Organic materials repurposing, a data set for theoretical predictions of new applications for existing compounds

Ömer H. Omar1, Tahereh Nematiaram1, Alessandro Troisi1,2 ✉ & Daniele Padula2 ✉

We present a data set of 48182 organic semiconductors, constituted of molecules that were prepared with a documented synthetic pathway and are stable in solid state. We based our search on the Cambridge Structural Database, from which we selected semiconductors with a computational funnel procedure. For each entry we provide a set of electronic properties relevant for organic materials research, and the electronic wavefunction for further calculations and/or analyses. This data set has low bias because it was not built from a set of materials designed for organic electronics, and thus it provides an excellent starting point in the search of new applications for known materials, with a great potential for novel physical insight. The data set contains molecules used as benchmarks in many fields of organic materials research, allowing to test the reliability of computational screenings for the desired application, “rediscovering” well-known molecules. This is demonstrated by a series of different applications in the field of organic materials, confirming the potential for the repurposing of known organic molecules.

Background & Summary

High Throughput Virtual Screenings (HTVSs)1,2 have recently been exploited to a great extent to identify promising materials in the domain of organic electronics. This powerful technique has often been used in combination with domain knowledge of the problem, carrying out screenings of modifications of known motifs or architectures known to work for a specific problem e.g. functionalisation for dye-sensitized solar cells3, donor-acceptor motifs for thermally activated delayed fluorescence (TADF)4, singlet fission (SF)5, and for general photovoltaic architectures6. This strategy translates in computational terms the process of experimental discovery exploiting chemical intuition7, and allows the reduction of the chemical space to explore8. However, the findings are bound to fall within the domain of what is already known and prevent the discovery of new motifs and design rules. Studies based on exploiting domain knowledge like biradical character for SF9,10 or donor-acceptor motifs for TADF11,12 will not find new design rules. Generative models also tend to find motifs similar to those already known13. Additionally, the identified candidates may not be easy to synthesise in the laboratory or be stable enough to be characterised, despite recent progresses in introducing measures of synthetic accessibility in HTVSs14.

In this study, we aim at providing a starting point for computational searches overcoming the mentioned limitations by presenting a data set of 48182 organic semiconductors (OSCs) constituted of molecules that were prepared with a documented synthetic pathway, and are stable in solid-state, enabling their crystallographic characterisation. The data set is therefore an excellent starting point to identify OSCs for various applications that can guide experimental research. We based our search on the Cambridge Structural Database (CSD)16, from which we selected OSCs with a computational strategy described in the following sections. The CSD dates back to the 60–70s, and contains crystal structure data for > 1M samples prepared for various purposes. Excluding polymorphs17–19 or samples measured in different experimental conditions20, the vast majority of molecules in the data set has characterisation data available. As it was not built with organic materials applications in mind though, of course, it does contain entries related to this field, any data set derived from the CSD21 is, therefore,

1University of Liverpool, Department of Chemistry, Liverpool, L69 7ZD, UK. 2Università di Siena, Dipartimento di Biotecnologie, Chimica e Farmacia, Siena, 53100, Italy. ✉e-mail: a.troisi@liverpool.ac.uk; daniele.padula@unisi.it
unbiased with respect to the application, though some bias is present due to choices of research groups in the study of a certain molecule or experimental contraints with respect to the ability to crystallise the sample and characterise it. This low bias provides a great potential for novel physical insight: setting different criteria for the ideal candidates based on experimental benchmarks, the more stringent ones (i.e. more rare) can be used to translate results into design principles. Additionally, the fact that it contains molecules used for benchmarks in many fields of organic materials research allows testing the reliability of computational screenings for the desired application, “rediscovering” well known molecules.

Studies of OSCs for technological applications exploit the analyses of various electronic properties, ranging from frontier orbital energies to excited state energies and oscillator strengths. For instance, early searches of materials for organic photovoltaics exploited HOMO and LUMO energies, high performance non-fullerene acceptors are known to possess a low LUMO-LUMO + 1 gap, luminescent materials for new generation organic light emitting diodes (OLEDs) based on TADF as well as singlet fission candidates have been identified by calculating the S1-T1 gap ($\Delta E_{ST}$), and high mobility semiconductors were discovered by looking at electronic couplings, reorganisation energies and electron-phonon couplings. Providing wavefunctions and basic excited state properties for the first few states will enable other researchers to carry out systematic investigations for applications that, to the best of our knowledge, are yet to be explored through computational screenings, such as aggregation induced emitters (AIEngens), but also for extremely innovative applications based on higher excited states, for which chemical intuition is still limited, e.g. designing anti-Kasha fluorophores, even displaying delayed fluorescence.

