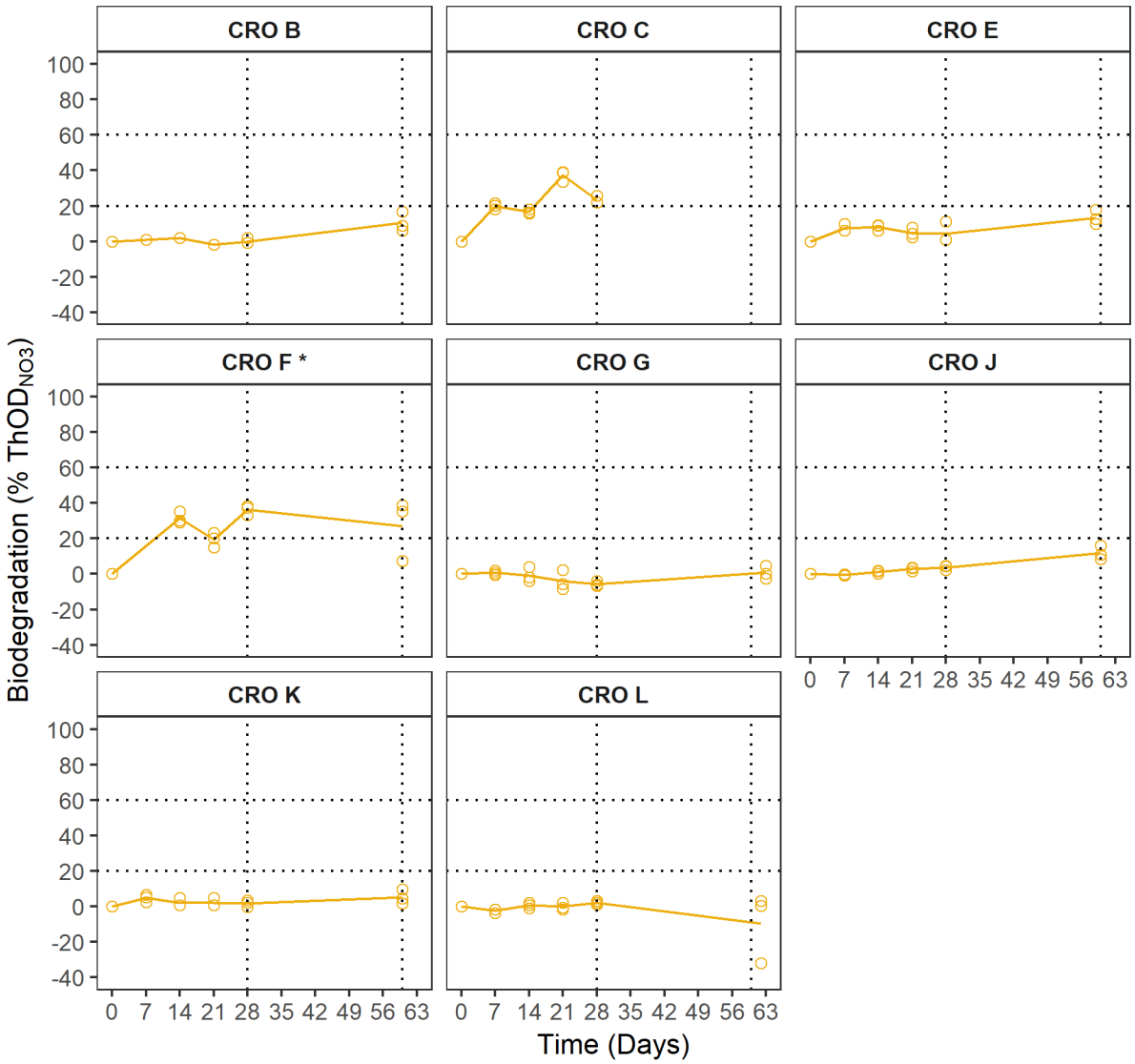


Figure S10. Biodegradation of anionic polyacrylamide in the $mBST_{MR}$ and $imBST_{MR}$.



121

122 **Figure S11.** Biodegradation of anionic polyacrylamide in the OECD306$_{CB}$. * For removed outlier, see
123 Figure S19.

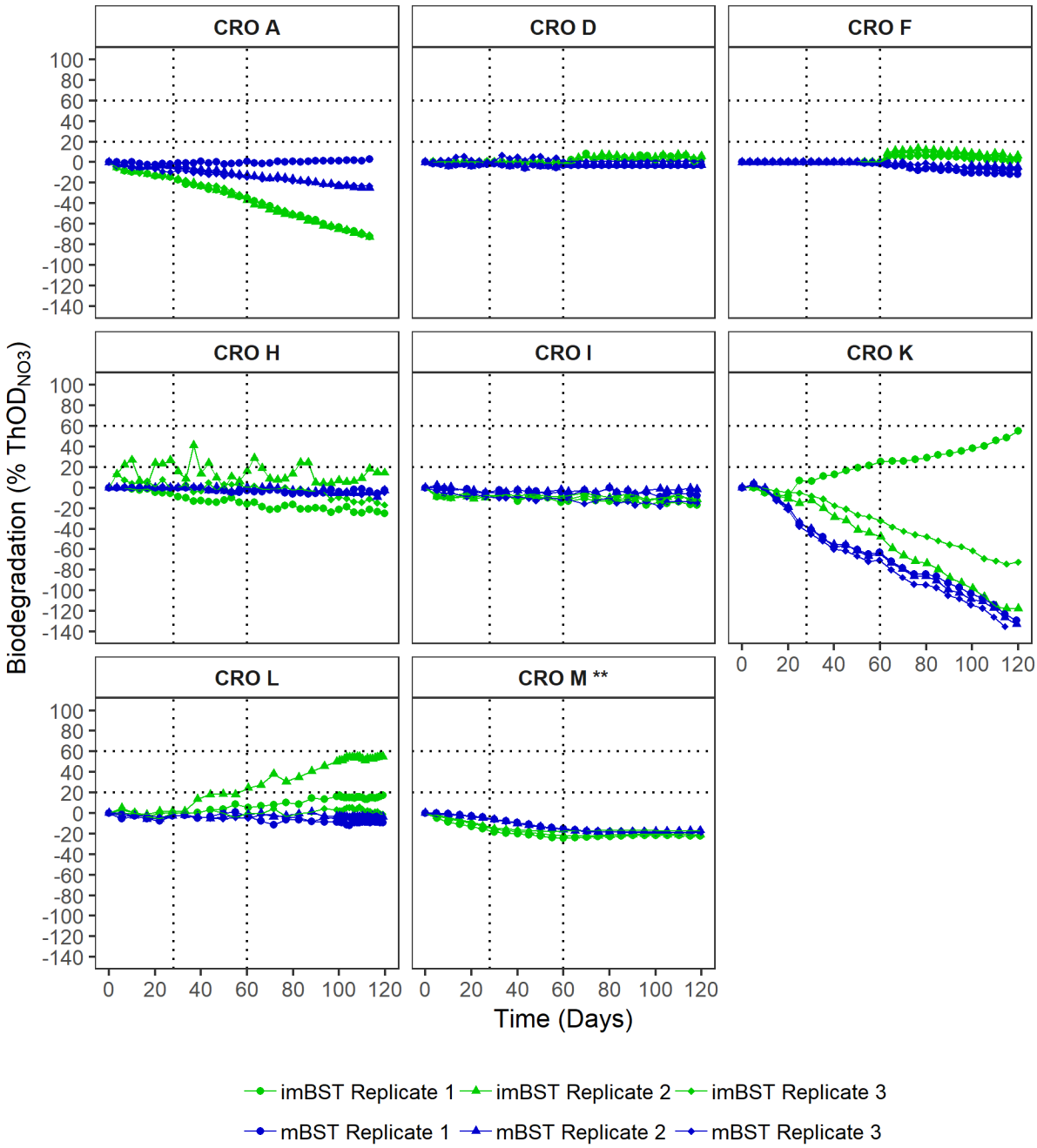
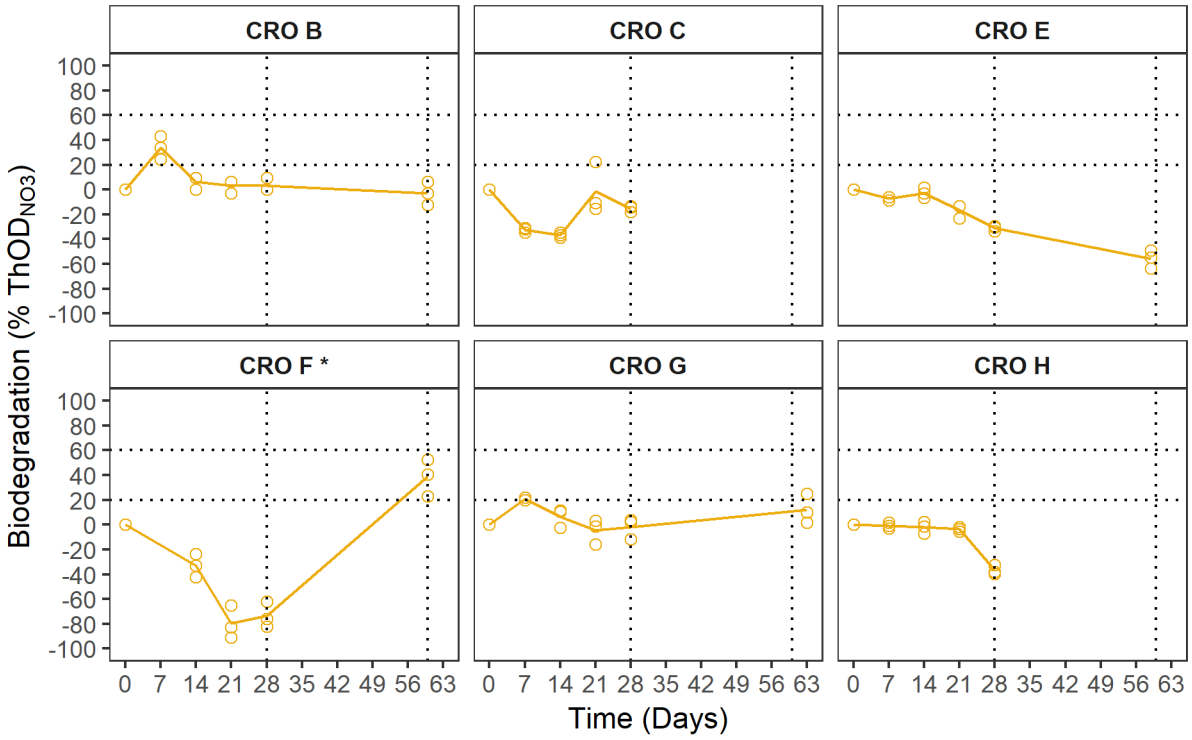


