New tyrosinases with a broad spectrum of action against contaminants of emerging concern: Insights from *in silico* analyses

Marcus V. X. Senra[†] & Ana Lúcia Fonseca

Instituto de Recursos Naturais, Universidade Federal de Itajubá, 37500-903, Itajubá, Minas Gerais – Brazil

[†]Corresponding author – msenra@unifei.edu.br; Orcid - 0000-0002-3866-8837

Abstract

Tyrosinases (EC 1.14.18.1) are type-3 copper metalloenzymes with strong oxidative capacities and low allosteric selectivity to phenolic and non-phenolic aromatic compounds that have been used as biosensors and biocatalysts to mitigate the impacts of environmental contaminants over aquatic ecosystems. However, the widespread use of these polyphenol oxidases is limited by elevated production costs and restricted knowledge on their spectrum of action. Here, six tyrosinase homologs were identified and characterized from the genomes of 4 widespread freshwater ciliates using bioinformatics. Binding energies between 3D models of these homologs and ~1000 contaminants of emerging concern (CECs), including fine chemicals, pharmaceuticals, personal care products, illicit drugs, natural toxins, and pesticides were estimated through virtual screening, suggesting their spectrum of action and potential uses in environmental biotechnology might be considerably broader than previously thought. Moreover, considering that many ciliates, including those carrying tyrosinase genes within their genomes are fast-growing unicellular microeukaryotes that can be efficiently culturable at large-scales under *in vitro* conditions, should be regarded as potential low-cost sources for the production of relevant biotechnological molecules.

Keywords. Genomics, bioremediation, biocatalysts, phenoloxidases, emerging contaminants, virtual screening

1. Introduction

Contamination of aquatic environments with toxic compounds has become a major global threat and one of our greatest challenges toward biodiversity and hydric resources conservation (Galindo-Miranda et al., 2019). Contaminants of emerging concern (CECs) are pervasive and unregulated pollutants of natural and synthetic origin, which include fine

chemicals, dyes, pharmaceuticals, cosmetics, personal care products, cyanotoxins, and pesticides among many others (Sauvé et al., 2014; Visanji et al., 2020). CECs usually reach aquatic compartments via domestic and industrial sewage or agriculture runoffs (Lapworth et al., 2012), where they are detected within nano (ng.L^{-1}) or micro ($\mu\text{g.L}^{-1}$) scales (Wilkinson et al., 2017) and frequently associated with a variety of short and long-term detrimental effects over biodiversity, community structure and human health (Brühl and Zaller, 2019; Lapworth et al., 2012; Saaristo et al., 2018; Sanchez and Egea, 2018; Santangeli et al., 2019). Because of the general inefficiency of conventional water treatment systems in the removal of such pollutants (Bolong et al., 2009), the development of novel efficient and cost-effective technologies able to detect and/or remove the great diversity CECs from natural systems is greatly desirable (Bilal et al., 2019).

Tyrosinases (EC 1.14.18.1) are widespread polyphenol oxidases of bacteria, fungi, plants, and animals (Zaidi et al., 2014) that participate of critical biological functions, such as photoprotection (mammals), immune system and wound-healing processes (plants, fungi, and insects), and in spore and sexual organs maturation (fungi) (Agarwal et al., 2019). These metalloenzymes carry a characteristic type-3 copper center formed of two copper ions (Cu^{A} and Cu^{B}), each one connected to three conserved histidine residues (Goldfeder et al., 2014), which catalyzes hydroxylation monophenols to diphenols (monophenolase reaction) and oxidation diphenols to quinones (diphenolase reaction), using molecular oxygen as electron acceptor in both reactions (Kaintz et al., 2014). This process occurs as a catalytic cycle, where the active site shifts from active *oxy* ($\text{Cu}^{\text{II}} - \text{O}_2 - \text{Cu}^{\text{II}}$) state to intermediate *met* ($\text{Cu}^{\text{II}} - \text{OH} - \text{Cu}^{\text{II}}$) state, and then, to a resting *deoxy* ($\text{Cu}^{\text{I}} - \text{Cu}^{\text{I}}$) state that requires a new dioxygen for reactivation (Ba and Vinoth Kumar, 2017).