The data set presented in this work thus contains a collection of simulated spectroscopic properties on the X-ray geometries of existing organic molecules, showing a simulated HOMO-LUMO gap ($E_{bg}$) falling below 4 eV, which we therefore define as organic semiconductors, and can be searched for relevant properties in various technological applications. Some data sets offer interesting properties for OSCs relevant for specific applications, e.g. the HOPV for organic photovoltaics, but they are limited to boundaries within the chemical space, i.e. they exploit domain knowledge about what is known to work. Other data sets offer spectroscopic properties of molecules, such as e.g. the QM8 or the OE62 data sets, but the former is limited in the number and type of heavy atoms and excited states considered, while the latter provides spectroscopic data only for a small fraction of the data set ($\approx 5K$ entries). The data set we present in this work is thus aimed at complementing the currently available ones in these aspects, which we describe in more detail in the following sections.

**Methods**

The data set of OSCs we present here has been built starting from the python application programming interface provided within the CSD distribution. To identify OSCs, we started by removing polymeric molecules, disordered solids, and co-crystals from the entries containing X-ray structures. We further reduced the structures to be retained by:

1. including only the most commonly used elements in typical OSCs in the selection of molecules (H, B, C, N, O, F, Si, P, S, Cl, As, Se, Br, I);
2. removing entries with more than one molecule type in the unit cell;
3. removing duplicate entries

X-ray geometries include all heavy atoms, while hydrogen atoms are added and normalised (i.e. placed at a typical X-H distance using statistical surveys of neutron diffraction data) using the CSD library’s built-in functions exploiting such literature data. Due to errors in the procedure, e.g. missing hydrogens in diborane molecules, the structurally erroneous entries are filtered out by comparing the heavy atom connectivity layer of InChI strings of the CSD entry and the extracted geometry, followed by comparison of the chemical formulae between the CSD entry and the extracted geometry. The data set is up to date with the 2020 version of the CSD, thus updates starting from the 2021 version are possible.

This procedure resulted in a reduction of the data set from $\approx 1M$ to $\approx 265K$ molecules. To identify OSCs we adopted a three-step computational funnel strategy in combination with a calibration procedure, aimed at estimating the HOMO-LUMO gap ($E_{bg}$) with quantum mechanical (QM) methods of reasonable accuracy. First of all, we selected three methods of increasing accuracy for our computational funnel: PM7, B3LYP/3 – 21 G*, and B3LYP/6–31 G*. Second, we picked a subset of 550 molecules on which we performed single point calculations including only the most commonly used elements in typical OSCs (H, B, C, N, O, F, Si, P, S, Cl, As, Se, Br, I) on the X-ray geometries provided within the CSD, obtaining orbital energies with all three methods. This allowed us to compute calibration curves to estimate the B3LYP/6–31 G* HOMO-LUMO gap ($E_{bg}$) from low accuracy ones (see panels b), c), e), and f) in Fig. 1), and the associated error distribution. With calibration curves available, we proceeded to compute HOMO-LUMO gaps ($E_{bg}$) for the entire data set of $\approx 265K$ molecules (panel d) in Fig. 1), estimating the gap that we would obtain if we ran a higher level calculation. Considering the distribution of errors of the calibration curve, at the PM7 level we retained any molecule showing $E_{bg} \leq 4$ eV as OSC, reducing the data set from $\approx 265K$ to $\approx 200K$ molecules. On these molecules, we recomputed the gap at the B3LYP/3–21 G* level (panel g) in Fig. 1), considering any molecule showing $E_{bg} \leq 4$ eV as OSC, resulting in the $\approx 50K$ molecules that constitute the data set presented here. 4 eV is a conventional upper limit for semiconductors, and all the best performing molecules across various applications have a smaller gap. On these molecules, we computed excited states properties at TD-DFT/M06-2X/def2-SVP (see Fig. 2), releasing, as part of the data set, the converged ground state wavefunction, and the results for the first three singlet ($S_1$-$S_3$) and triplet ($T_1$-$T_3$) states. A calibration of the TD-DFT method for $S_1$ and $T_1$ excitation energies for $\approx 100$ data points with available experimental data is presented elsewhere.
and guarantees the reliability of the method (RMSE $\approx 0.05$ eV). All QM calculations were carried out with the Gaussian16 software\(^43\), and the data provided as part of this release were extracted from output and checkpoint files using the Multiwfn software\(^44\) and the CClib python library\(^45\).