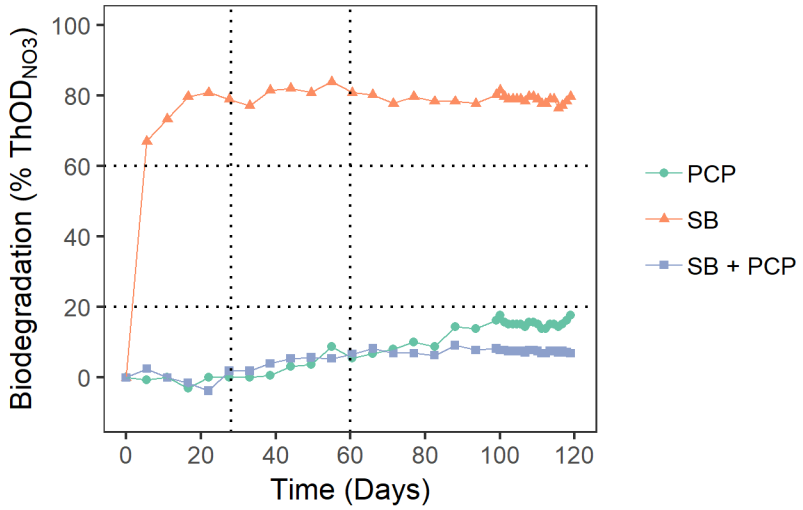
Figure S12. Biodegradation of pentachlorophenol in $mBST_{MR}$ and $imBST_{MR.}$ ** Biodegradation based on CO_2 production instead of O_2 consumption.



129

130 **Figure S13.** Biodegradation of pentachlorophenol in the OECD306$_{CB}$. * For removed outlier, see Figure
131 S19.

132



133

134 **Figure S14.** Pentachlorophenol (PCP) toxicity control with sodium benzoate (SB) for the imBST$_{MR}$ by
135 CRO L.

136

137

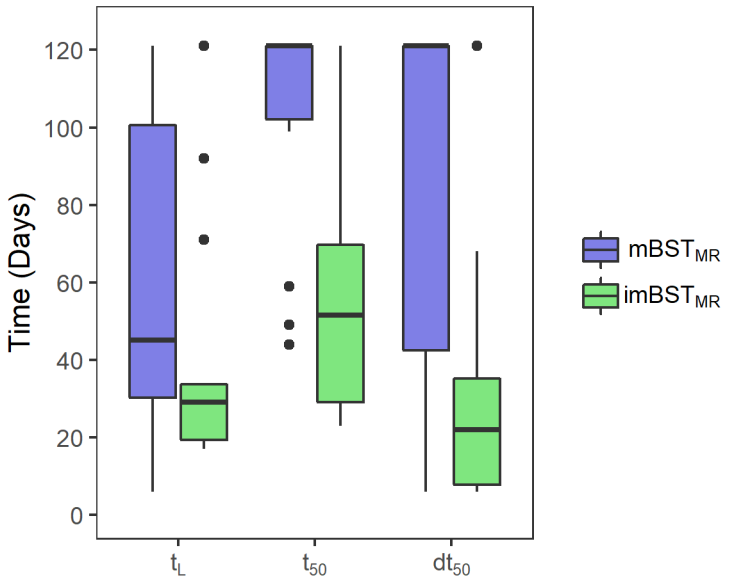


Figure S15. Increased cell numbers in the new test reduce t_L (time to 10% degradation), t_{50} (time to 50% degradation) and dt_{50} (t_{50} – t_L) for triethanolamine. For non-degrading $mBST_{MR}$ and $imBST_{MR}$ replicates, descriptor values were set to 121 days.

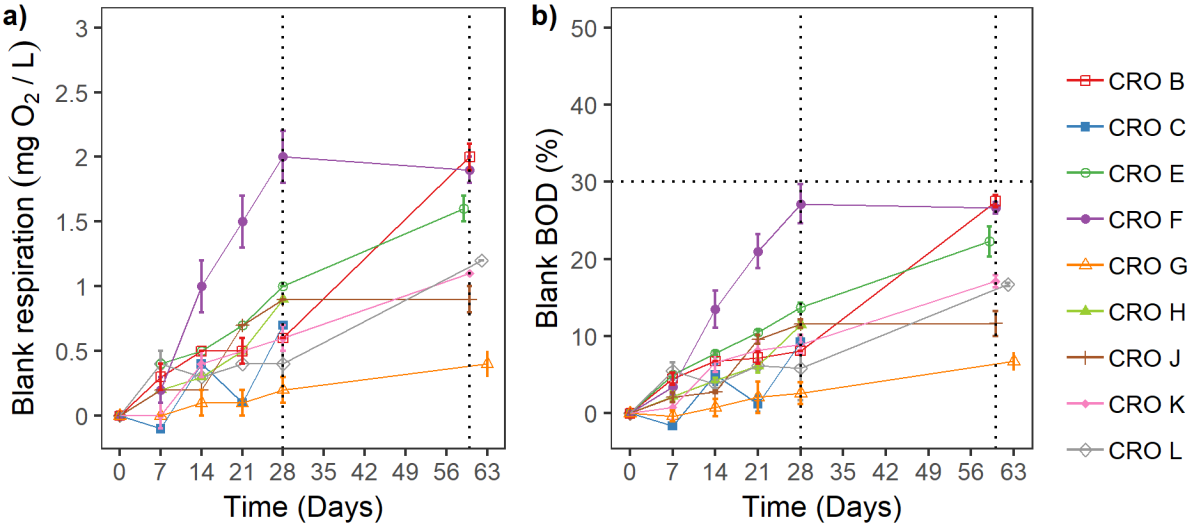
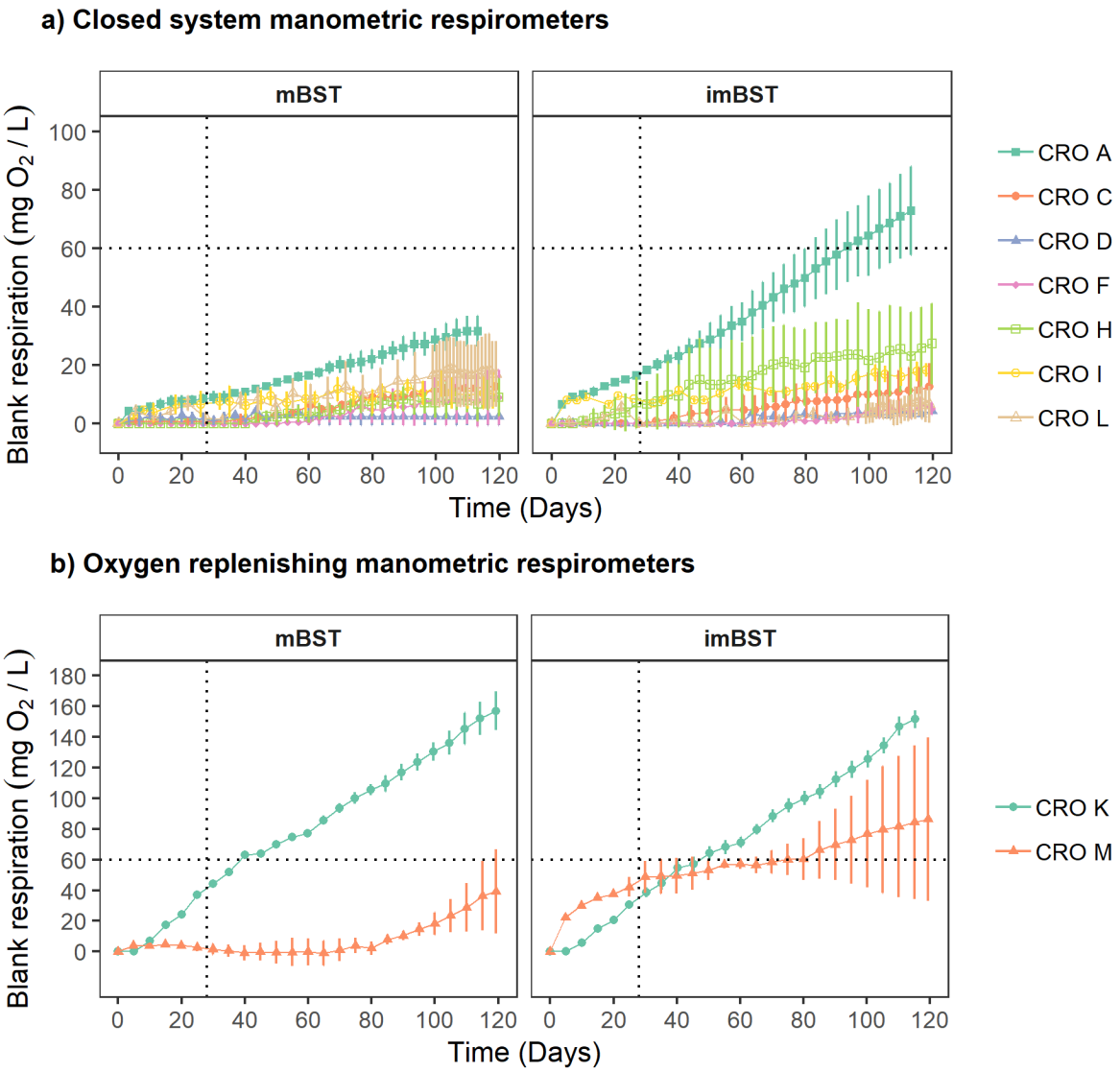