This biocatalyst have been largely used in food, cosmetic and pharmaceutical industries (Agarwal et al., 2019) and more recently, because of its low allosteric specificity and strong oxidative capacities over a diversity of phenolic and nonphenolic aromatic compounds (Asgher et al., 2014), emerged as a versatile biosensor and biocatalyst for monitoring and removing environmental contaminants, such as cresols, chlorophenols, phenylacetates, and bisphenols from natural systems (Ba and Vinoth Kumar, 2017; Harms et al., 2011). However, despite this great potential to mitigate the environmental impacts of CECs, most studies remain restricted to commercially available enzymes, purified from bacteria or fungi and to tests against a narrow diversity of pollutants. Moreover, their high production cost limits the widespread use of tyrosinases in field applications (Ba and Vinoth Kumar, 2017).

Bioinformatic tools represent a powerful, fast and cost-effective method for enzyme discovery (Vanacek et al., 2018). As a consequence of diffusion high-throughput sequencing technologies thousands of genomes from model and non-model organisms from all over the tree of life are now available within public sequence databases, which in turn have become a fruitful source of biological information. Despite of extensively used in drug discovery processes, including for detection and rational design of novel molecules and prediction of their potential cellular and sub-cellular targets (Xia, 2017), genomics tools (protein 3D structure modeling, molecular dynamics (MD), and docking assays) have only sparsely used in environmental biotechnology applications for identification and characterization of novel biocatalysts and prediction of potential targets. In an interesting work, Srinivasan and coworkers (Srinivasan et al., 2019) applied a molecular docking approach to indicate a potential a broad spectrum of action of laccases (EC 1.10.3.2) (polyphenol oxidases) against a diversity of textile dyes, and in a different study, Ahlawat and collaborators (Ahlawat et al., 2019), corroborated this data using MD and docking assays, but also validated their results through experimental procedures, highlighting potential applications of bioinformatics to speed up and reduce costs related to development of novel biocatalysts

Here, we screened the genomes of 23 ciliate species (Ciliophora) for the presence of tyrosinase homologs. Ciliates are ancient (~1180 Ma) (Fernandes and Schrago, 2019), highly diversified (> 8,000 valid species), and widespread unicellular microeukaryotes able to colonize almost every aquatic and terrestrial ecosystem (Lynn, 2008), including those pristine and highly impacted ones (Foissner and Berger, 1996), indicating adaptations to tolerate or even to degrade and metabolize environmental contaminant. Since many species are fast-growing and culturable under *in vitro* conditions and some of them are amenable to genetic modification approaches a great (yet underexplored) biotechnological potential is foreseen. In fact, six new tyrosinase homologs from four freshwater culturable ciliate species were identified and characterized using bioinformatics. Theoretical 3D models were generated, using comparative modeling, and the spectrum of action of these new tyrosinases were predicted against a set of ~1000 CECs structures through molecular docking (virtual screening) assays. Our data indicate that these tyrosinases of ciliates may target a wide variety of CECs, including antibiotics, licit and illicit drugs, pharmaceuticals, cosmetics, personal care products, cyanotoxins and pesticides, expanding the potential environmental applications of these versatile enzymes.