These calculations allow for interesting analyses regarding the time evolution of the CSD. For instance, since the deposition date of each entry is known, it is possible to follow how many OSCs were deposited over time, both in absolute and fractional terms. From these analyses (see Fig. 3) we see that, while the absolute number is naturally increasing over time, the fractional number of OSCs within the CSD is constant until $\approx 2010$, and since then it has basically doubled, rising from $\approx 3–4\%$ to $\approx 7\%$, which is in agreement with the evolution of research in the organic materials field.

**Data Records**

The curated data set is available from DataCat, the University of Liverpool repository\(^46\):

1. data extracted from QM calculations are provided at the University of Liverpool repository\(^46\) in comma-separated values (\(\text{.csv}\)) format, which can be easily read through common programs or programming languages. A description of the provided data is given in Table 1;
2. the wavefunctions for each entry are provided in a set of 31 sequential archives at the University of Liverpool repository allowing for sequential or partial download. Geometries are also given to facilitate analyses. Data are made available in .wfn format; Gaussian16 QM calculations output files are provided at the University of Liverpool repository to allow for additional wavefunction analyses, with the aim to characterise electronic states or transitions, as mentioned in the following sections.

Geometries and wavefunctions are provided in .wfn format, the AIM traditional format. We chose this format to provide interested users data for analyses or subsequent calculations that would be independent of the software we used. In fact, .wfn files can be generated or processed with a multitude of tools, among which the popular software Multiwfn, the python library IOData, ORCA, and others. Each .wfn file contains the molecular geometry, as well as occupied molecular orbitals expressed in the atomic basis and their energies. These data can be used for visualisation of e.g. geometries, occupied orbitals, but also to run QM calculations with an initial guess to obtain refined properties for applications of interest.

Gaussian16 output files are provided to allow for additional wavefunction analyses on electronic excitations, allowing interested users to avoid repeating calculations that we have already performed.

**Technical Validation**

The key idea is that new applications of existing molecules can be discovered by searching for useful properties computed for a large data set, thanks to a robust calibration between predicted and experimentally validated data. Crucially, the data set should be totally unbiased and not related to the property of interest: this way, discoveries are truly unexpected and have a large applicative and commercial value. We proved this concept through a range of demonstrations in recent works, covering various applications areas. These demonstrations considered an outdated data set consisting of ≈40K OSCs. The data set presented here is up to date with the 2020

![Graph](https://via.placeholder.com/150)

**Fig. 3** Time evolution of OSCs within the CSD. Left: total number of OSCs deposited each year. Right: fraction of OSCs deposited each year.

| Column name | Unit | Method | Description |
|-------------|------|--------|-------------|
| ID          | —    |        | unique CSD identifier |
| doi         | —    |        | doi of the experimental paper characterising the X-ray structure |
| formula     | —    |        | chemical formula |
| NAts        | —    |        | number of heavy atoms |
| SMILES      | —    |        | the SMILES string |
| HOMO        | eV   | TD-DFT/M06-2X/def2-SVP | computed HOMO energy on the X-ray geometry |
| LUMO        | eV   | TD-DFT/M06-2X/def2-SVP | computed LUMO energy on the X-ray geometry |
| E(S1)       | eV   | TD-DFT/M06-2X/def2-SVP | computed S1 energy on the X-ray geometry |
| f(S1)       | —    | TD-DFT/M06-2X/def2-SVP | computed S1 oscillator strength on the X-ray geometry |
| E(S2)       | eV   | TD-DFT/M06-2X/def2-SVP | computed S2 energy on the X-ray geometry |
| f(S2)       | —    | TD-DFT/M06-2X/def2-SVP | computed S2 oscillator strength on the X-ray geometry |
| E(S3)       | eV   | TD-DFT/M06-2X/def2-SVP | computed S3 energy on the X-ray geometry |
| f(S3)       | —    | TD-DFT/M06-2X/def2-SVP | computed S3 oscillator strength on the X-ray geometry |
| E(T1)       | eV   | TD-DFT/M06-2X/def2-SVP | computed T1 energy on the X-ray geometry |
| E(T2)       | eV   | TD-DFT/M06-2X/def2-SVP | computed T2 energy on the X-ray geometry |
| E(T3)       | eV   | TD-DFT/M06-2X/def2-SVP | computed T3 energy on the X-ray geometry |