Figure S16. $OECD306_{CB}$ blank respiration over 60 days across CROs expressed in mg O_2 L^{-1} (a) and % (b). Dotted horizontal line at 30% BOD (b) refers to blank threshold defined in test guideline OECD 306.[9]



146

147 **Figure S17.** imBST$_{MR}$ and mBST$_{MR}$ blank respiration in closed manometric respirometer systems (a and
148 b) and oxygen replenishing manometric respirometer systems (c and d). Dotted horizontal line at
149 60 mg O_2 L^{-1} blank respiration and 28 days refers to blank threshold defined in test guideline OECD 301F.[8]

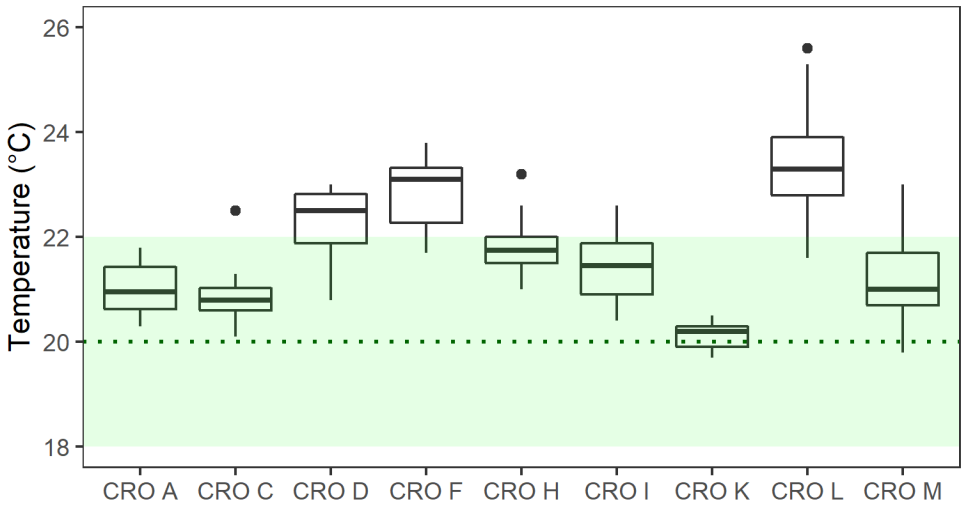


Figure S18. Boxplots showing temperatures measured in $mBST_{MR}$ and $imBST_{MR}$ test media after 120 day incubation period across CROs. Green indicates 20 ± 2°C range.

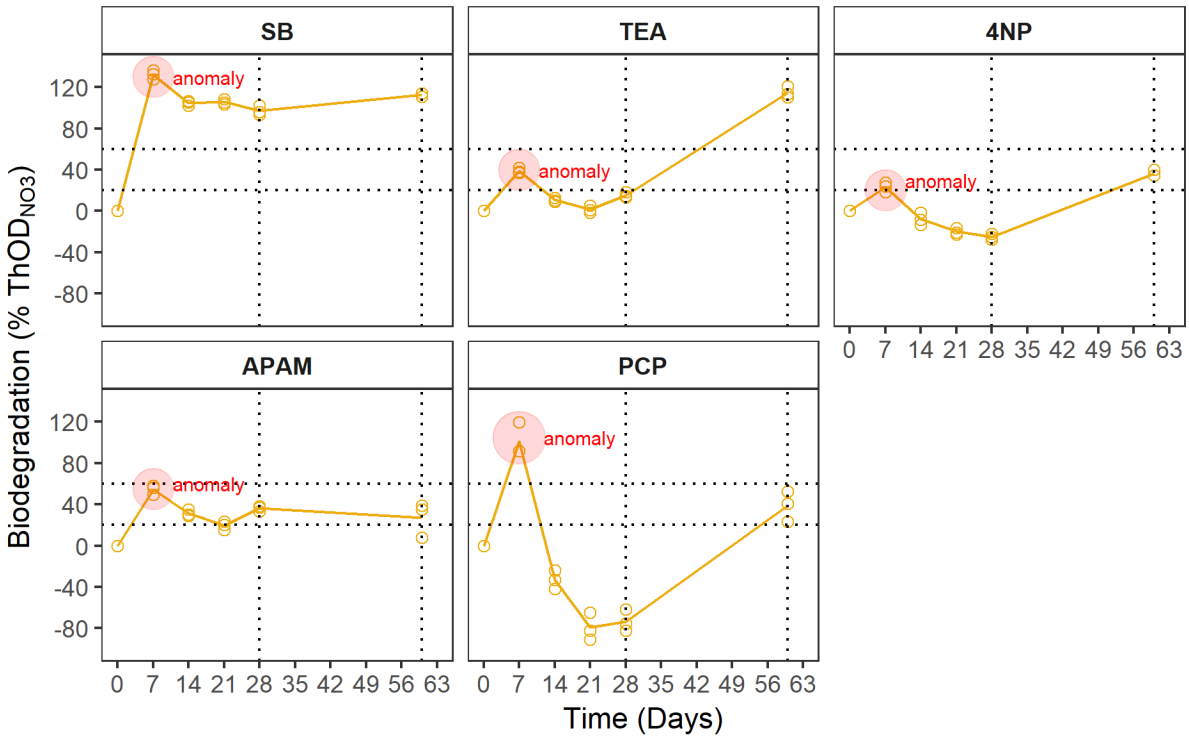
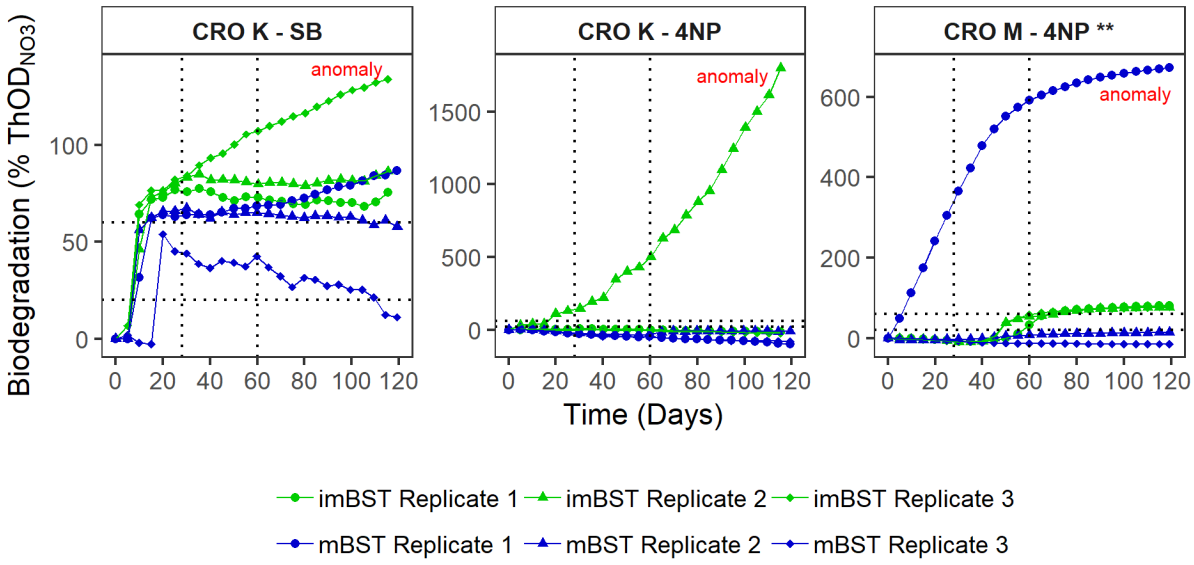


Figure S19. Systematic anomalous results (marked with a red circle) observed in the $OECD306_{CB}$ at CRO F. SB: sodium benzoate. TEA: triethanolamine. 4NP: 4-nitrophenol. APAM: anionic polyacrylamide. PCP: pentachlorophenol.



159

160 **Figure S20.** Outliers observed in the mBST$_{MR}$ and imBST$_{MR}$. SB: sodium benzoate. 4NP: 4-nitrophenol.
161 ** Biodegradation based on CO_2 production instead of O_2 consumption.

162

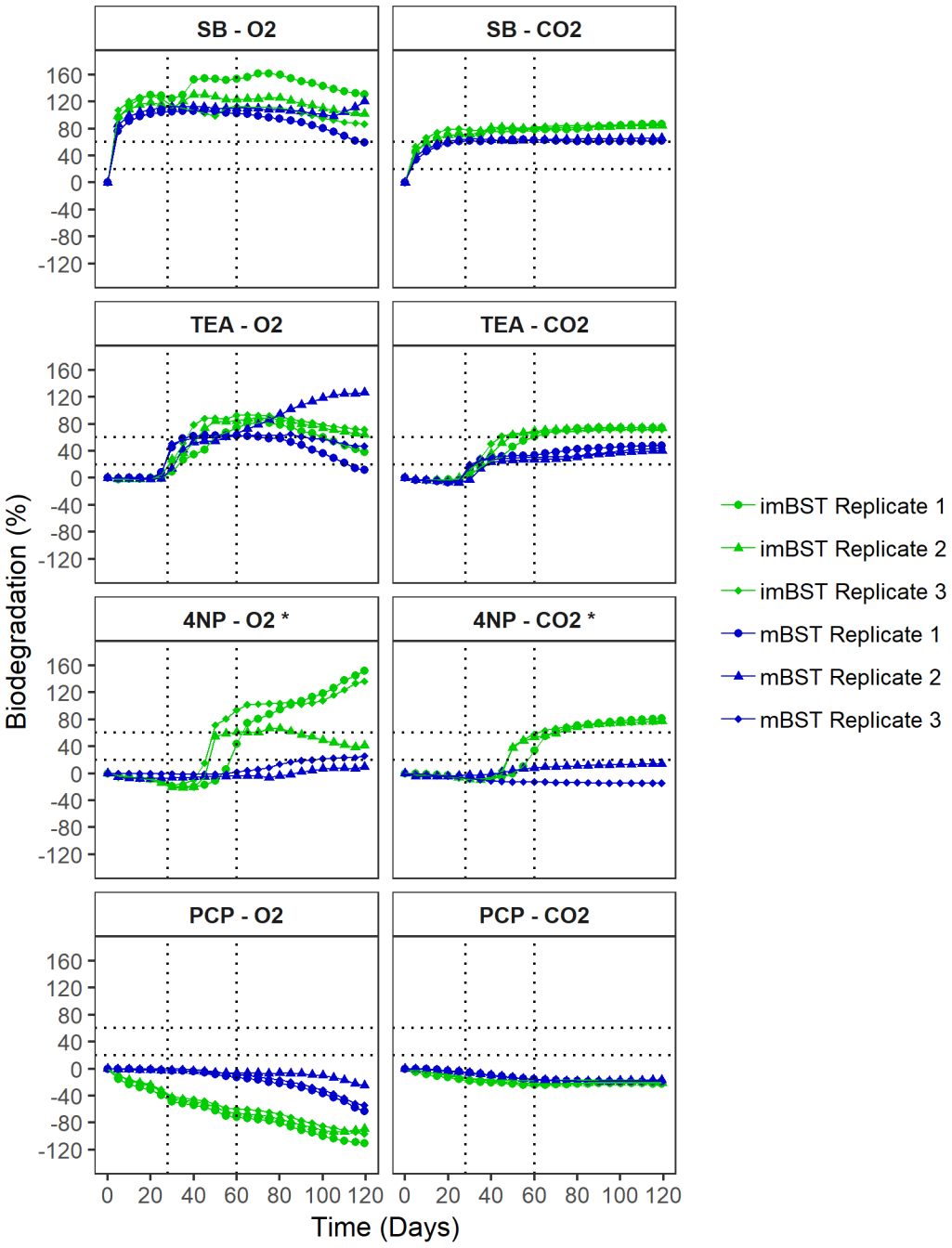


Figure S21. Comparison of biodegradation values calculated based on O_2 consumption and CO_2 production for CRO M. SB: sodium benzoate. TEA: triethanolamine. 4NP: 4-nitrophenol. PCP: pentachlorophenol. * For removed outlier, see Figure S20.

Table S1. Instruments and methods employed at the CROs for the $mBST_{MR,}$ $imBST_{MR}$ and $OECD306_{CB}$.

CRO →	A	B	C	D	E	F	G	H	I	J	K	L	M
$mBST_{MR}$ and $imBST_{MR}$													
Manometric respirometer	WTW OxiTop Control	—	WTW OxiTop Control	WTW OxiTop Control	—	WTW OxiTop IS	—	WTW OxiTop Control	WTW OxiTop Control	—	CES multi-channel aerobic respire-meter	WTW OxiTop Control	Columbus Instrument Micro-Oxymax Respiro-meter
$OECD306_{CB}$													
Removing coarse particles		Filtration (11 µm)	Filtration (10 µm)		Filtration	Not performed	Sedimen-tation	Sedimen-tation		Sedimen-tation	Filtration (coarse filter paper)	Sedimen-tation and siphoning	
Ageing conditions	—	7 days ageing with 3 days aeration; 20°C; dark	6 days with full aeration; 20°C; dark	—	7 days with full aeration; 20°C; dark	7 days with full aeration; 20°C; dark	7 days with full aeration 18°C ± 2°C; dark	7 days with full aeration; 20°C; dark	—	7 days with no aeration; 18.4-19°C; dark	10 days with full aeration; 21°C; dark	6 days with aeration for 2h 15 min; 20°C; dark	—
DO (mg/L)		YSI 58	Days 0-14: YSI DO; Days 21-28: Mettler Toledo SevenGo pro DO		Hach HQ40d LDO101	Winkler Titration Method	YSI Oximeter model 5100	WTW Oxi 1970i		Hach HQ30d	YSI Model 57	WTW inoLab Oxi 7310	

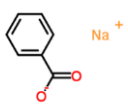
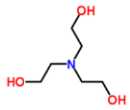
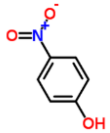
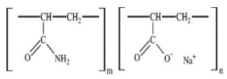
—: test setup not conducted. DO: dissolved oxygen.

Table S2. Instruments and methods employed at the CROs to characterize the seawater.

CRO →	A	B	C	D	E	F	G	H	I	J	K	L	M
pH	WTW Multi 350i	Orion Star A111	Hanna HI113 pH/ mV	HM-25R, DKK-TOA Corpor-ation	Hach PHC101 probe	Fisher Scientific AP 115	WTW InoLab pH 730	Fisher Scientific Meter 0503	See CRO A as same seawater was used	Handylab pH	Hach HQ30D	WTW pH 340i, PHM220 lab pH	Orion Star A221
T (°C)	WTW Multi 350i	YSI Pro 30	Mercury thermo-meter	Alcohol thermo-meter	Hach CDC401 probe	Hach sension5	Total immersion glass thermo-meter	Thermo Scientific Orion Star		Testo 110	Hach HQ30D	WTW Multi 3430, WTW InoLab Oxi 7310	Alcohol thermo-meter
DO (mg/L)	Hach HQ 40d	YSI 58 DO Meter	YSI 55 DO	ID-150, Iijima Electronics	Hach LDO101 probe	Hach sension5	YSI Oximeter 5100	Fisher Scientific Meter 0503		Hach HQ30d	Hach HQ30D	WTW Multi 3430, WTW Inolab Oxi 7310	HQ40d meter LBOD101r
Conduc-tivity (mS/cm)	WTW Multi 340i	YSI Pro 30	Mettler Toledo Seven Multi	CM-31P, DKK-TOA Corpor-ation	Hach CDC401 probe	Hach sension5	Not measured	Fisher Scientific Meter 0503		WTW Conducto-meter	Hach HQ30D	WTW Multi 3430, WTW inoLab Terminal Level 3 Tetracon 325 probe	YSI 3200
Salinity (ppt)	WTW Multi 340i	YSI Pro 30	Mettler Toledo Seven Multi	CM-31P, DKK-TOA Corpor-ation	Hach CDC401 probe	Hach sension5	Thermo-balance Satorius MA35	Fisher Scientific Meter 0503		WTW Conducto-meter	Hach HQ30D	WTW Multi 3430, WTW inoLab Terminal Level 3 Tetracon 325 probe	YSI 3200
HPC/mL	DEV nutrient agar	Serial extinction marine broth bottle test	np.	Trypticase soy agar	Marine Agar	APHA Method 9215	Total viable count	Marine agar		np.	Trypticase soy agar	PCA with seawater	np.