**Table 1.** Description of metadata and electronic properties gathered in the database.
version of the CSD, thus containing entries that were not the objects of our previous studies; the same strategies can be used on the fraction of molecules not previously considered to discover more potential candidates, in line with our previous findings.

The key applications demonstrated in our previous works are the following:

1. we showed that it is possible to identify completely new molecules that undergo singlet fission (a property of relevance for solar cells) by calibrating a computational method to yield accurate energies of singlet and triplet excited states and found molecules with the ideal energy level alignment. The method rediscovered known molecules for singlet fission (true positives), and identified several different families of known compounds with this desirable property;

2. we proposed a related screening protocol to identify molecules undergoing TADF, a relevant property in the area of display technologies. The protocol indicated without any adjustable parameter that 0.3% of the molecules considered may undergo TADF. About half of them were known TADF emitters, providing great confidence in the quality of the prediction. The other half of the hits were totally unknown to the field, illustrating in parallel how this approach can lead to completely novel design rules;

3. we showed that a similar approach can be used to identify novel electron acceptors to be used in organic solar cells to replace expensive and inefficient fullerene derivatives. Also in this case, about half of the “discovered” molecules were known, the other half being totally novel ones. This work showed that database searching is only the first step and it is possible to modify lead compounds to have other desirable properties, like solubility;

4. we showed that we can screen for luminescent crystals displaying superradiance or near IR emission, properties of interest in the areas of light-emitting diodes, organic lasers, and biological imaging. A common theme of all applications particularly well exemplified by this one is the ability of large screenings to identify plausible optima for any properties; in this case, what is the maximum red shift that can be observed when a particular molecule is studied in its crystal.

The basis of similar studies can be laid by analysing properties provided in this database similarly to what is shown in Fig. 4. In the left panel, we report T1 vs S1 energies. Potential singlet fission materials fall in the proximity of the dashed black line, representing S1 = 2 T1. Similarly, potential TADF materials fall in the proximity of the dashed blue line, representing S1 = T1. Colours encode the S1 oscillator strength (fS1) through a logarithmic scale, assuming bright states are desirable.

Right: S2 vs S1 energies. Potential anti-Kasha materials fall in the proximity of the dashed black line, representing S2 = 2 S1. Colours encode the S2 oscillator strength (fS2) through a logarithmic scale, assuming bright states are desirable.
Usage Notes
Above, we have listed some applications deriving from the data presented here. In general, the starting point for each of those applications consisted of a calibration of the computational method used to carry out further analyses with available experimental data. Thanks to the fact that we provide the ground state wavefunction for each of our entries, not only will these calibrations be faster because we provide an initial guess for QM calculations, but also many more analyses are accessible. For instance, electronic states or transitions can be thoroughly characterised with packages such as Multiwfn[32] or TheoDORF[56], which can provide detailed information regarding the nature of an electronic transition (e.g. Charge Transfer metrics[57,58], *ghost states*[59], electronic density difference[60], exciton delocalisation[61,62] etc). Additionally, this data set can form the basis for training sets to Machine Learning models aiming at reproducing the electronic density of molecules[63,64], based on experimental X-ray geometries. The availability of CSD identifiers enables the expansion of analyses to molecules in their crystals[65], which is fundamental for technological applications of organic semiconductors. Finally, the synthetic approaches that make molecules within the CSD accessible can be easily tracked down thanks to references provided within the data set. This allows not only for a prompt source of synthetic routes to be exploited in case of experimental validation of the results, but is also useful in combination with retrosynthetic planning strategies[66,67].