DO: dissolved oxygen. HPC: heterotrophic plate counts. np: not performed. T: temperature.

Table S3. Chemical and physical properties of reference chemicals. All data for APAM provided by chemical supplier SNF. Information for other chemicals obtained from PhysProp [12] , except for calculated $ThCO_2$ and $ThOD_{NH3/NO3}$ values [9] and chemical structures (obtained from ChemSpider [13]). All chemicals except APAM purchased from Sigma Aldrich, St. Louis, USA.

	Positive control: Sodium benzoate (SB)	**Variable degradation:** Triethanolamine (TEA)	4-Nitrophenol (4NP)	Anionic polyacrylamide (APAM)	**Negative control:** Pentachlorophenol (PCP)
CAS Formula Purity Structure	532-32-1 $C_7H_5NaO_2$ ≥ 99.0% 	102-71-6 $C_6H_{15}NO_3$ 98% 	100-02-7 $C_6H_5NO_3$ >=99% 	25937-30-8 $[C_3H_5NO]_m$ $[C_3H_3NaO_2]_l$ / 	87-86-5 $C_6H_5Cl_5O$ 97% 
Molecular weight (g/mol)	144.11	149.19	139.11	7.6 M Da	266.34
Water solubility (mg/L)	5.56 x 10^5 at 25ºC, exp.	1.00 x 10^6 at 22ºC, exp.	1.16 x 10^4 at 20ºC, exp.	100%	14 at 25ºC, exp.
Vapour pressure (mm Hg)	3.67 x 10^{-9} at 25ºC, est.	3.59 x10^{-6} at 25ºC, exp.	9.79 x 10^{-5} at 20ºC, exp.	information not available	1.10 x 10^{-4} at 25ºC, exp.
Henry's law constant at 25ºC (atm-m^3/mol)	1.09 x 10^{-7}, est.	7.05 x 10^{-13}, est.	4.15 x 10^{-10}, exp.	information not available	2.45 x 10^{-8}, exp.
Log K_{ow}	-2.27, est.	-1, exp.	1.19, exp.	-2.34, exp.	5.12, exp.
$ThOD_{NH3}$ and $ThOD_{NO3}$ (mg O_2/mg test substance)	1.67 1.67	1.61 2.04	1.15 1.61	1.25 1.88	0.54 0.54
$ThCO_2$ (mg CO_2/mg test substance)	2.14	1.77	1.90	information not available	0.99

est: estimated data. exp: experimental data.

Table S4. Explanation on test chemical selection and assigned "correct" biodegradation classification to compare the results of the standard OECD 306 test, the revised test and the new test. Note that these assigned biodegradation classifications are not definitive as they are restricted by the quality and scope of the evaluated data.[1,14]

Assigned reference biodegradation classification	Previously reported biodegradation data and explanation on test chemical selection
Sodium benzoate (SB); rapidly biodegradable – non persistent	– ECHA database: Readily biodegradable;[15] – Comber and Holt (2010) grouped SB in bin 1 (would normally pass a BST and enhanced BST);[16] – Positive control in BSTs OECD 301, 306, 310;[4,8,9]
Triethanolamine (TEA); rapidly biodegradable – non persistent	– ECHA database: Readily biodegradable;[17] – Recommended by regulators for testing in ring test; – Variable degradation observed in BSTs ranging from 0-100%: o Eide-Haugmo et al. (2012) found TEA to degrade 20% in 28 days in OECD 306 Closed Bottle test;[18] o Unpublished results vary from under 20% to over 60% biodegradation after 28 days for OECD 306 Closed Bottle test (Cefas, personal communication, 2016); o Gerike and Fisher (1979) found TEA to degrade 91-100% in 28 days in Sturm test, 97% in 42 days in AFNOR test, 96% in 19 days in precursor to OECD 301E test, 0-2% in 14 days in MITI test and 0-9% in 30 days in Closed Bottle test;[19]
4-nitrophenol (4NP); inherently biodegradable – non persistent	– ECHA database: Inherently biodegradable;[20] – Comber and Holt (2010) grouped 4NP in bin 2 (would normally fail a current BST, but pass an enhanced BST);[16] – Previously tested during intra-laboratory activated sludge and marine BST validation;[1,21] – Variable degradation observed in BSTs ranging from 0-100%: o Nyholm and Kristensen (1987) found 4NP to degrade in OECD 306 Closed Bottle tests 38% in 28 days and 0-64% in 60 days; 4NP degraded in OECD 306 Shake Flask tests 35-54% in 28 days and 0-100% in 60 days (results from OECD 306 ring test 1984-85);[22,23] o Ott et al. (2019) found 4NP to degrade 3-91% in 60 days in marine OECD 301B tests with varying cell concentrations;[1] o Martin et al. (2017) found 4NP to degrade 84-91% in 60 days in activated sludge OECD 301B tests with varying cell concentrations;[21] o Gerike and Fisher (1979) found 4NP to degrade 90-98% in 28 days in Sturm test, 97% in 42 days in AFNOR test, 100% in 19 days in precursor to OECD 301E test, 1-3% in 14 days in MITI test and 0-60% in Closed Bottle test;[19]
Anionic polyacrylamide (APAM); no reference biodegradation classification assigned	– No information available in ECHA database as polymers are exempt from REACH;[24] – Recommended by industry for testing in ring test: polyacrylamides (PAMs) are widely used in several industrial fields such as for water treatment, agriculture and oil recovery;[25] – Previous research found PAM macromolecules resistant to microbial attack, requiring initial physical-chemical break-down;[26,27] – Unpublished biodegradability data shows no degradation for OECD 306 Closed Bottle test, marine BODIS test or Zahn Wellens test (SNF, personal communication, 2018); – Variable degradation reported in unpublished imBST_{MR}-similar industry study with 100-fold increased bacterial cell concentrations from seawater measuring O_2 consumption with MRs and 400 mg/L APAM (Equinor, personal communication, 2016): o Study 1, April: over 20% biodegradation measured in 120 days; o Study 2, November: no biodegradation detected in 90 days; – Due to a lack of peer-reviewed reference literature for APAM, it was not possible to assign a "correct" biodegradation classification; consequently, APAM results in the ring test were discussed separately to data of SB, TEA, 4NP and PCP;
Pentachlorophenol (PCP); potentially persistent	– Not registered under REACH [28], but the Finish Environment Institute (SYKE) database indicates potential persistence based on BST results;[29] – Comber and Holt (2010) grouped PCP in bin 3 (should normally fail a BST and enhanced BST);[16] – Previously tested during intra-laboratory activated sludge and marine BST validation;[1,21] – Variable degradation observed in different biodegradation test, depending on PCP concentration and adaptation: o Ott et al. (2019) found radiolabeled PCP at 10 mg/L to not degrade (0-1%) in 60 days in marine OECD 301B tests with varying cell concentrations;[1] o Martin et al. (2017) found radiolabeled PCP at 10 mg/L to not degrade (0-1%) in 60 days in activated sludge OECD 301B tests with varying cell concentrations;[21] o Lapertot and Pulgarin (2006) found PCP to not degrade (0%) in 28 days in inherent test OECD 302B, but concluded that this may have been the result of substrate inhibition;[30] o Ingerslev et al. (1998) observed PCP degradation in shake flask simulation tests in unadapted systems only after long acclimation phases (14-85 days in river water tests), but PCP degradation rates increased in adapted systems; no or little degradation was observed at inhibitory PCP concentrations above 20 mg/L, but PCP degraded quickly (t_{50} = 3-10 days) at concentrations under 2.5 mg/L;[31] – Toxicity [31,32] and low solubility concerns; however, PCP was most suitable negative control after screening 34 potential compounds proposed from regulators and recommendations from previous report;[16,33]

Table S5. Chemical and test strategy**.** Overview of the test setups and chemicals tested at each anonymised CRO, labelled CRO A-M. The total number of each test method, per chemical, is included in the last row of the table.

CRO	OECD306$_{CB}$						mBST$_{MR}$						imBST$_{MR}$					
	B	SB	TEA	4NP	APAM	PCP	B	SB	TEA	4NP	APAM	PCP	B	SB	TEA	4NP	APAM	PCP
A							X	X	X			X	X	X	X			X
B	X	X	X	X	X	X												
C	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
D							X	X	X	X	X	X	X	X	X	X	X	X
E	X	X	X	X	X	X												
F	X	X	X	X	X	X	X	X		X		X	X	X		X		X
G	X	X	X	X	X	X												
H	X	X	X	X		X	X	X		X		X	X	X		X		X
I							X	X	X	X	X	X	X	X	X	X	X	X
J	X	X	X	X	X													
K	X	X	X	X	X	X	X	X		X	X	X	X	X		X	X	X
L	X	X	X	X	X		X		X			X	X	X	X			X
M							X	X	X	X		X	X	X	X	X		X
Total:	9	9	9	9	8	7	9	8	6	7	4	9	9	9	6	7	4	9

B: blank. SB: sodium benzoate. TEA: triethanolamine. 4NP: 4-nitrophenol. APAM: anionic polyacrylamide. PCP: pentachlorophenol.

191 **Table S6.** Oxygen available in the $OECD306_{CB}$ and closed system MR systems.

Assumptions	
– "At 15°C and 20°C and 32 parts per thousand salinity (ocean water), the solubility of dissolved oxygen is about 8.1 and 7.4 mg/l, respectively." [9] – $OECD306_{CB}$: fill volume 300 mL, no headspace, incubation temperature 20°C; – $mBST_{MR}$ and $imBST_{MR}$: fill volume 250 mL, headspace 260 mL, incubation temperature 20°C; – For the $imBST_{MR}$ and $mBST_{MR}$, calculations are only relevant for closed MR systems (OxiTop), as the other MR systems (CES respirometer and Micro-Oxymax) replenish oxygen immediately after consumption; – Molecular mass O_2: 32 g/mol; 21% O_2 in air; ideal gas at 20°C, 1 atm: 24.04 L/ mol;	
$OECD306_{CB}$:	**$mBST_{MR}$ and $imBST_{MR}$**
O_2 in liquid phase: 0.3 L x 7.4 mg O_2/L = 2.22 mg O_2 O_2 in headspace: / **Total O_2 in $OECD306_{CB}$ bottle:** 2.22 mg O_2 + 0 mg O_2 = **2.22 mg O_2**	O_2 in liquid phase: 0.25 L x 7.4 mg O_2/L = 1.85 mg O_2 O_2 in headspace: Volume O_2 in headspace: 0.26 L x 0.21 = 0.055 L O_2; n(O_2) = 0.055 L O_2 ÷ 24.04 L/mol = 2.29 x 10^{-3} mol O_2 m(O2) = 32 g/mol x 2.29 x 10^{-3} mol O_2 = 7.33 x 10^{-2} g = 73.28 mg O_2 **Total O_2 in $imBST_{MR}$ or $mBST_{MR}$ bottle:** 1.85 mg O_2 + 73.28 mg O_2 = **75.13 mg O_2**
75.13 mg O_2 ÷ 2.22 mg O_2 = 33.84 In this study, MR test setups provide at least 34-times more O_2 than the $OECD306_{CB}$ test setup.	

192

Table S7. Raw and processed seawater characterization. CRO A and I used seawater collected and processed from the same source. All analysis except TCC performed by CROs (methods see Table S2). Temperature measurement S1 does not always represent original seawater temperature.

Description		CRO A	CRO B	CRO C	CRO D	CRO E	CRO F	CRO G	CRO H	CRO I	CRO J	CRO K	CRO L	CRO M
Seawater collection	Collection date $OECD306_{CB}$	—	01.06.17	09.03.17	—	30.05.17	01.05.17	23.05.17	06.04.17	See CRO A	02.06.17	24.04.17	14.03.17	—
	Collection date MR tests	27.03.17	—	07.03.17	08.05.17	—	01.05.17	—	04.04.17			24.04.17	14.03.17	14.08.17
	Depth (m)	6	3	nr.	10	2	10	50	nr.		10	nr.	60	0.5
	Distance offshore (m)	40-50	45	67	300	100	250	5000	nr.		100	nr.	nr.	200
	Water appearance	Clear	Clear	Clear	Clear	Clear	Slightly turbid	Clear	Clear		Clear	Clear	Clear	Clear
	Date setup $OECD306_{CB}$	—	14.06.17	15.03.17	—	06.06.17	11.05.17	30.05.17	13.04.17	—	09.06.17	04.05.17	21.03.17	—
	Date setup MR tests	31.03.17	—	08.03.17	13.05.17	—	04.05.17	—	06.04.17	31.03.17	—	26.04.17	15.03.17	17.08.17
Raw seawater (S1)	pH	8.0	7.8	8	8.1	7.9	7.40	8	7.70	See CRO A	8.2	7.8	8	7.9
	T (°C)	10.4	24.9	18.7	17.8	19.2	9.0	22.0	15.6		12.0	10.4	14.9	2.8
	DO (mg/L)	10.3	6.0	8	9.5	9.19	7.9	7.4	7.85		9.6	11.1	7.9	12.
	Conductivity (mS/cm)	24.0	44.1	45.3	44.3	46.7	45.0	np.	48.10		43.8	45.8	53.3	42.7
	Salinity (ppt)	16.1	28.7	32.2	27.5	34.7	28.0	34.1	30.60		31.1	29.6	34.6	27.5
	HPC x 10^3/ mL	82	10	np.	0.92	0.48	0.5	2	4.5		np.	2	Not countable	np.
	TCC x 10^5/ mL	5.4 ± 0.4	—	2 ± 0.094	2.8 ± 0.21	—	3.1 ± 0.49	—	0.6 ± 0.034		—	0.7 ± 0.07	1.1 ± 0.04	7.5 ± 0.21
10 µm filtered seawater for $mBST_{MR}$ (S2)	pH	8.7		8					7.80	See CRO A		6.76	8.	8
	T (°C)	19.1		18.7					16.00			19.7	10.7	0.9
	DO (mg/L)	8.8		8	np.		np.		7.72			9.3	8.3	13.4
	Conductivity (mS/cm)	24.5	—	45.3		—		—	48.10		—	48.4	53.5	41.1
	Salinity (ppt)	16.7		32.2					30.60			30.3	34.4	26.3
	HPC x 10^3/ mL	Not countable		np.	0.39		0.2		2.1			0.65	4	np.
	TCC x 10^5/ mL	4.8 ± 0.36		2.4 ± 0.37	1.6 ± 0.12		1.2 ± 0.0094		0.4 ± 0.02			0.86 ± 0.12	1.87± 0.12	5.4 ± 0.15
TFF processed seawater for	pH	8.8		7.9					7.80	See CRO A		7.1	8	7.6
	T (°C)	19.0	—	18.9	np.	—	np.	—	16.10		—	19.8	12.6	6.0
	DO (mg/L)	8.5		8.1					7.23			9	8.5	11.4

Description		CRO A	CRO B	CRO C	CRO D	CRO E	CRO F	CRO G	CRO H	CRO I	CRO J	CRO K	CRO L	CRO M
$imBST_{MR}$ (S3)	Conductivity (mS/cm)	24.4		45.7					48.00			48.1	53.6	42.6
	Salinity (ppt)	16.6		33.2					30.60			31.3	34.6	27.4
	HPC x 10^4/ mL	140		np.	0.19		0.37		49			Not countable	20	np.
	TCC x 10^7/ mL	7.6 ± 0.14		0.37 ± 0.041	2.4 ± 0.096		1.3 ± 0.035		0.71 ± 0.0036			0.16 ± 0.0054	0.26 ± 0.013	12 ± 0.99
Aged seawater for $OECD306_{CB}$ (S4)	pH		8.00	8		7.9	7.3	8	8.2		8.2	8.3	7.8	
	T (°C)		19.80	19.6		20.0	20.3	19.0	19.7		18.6	21.2	19.7	
	DO (mg/L)		7.40	7.5		9.0	7.7	6.4	7.6		7.8	9	7.6	
	Conductivity (mS/cm)	—	44.1	49.0	—	46.1	44.6	np.	48.8	—	43.6	44.6	52.5	—
	Salinity (ppt)		28.50	34.7		33.1	31.0	34.1	31.6		31.4	31.5	34.4	
	HPC x 10^4/ mL		10	np.		6.8	0.012	0.06	0.3		np.	0.33	10	

—: test setup not conducted. HPC: heterotrophic plate counts. nr: not recorded. np: not performed. T: temperature. TCC: total cell counts.

Table S8. Effect of pretreatment on bacteria concentrations in $OECD306_{CB}$ and $imBST_{MR}$. Coloring indicates fold cell increase (green) and fold cell reduction (red) between treatment steps. CRO A and I used the same seawater.

Test	Fold change	CRO A/I	CRO B	CRO C	CRO D	CRO E	CRO F	CRO G	CRO H	CRO J	CRO K	CRO L	CRO M
$OECD306_{CB}$	S1→S4	—	94	np.	—	141.7	0.2	0.3	0.7	np.	1.7	25	—
$imBST_{MR}$	S1→S3	140.4	—	18.8	88	—	42	—	118.9	—	23.3	23.7	160.2
	S2→S3	156.2	—	14.8	148	—	103	—	180.6	—	19.1	14.1	221.8

—: test setup not conducted. S1: raw seawater. S2: 10 µm filtered seawater. S3: 10 µm filtered and TFF treated seawater to increase bacteria concentrations 100-fold nominally. S4: seawater after OECD 306 pretreatment (filtered/sedimented and aged). np: analysis not performed.

Table S9. Chemical degradation of reference compounds in the three test systems in respect to CROs as evaluated against two regulatory persistence thresholds. Cursive brackets state the number of CROs out of all CROs where the reference compound degraded in at least 2/3 replicates to pass the stated persistence criteria and classify as non-persistent.