Code availability
Scripts to obtain plots starting from the database are available at the University of Liverpool repository[68].

Received: 8 October 2021; Accepted: 21 December 2021;
Published online: 14 February 2022

References
1. Pyzer-Knapp, E. O., Suh, C., Gómez-Bombarelli, R., Aguilera-Iparraguirre, J. & Aspuru-Guzik, A. What is high-throughput virtual screening? a perspective from organic materials discovery. *Annu. Rev. Mater. Res.*, **45**, 195–216, https://doi.org/10.1146/annurev-матеrics-070214-020823 (2015).
2. Omar, O. H., del Cueto, M., Nematiaram, T. & Troisi, A. High-throughput virtual screening for organic electronics: A comparative study of alternative strategies. *J. Mater. Chem. C*, **9**, 13357–13343, https://doi.org/10.1039/d1tc03256a (2021).
3. Ormø, K. B., Pedersen, C.-S., Garcia-Lastra, J. M. & Thygesen, K. S. Optimizing porphyrins for dye sensitized solar cells using large-scale ab initio calculations. *Phys. Chem. Chem. Phys.*, **16**, 16246–16254, https://doi.org/10.1039/C4CP01289E (2014).
4. Shu, Y. & Levine, B. G. Simulated evolution of fluorophores for light emitting diodes. *J. Phys. Chem. Phys.*, **57**, 104104, https://doi.org/10.1021/ACS.chemphys.5b01784 (2015).
5. Blaskovits, J. T., Fumanal, M., Vela, S. & Corminboeuf, C. Designing singlet fission candidates from donor–acceptor copolymers. *Chemistry of Materials*, **32**, 6515–6524, https://doi.org/10.1021/acs.chemmater.0c02882 (2020).
6. Jørgensen, P. B. Concurrent optimisation of organic donor-acceptor pairs through machine learning. *Adv. Energy Mater.*, **9**, 1902463, https://doi.org/10.1002/aenm.201902463 (2019).
7. Padula, D., Simpson, J. D. & Troisi, A. Combining electronic and structural features in machine learning models to predict organic solar cells properties. *Mater. Horiz.*, **6**, 343–349, https://doi.org/10.1039/C8MH01135D (2019).
8. Padula, D., Stuke, A. & Aspuru-Guzik, A. Inverse molecular design using machine learning:Generative models for matter engineering. *Science*, **361**, 360–365, https://doi.org/10.1126/science.aat2668 (2018).
9. von Lilienfeld, O. A. Quantum machine learning in chemical compound space. *Angew. Chem. Int. Ed.*, **57**, 4164–4169, https://doi.org/10.1002/anie.201709686 (2018).
10. Minami, T. & Nakano, M. Diradical character view of singlet fission. *J. Phys. Chem. Lett.*, **3**, 145–150, https://doi.org/10.1021/acs.jpclett.2c00097 (2011).
11. Omar, O. H., Padula, D. & Troisi, A. Elucidating the relationship between multiradical character and predicted singlet fission activity. *ChemPhotoChem* **4**, 5223–5229, https://doi.org/10.1002/cptc.202000098 (2020).
12. Tanaka, H., Shizu, K., Nakanotani, H. & Adachi, C. Twisted intramolecular charge transfer state for long-wavelength thermally activated delayed fluorescence. *Chem. Mater.*, **25**, 3766–3771, https://doi.org/10.1021/acs.chemmater.3c02428a (2013).
13. Zhang, Y. et al. Supramolecular structure-dependent thermally-activated delayed fluorescence (TADF) properties of organic polymorphs. *J. Phys. Chem. C*, **120**, 19759–19767, https://doi.org/10.1021/acs.jpcc.6b05537 (2016).
14. Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning:Generative models for matter engineering. *Science*, **361**, 360–365, https://doi.org/10.1126/science.aat2668 (2018).
15. Wen, Y., Fu, L., Li, G., Ma, J. & Ma, H. Accelerated Discovery of Potential Organic Dyes for Dye-Sensitized Solar Cells by Interpretable Machine Learning Models and Virtual Screening. *Solar RRL*, **4**, https://doi.org/10.1002/solr.202000110 (2020).
16. Groom, C. R., Bruno, J. I., Lightfoot, M. P. & Ward, S. C. The Cambridge structural database. *Acta Cryst.*, **B72**, 171–179, https://doi.org/10.1107/S2052520616003954 (2016).
17. Landi, A., Troisi, A. & Peluso, A. Explaining different experimental hole mobilities: influence of polymorphism on dynamic disorder in pentacene. *J. Mater. Chem. C*, **7**, 9665–9670, https://doi.org/10.1039/c9tc03174a (2019).
18. Matthews, C. C., de Wij, G. A. de Groot, R. A. & Palstra, T. T. M. Modeling the polymorphism of pentacene. *J. Am. Chem. Soc.*, **125**, 6323–6330, https://doi.org/10.1021/ja011449n (2003).
19. Matthews, C. C. et al. Polymorphism in Pentacene. *Acta Cryst. Sect. C*, **57**, 939–941, https://doi.org/10.1107/S010827010100703X (2001).
20. Siegrist, T. et al. A polymorph lost and found: The high-temperature crystal structure of pentacene. *Adv. Mater.*, **19**, 2079–2082, https://doi.org/10.1002/adma.200602072 (2007).
21. Stuke, A. et al. Atomic structures and orbital energies of 61,489 crystal-forming organic molecules. *Scientific Data*, **7**, https://doi.org/10.1038/s41597-020-0385-y (2020).
22. Pyzer-Knapp, E. O., Li, K. & Aspuru-Guzik, A. Learning from the harvard clean energy project: The use of neural networks to accelerate materials discovery. *Adv. Funct. Mater.*, **25**, 6495–6502, https://doi.org/10.1002/adfm.201501919 (2015).
23. Hachmann, J. et al. Lead candidates for high-performance organic photovoltaics from high-throughput quantum chemistry – the harvard clean energy project. *Energy Environ. Sci.*, **7**, 698–704, https://doi.org/10.1039/C3ee42756k (2014).
24. Hachmann, J. et al. The harvard clean energy project: Large-scale computational screening and design of organic photovoltaics on the world community grid. *J. Phys. Chem. Lett.*, **2**, 2241–2251, https://doi.org/10.1021/acs.jpclett.1b00866 (2011).
25. Kanal, I. Y., Owens, S. G., Bechtel, J. S. & Hutchison, G. R. Efficient computational screening of organic polymer photovoltaics. *J. Phys. Chem. Lett.*, **4**, 1613–1623, https://doi.org/10.1021/acs.jpclett.1j00215 (2013).
66. Genheden, S. et al. AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning. J. Cheminformatics 12, https://doi.org/10.1186/s13321-020-00472-1 (2020).
67. Weininger, D. SMILES, a chemical language and information system. 1. introduction to methodology and encoding rules. J. Chem. Inf. Comput. Sci 28, 31–36, https://doi.org/10.1021/ci00057a005 (1988).
68. Weininger, D., Weininger, A. & Weininger, J. L. SMILES, 2. algorithm for generation of unique SMILES notation. J. Chem. Inf. Comput. Sci 29, 97–101, https://doi.org/10.1021/ci00062a008 (1989).
69. Weininger, D. SMILES, 3. DEPICT, graphical depiction of chemical structures. J. Chem. Inf. Comput. Sci 30, 237–243, https://doi.org/10.1021/ci00067a005 (1990).

Acknowledgements
The authors gratefully acknowledge the Cambridge Crystallographic Data Centre for their support in the development of this research. D.P. acknowledges support from the Italian Ministry of Education, University, and Research (MIUR) through a Rita Levi Montalcini grant, and hpc@dbcf (Università di Siena) for computational resources. A.T. acknowledges support by EPSRC (U.K.).

Author contributions
A.T. and D.P. planned the project. Ö.H.O., T.N. and D.P. carried out calculations and analyses. All authors discussed the results and reviewed the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to A.T. or D.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access
This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

The Creative Commons Public Domain Dedication waiver http://creativecommons.org/publicdomain/zero/1.0/ applies to the metadata files associated with this article.

© The Author(s) 2022