	Current test: $OECD306_{CB}$				**Revised test: $mBST_{MR}$**				**New test: $imBST_{MR}$**			
	Not persistent under OSPAR [a]		Not persistent under REACH [b]		Not persistent under OSPAR [a]		Not persistent under REACH [b]		Not persistent under OSPAR [a]		Not persistent under REACH [b]	
SB	**100%**	*(9/9)*	**100%**	*(7/7)*	**100%**	*(8/8)*	**100%**	*(8/8)*	**100%**	*(9/9)*	**100%**	*(9/9)*
TEA	**0%**	*(0/9)*	**14%**	*(1/7)*	**0%**	*(0/6)*	**17%**	*(1/6)*	**33%**	*(2/6)*	**50%**	*(3/6)*
4NP	**11%**	*(1/9)*	**0%**	*(0/7)*	**0%**	*(0/7)*	**0%**	*(0/7)*	**0%**	*(0/7)*	**14%**	*(1/7)*
APAM	**25%**	*(2/8)*	**0%**	*(0/7)*	**0%**	*(0/4)*	**0%**	*(0/4)*	**0%**	*(0/3)*	**0%**	*(0/3)*
PCP	**33%**	*(2/6)*	**0%**	*(0/4)*	**0%**	*(0/8)*	**0%**	*(0/8)*	**0%**	*(0/8)*	**0%**	*(0/8)*

[a] OSPAR: Biodegradation ≥ 20% over 28 days = non-persistent; biodegradation < 20% over 28 days = persistent [34]

[b] REACH: Biodegradation ≥ 60% over 60 days = non-persistent; biodegradation < 60% over 60 days = potentially persistent [35]

Table S10. Overview of sodium benzoate (SB) degradation in the three test systems based on replicates. The mean biodegradation values recorded on day 28, 60 and 120 are stated. Lag phase (t_L), time to reach 50% degradation (t_{50}) and dt_{50} (t_{50}-t_L) were only determined for the $mBST_{MR}$ and $imBST_{MR}$ tests. Cursive values state the number of SB replicates out of all performed SB replicates, which were used to calculate the respective benchmark criteria.

	OECD306$_{CB}$		**mBST$_{MR}$**		**imBST$_{MR}$**	
	Mean ± SD	*R*	Mean ± SD	*R*	Mean ± SD	*R*
Day 28	73 ± 15 %	*27/27*	73 ± 14 %	*22/22*	77 ± 9 %	*26/26*
Day 60	82 ± 15 %	*21/21*	77 ± 15 %	*22/22*	80 ± 9 %	*26/26*
Day 120	ND		76 ± 20 %	*22/22*	81 ± 16 %	*26/26*
t_L	ND		4 ± 3 d	*22/22*	2 ± 1 d	*26/26*
t_{50}	ND		7 ± 4 d	*22/22*	4 ± 2 d	*26/26*
dt_{50}	ND		3 ± 3 d	*22/22*	2 ± 1 d	*26/26*

ND: not defined. R: replicate numbers. SD: standard deviation.

Table S11. Overview of triethanolamine (TEA) degradation in the three test systems in respect to replicates. The mean biodegradation values recorded on day 28, 60 and 120 are stated. Lag phase (t_L), time to reach 50% degradation (t_{50}) and dt_{50} (t_{50}-t_L) were only determined for the $mBST_{MR}$ and $imBST_{MR}$ tests. Cursive values state the number of TEA replicates out of all performed TEA replicates, which were used to calculate the respective benchmark criteria.

	OECD306$_{CB}$		**mBST$_{MR}$**		**imBST$_{MR}$**	
	Mean ± SD	*R*	Mean ± SD	*R*	Mean ± SD	*R*
Day 28	6 ± 7 %	*27/27*	4 ± 6 %	*18/18*	20 ± 24 %	*18/18*
Day 60	28 ± 33 %	*20/20*	24 ± 25 %	*18/18*	51 ± 28 %	*18/18*
Day 120	ND		43 ± 31 %	*18/18*	61 ± 24 %	*18/18*
t_L	ND		42 ± 19 d	*14/18*	32 ± 20 d	*17/18*
t_{50}	ND		82 ± 30 d	*7/18*	50 ± 26 d	*16/18*
dt_{50}	ND		30 ± 21 d	*7/18*	21 ± 17 d	*16/18*

ND: not defined. R: replicate numbers. SD: standard deviation.

Table S12. Overview of 4-nitrophenol (4NP) degradation in the three test systems in respect to replicates. The mean biodegradation values recorded on day 28, 60 and 120 are stated. Lag phase (t_L), time to reach 50% degradation (t_{50}) and dt_{50} (t_{50}-t_L) were only determined for the $mBST_{MR}$ and $imBST_{MR}$ tests. Cursive values state the number of 4NP replicates out of all performed 4NP replicates, which were used to calculate the respective benchmark criteria.

	OECD306$_{CB}$		**mBST$_{MR}$**		**imBST$_{MR}$**	
	Mean ± SD	*R*	Mean ± SD	*R*	Mean ± SD	*R*
Day 28	3 ± 4 %	*27/27*	0 ± 1 %	*20/20*	6 ± 18 %	*20/20*
Day 60	8 ± 12 %	*21/21*	4 ± 13 %	*20/20*	21 ± 30 %	*20/20*
Day 120	ND		5 ± 13%	*20/20*	38 ± 36 %	*20/20*
t_L	ND		73 ± 38 d	*3/20*	53 ± 25 d	*11/20*
t_{50}	ND		39 d	*1/20*	56 ± 23 d	*10/20*
dt_{50}	ND		3 d	*1/20*	6 ± 3 d	*10/20*

ND: not defined. R: replicate numbers. SD: standard deviation.

Table S13. Overview of anionic polyacrylamide (APAM) degradation in the three test systems in respect to replicates. The mean biodegradation values recorded on day 28, 60 and 120 are stated. Lag phase (t_L), time to reach 50% degradation (t_{50}) and dt_{50} (t_{50}-t_L) were only determined for the $mBST_{MR}$ and $imBST_{MR}$ tests. Cursive values state the number of APAM replicates out of all performed APAM replicates, which were used to calculate the respective benchmark criteria.

	OECD306$_{CB}$		**mBST$_{MR}$**		**imBST$_{MR}$**	
	Mean ± SD	*R*	Mean ± SD	*R*	Mean ± SD	*R*
Day 28	9 ± 13 %	*24/24*	0 ± 0 %	*12/12*	3 ± 4 %	*10/10*
Day 60	10 ± 11 %	*21/21*	0 ± 1 %	*12/12*	6 ± 6 %	*10/10*
Day 120	ND		2 ± 2 %	*12/12*	8 ± 8 %	*10/10*
t_L	ND		ND	*0/12*	62 ± 30 d	*5/10*
t_{50}	ND		ND	*0/12*	ND	*0/10*
dt_{50}	ND		ND	*0/12*	ND	*0/10*

ND: not defined. R: replicate numbers. SD: standard deviation.

Table S14. Overview of pentachlorophenol (PCP) degradation in the three test systems in respect to replicates. The mean biodegradation values recorded on day 28, 60 and 120 are stated. Lag phase (t_L), time to reach 50% degradation (t_{50}) and dt_{50} (t_{50}-t_L) were only determined for the $mBST_{MR}$ and $imBST_{MR}$ tests. Cursive values state the number of PCP replicates out of all performed PCP replicates, which were used to calculate the respective benchmark criteria.

	$OECD306_{CB}$		**$mBST_{MR}$**		**$imBST_{MR}$**	
	Mean ± SD	*R*	Mean ± SD	*R*	Mean ± SD	*R*
Day 28	1 ± 2 %	*18/18*	0 ± 0 %	*24/24*	1 ± 4 %	*24/24*
Day 60	13 ± 18 %	*12/12*	0 ± 0 %	*24/24*	3 ± 8 %	*24/24*
Day 120	ND		0 ± 0 %	*24/24*	6 ± 14 %	*24/24*
t_L	ND		ND	*0/24*	35 ± 29 d	*6/24*
t_{50}	ND		ND	*0/24*	ND	*0/24*
dt_{50}	ND		ND	*0/24*	ND	*0/24*

ND: not defined. R: replicate number. SD: standard deviation.

Table S15. Chemical degradation of reference compounds in the three test systems in respect to replicates as evaluated against two regulatory persistence thresholds. Cursive brackets state the number of replicates out of all replicates where the reference compound degraded to pass the stated persistence criteria and classify as non-persistent.

	Current test: $OECD306_{CB}$				**Revised test: $mBST_{MR}$**				**New test: $imBST_{MR}$**			
	Not persistent under OSPAR [a]		Not persistent under REACH [b]		Not persistent under OSPAR [a]		Not persistent under REACH [b]		Not persistent under OSPAR [a]		Not persistent under REACH [b]	
SB	**100%**	*27/27*	**100%**	*21/21*	**100%**	*22/22*	**95%**	*21/22*	**100%**	*26/26*	**100%**	*26/26*
TEA	**4%**	*1/26*	**11%**	*2/19*	**0%**	*0/18*	**11%**	*2/18*	**33%**	*6/18*	**50%**	*9/18*
4NP	**7%**	*2/27*	**0%**	*0/21*	**0%**	*0/21*	**0%**	*0/21*	**10%**	*2/20*	**15%**	*3/20*
APAM	**25%**	*6/24*	**0%**	*0/21*	**0%**	*0/12*	**0%**	*0/12*	**0%**	*0/10*	**0%**	*0/10*
PCP	**39%**	*7/18*	**0%**	*0/12*	**0%**	*0/24*	**0%**	*0/24*	**4%**	*1/24*	**0%**	*0/24*

[a] OSPAR: Biodegradation ≥ 20% over 28 days = non-persistent; biodegradation < 20% over 28 days = persistent [34]

[b] REACH: Biodegradation ≥ 60% over 60 days = non-persistent; biodegradation < 60% over 60 days = potentially persistent [35]

Table S16. Test variation per chemical across tests described by the coefficient of variation.

	Current test: $OECD306_{CB}$	Revised test: $mBST_{MR}$	New test: $imBST_{MR}$
SB	5%	11%	9%
TEA	55%	51%	25%
4NP	75%	69%	50%
APAM	57%	57%	36%
PCP	52%	21%	56%
Mean	49%	42%	35%
Mean excl. PCP	48%	47%	30%

REFERENCES.

(1) Ott, A.; Martin, T. J.; Snape, J. R.; Davenport, R. J. Increased Cell Numbers Improve Marine Biodegradation Tests for Persistence Assessment. *Sci. Total Environ.* **2020**, *706*, 135621. https://doi.org/10.1016/j.scitotenv.2019.135621.

(2) Martin, T. J.; Goodhead, A. K.; Snape, J. R.; Davenport, R. J. Improving the Ecological Relevance of Aquatic Bacterial Communities in Biodegradability Screening Assessments. *Sci. Total Environ.* **2018**, *627*, 1552–1559. https://doi.org/10.1016/j.scitotenv.2018.01.264.

(3) Ott, A.; Martin, T. J.; Whale, G. F.; Snape, J. R.; Rowles, B.; Galay-Burgos, M.; Davenport, R. J. Improving the Biodegradability in Seawater Test (OECD 306). *Sci. Total Environ.* **2019**, *666*, 399–404. https://doi.org/10.1016/j.scitotenv.2019.02.167.

(4) OECD. *OECD 310 - Ready Biodegradability - CO2 in Sealed Vessels (Headspace Test)*; Paris, 2014. https://doi.org/10.1787/9789264016316-en.

(5) Weytjens, D.; Van Ginneken, I.; Painter, H. The Recovery of Carbon Dioxide in the Sturm Test for Ready Biodegradability. *Chemosphere* **1994**, *28* (4), 801–812. https://doi.org/10.1016/0045-6535(94)90232-1.

(6) Fiebig, S.; Maischak, H.; Maeß, C. *Investigation of Alternative Test Designs for Testing the Biodegradability in Seawater - Poster Presented at SETAC Europe 27th Annual Meeting, Brussels, Belgium*; 2017.

(7) Xie, W. H.; Shiu, W. Y.; Mackay, D. A Review of the Effect of Salts on the Solubility of Organic Compounds in Seawater. *Mar. Environ. Res.* **1997**, *44* (4), 429–444. https://doi.org/10.1016/S0141-1136(97)00017-2.

(8) OECD. *OECD Guideline for Testing of Chemicals (301): Ready Biodegradability*; Paris, 1992.

(9) OECD. *OECD Guideline for Testing of Chemicals (306): Biodegradability in Seawater*; Paris, 1992. https://doi.org/http://dx.doi.org/10.1787/9789264070486-en.

(10) Painter, H. A. Detailed Review Paper on Biodegradability Testing. *OECD Environ. Monogr.* **1995**, *98*, 1–162.

(11) Courtes, R.; Bahlaoui, a; Rambaud, a; Deschamps, F.; Sunde, E.; Dutrieux, E. Ready Biodegradability Test in Seawater: A New Methodological Approach. *Ecotoxicology and environmental safety*. 1995, pp 142–148. https://doi.org/10.1006/eesa.1995.1054.

(12) Syracuse Research Corporation. Physprop http://esc.syrres.com/fatepointer/search.asp (accessed Jul 23, 2019).

(13) Royal Society of Chemistry. ChemSpider https://www.chemspider.com/Default.aspx (accessed Jul 23, 2019).

(14) Ingre-Khans, E.; Ågerstrand, M.; Beronius, A.; Rudén, C. Reliability and Relevance Evaluations of REACH Data. *Toxicol. Res. (Camb).* **2019**, *8* (1), 46–56. https://doi.org/10.1039/c8tx00216a.

(15) ECHA. Brief profile sodium benzoate https://echa.europa.eu/brief-profile/-/briefprofile/100.007.760 (accessed Nov 5, 2019).

(16) Comber, M.; Holt, M. *Developing a Set of Reference Chemicals for Use in Biodegradability Tests for Assessing the Persistency of Chemicals*; 2010.

(17) ECHA. Brief profile triethanolamine https://echa.europa.eu/fr/brief-profile/-/briefprofile/100.002.773 (accessed Nov 5, 2019).

(18) Eide-Haugmo, I.; Brakstad, O. G.; Hoff, K. A.; da Silva, E. F.; Svendsen, H. F. Marine Biodegradability and Ecotoxicity of Solvents for CO2-Capture of Natural Gas. *Int. J. Greenh. Gas Control* **2012**, *9*, 184–192. https://doi.org/10.1016/j.ijggc.2012.03.006.

(19) Gerike, P.; Fischer, W. K. A Correlation Study of Biodegradability Determinations with Various Chemicals in Various Tests. *Ecotoxicol. Environ. Saf.* **1979**, *3* (1), 159–173. https://doi.org/10.1016/0147-6513(81)90045-2.
(20) ECHA. Brief profile 4-nitrophenol https://echa.europa.eu/fr/brief-profile/-/briefprofile/100.002.556 (accessed Nov 5, 2019).
(21) Martin, T. J.; Snape, J. R.; Bartram, A.; Robson, A.; Acharya, K.; Davenport, R. J. Environmentally Relevant Inoculum Concentrations Improve the Reliability of Persistent Assessments in Biodegradation Screening Tests. *Environ. Sci. Technol.* **2017**, *51* (5), 3065–3073. https://doi.org/10.1021/acs.est.6b05717.
(22) Nyholm, N.; Kristensen, P. *Screening Test Methods for Assessment of Biodegradability of Chemical Substances in Sea Water: Ring Test Programme 1984- 1985. Final Report*; Hørsholm, Denmark, 1987.
(23) Nyholm, N.; Kristensen, P. Screening Methods for Assessment of Biodegradability of Chemicals in Seawater - Results from a Ring Test. *Ecotoxicol. Environ. Saf.* **1992**, *23* (2), 161–172. https://doi.org/10.1016/0147-6513(92)90056-9.
(24) ECHA. Guidance for Monomers and Polymers - Guidance for the Implementation of REACH. **2012**, No. April, 1–25.
(25) Caulfield, M. J.; Qiao, G. G.; Solomon, D. H. Some Aspects of the Properties and Degradation of Polyacrylamides. *Chem. Rev.* **2002**, *102* (9), 3067–3084. https://doi.org/10.1021/cr010439p.
(26) Seybold, C. A. Polyacrylamide Review: Soil Conditioning and Environmental Fate. *Commun. Soil Sci. Plant Anal.* **1994**, *25* (11–12), 2171–2185.
(27) El-Mamouni, R.; Frigon, J. C.; Hawari, J.; Marroni, D.; Guiot, S. R. Combining Photolysis and Bioprocesses for Mineralization of High Molecular Weight Polyacrylamides. *Biodegradation* **2002**, *13* (4), 221–227. https://doi.org/10.1023/A:1021272519714.
(28) ECHA. Substance information pentachlorophenol https://echa.europa.eu/fr/substance-information/-/substanceinfo/100.001.617 (accessed Nov 5, 2019).
(29) SYKE. Data bank of environmental chemicals - Pentachlorophenol http://wwwp.ymparisto.fi/scripts/Kemrek/Kemrek_uk.asp?Method=MAKECHEMdetailsform&txtChemId=245 (accessed Nov 5, 2019).
(30) Lapertot, M. E.; Pulgarin, C. Biodegradability Assessment of Several Priority Hazardous Substances: Choice, Application and Relevance Regarding Toxicity and Bacterial Activity. *Chemosphere* **2006**, *65* (4), 682–690. https://doi.org/10.1016/j.chemosphere.2006.01.046.
(31) Ingerslev, F.; Baun, A.; Nyholm, N. Aquatic Biodegradation Behavior of Pentachlorophenol Assessed through a Battery of Shake Flask Die-Away Tests. *Environ. Toxicol. Chem.* **1998**, *17* (9), 1712–1719. https://doi.org/10.1897/1551-5028(1998)017<1712:ABBOPA>2.3.CO;2.
(32) King, E. F.; Painter, H. A. The Effect of Acclimatisation on the Toxicity of Chemicals to Activated Sludge Microorganisms. *Environ. Pollut.* **1985**, *39* (3), 267–280. https://doi.org/10.1016/0143-1471(85)90101-1.
(33) ECETOC. *Refined Approaches for Risk Assessment of PBT/VPvB Chemicals Technical Report No.112*; Brussels, 2011.
(34) OSPAR. *Guidelines for Completing the Harmonised Offshore Chemical Notification Format (HOCNF) (OSPAR Agreement: 2012/05. Update 2015)*; 2015.
(35) Regulation (EC) No 1907. *Of the European Parliament and of the Council of 18 December 2006 Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals*

352 *(REACH), Establishing a European Chemicals Agency, Amending Directive 1999/4*; 2006.
353