2. Materials and Methods

2.1. Genomic screenings and primary structure characterization of tyrosinase homologs

Draft genomes of 23 ciliate species were retrieved from GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) and screened for the presence of tyrosinase homologs. For each genome, open reading frames (orfs) (>300bp) and putative proteomes were inferred using *getorf* (Rice et al., 2000), applying the appropriate genetic codes, as suggested in analyses using *facil* (Dutilh et al., 2011). To identify tyrosinase homologs within these predicted proteomes two complementary approaches were used: pairwise sequence alignments, using *blastp* v2.6 (Camacho et al., 2009) against protein sequences from SwissProt (<https://www.uniprot.org/>), applying an e-value cutoff of 10^{-5} and query coverage filter of 80%; and profile searches, using *fuzzpro* (Rice et al., 2000) with conserved tyrosinase signatures from Prosite (Sigrist et al., 2012), and *Hmmer* v3 (Eddy, 2011) with Hidden Markov Models (HMMs) profiles of tyrosinases from PFAM (El-Gebali et al., 2019). Candidates were, subsequently, characterized using *InterProScan* v5 (Jones et al., 2014), *Phobius* (Käll et al., 2004) and *SignalP* v5 (Almagro Armenteros et al., 2019) for detection of conserved domains, presence of transmembrane terminals and secretion signals, respectively. Moreover, molecular weight (MW) and isoelectric points (pI) were estimated using *compute pI/Mw* (https://web.expasy.org/compute_pi/).

2.2. Phylogenetic analysis

The dataset used to infer phylogenetic relationships among the newly identified tyrosinases and previously characterized ones were generated using all amino acid sequences available from SwissProt (<https://www.uniprot.org/>) (consulted in Feb 2020), excluding fragment of sequences, and using the tyrosinase from *Rhizobium meliloti* (SwissProt:P33180) as outgroup. The dataset was, then, aligned using *MAFFT* v7 (Rozewicki et al., 2019), with default parameters, resulting a matrix of 39 sequences and 1207 characters. To determine the best model of sequence evolution within this dataset, *SMS* (Lefort et al., 2017) was used, and a maximum likelihood phylogenetic tree was inferred using *phym*l v7 (Guindon et al., 2010), applying the WAG +G +I model of sequence evolution, and branch support values was defined through the nonparametric bootstrap method of 1,000 pseudo-replicates.

2.3. Three-dimensional (3D) structure prediction of tyrosinase homologs

Comparative modeling, using *Modeller* v9.19 (Webb and Sali, 2016) were performed to predict the theoretical 3D structures of tyrosinase homologs of ciliates. Best templates for each homolog sequence were identified using *HHpred* (Zimmermann et al., 2017), which is defined

by high-quality pairwise amino acid sequence alignments between subjects and query proteins (with known 3D structures) deposited in the Protein Data Bank (PDB) (Berman et al., 2000). 100 models were generated for each homolog sequence and the one with the lowest DOPE (Discrete Optimized Protein Structure) score were selected for further analyses. Finally, the quality of the best theoretical models were evaluated using *MolProbity* (Chen et al., 2010). Visualization and manipulation of all 3D models were done using UCSF *Chimera* v1.14 (Pettersen et al., 2004).

2.4. Molecular dynamics (MD) simulations

To evaluate the stability of the generated 3D models of tyrosinases of ciliates, we performed MD simulations using *Gromacs* v5 (Abraham et al., 2015), applying OPLS-AA force field (Jorgensen et al., 1996). Models were centered into dodecahedral boxes with sizes of at least 2.0 nm, from the surface of the model to the side of the box. Systems were solvated using SPC/E water model (Mark and Nilsson, 2001) and neutralized with 0.15 mol/L sodium chloride. After a steepest *descent* minimization of 50,000 steps using position restraints, systems were equilibrated, sequentially, using NVT followed by NPT, considering temperature of 300 K and pressure of 1 bar, each ensemble for 100 ps. Then, models were released from their position restraints and the production steps ran for 50 ns using *leap-frog* as integrator. All-atom bond lengths were linked using *LINCS* (Hess et al., 1997). Electrostatic interactions and van der Waals were calculated by using particle-mesh Ewald (PME) method (Essmann et al., 1995) with cutoffs of 1.0 nm. For each atom, the list of neighbors was updated every 2 fs. The stability of the models during the run was accessed through root-mean-square deviation (RMSD) and radius gyrase plots, and secondary structures of the models were assigned by means of Dictionary of Protein Secondary Structures (DSSP), using the Gromacs tool, *do_dssp*.

2.5. Virtual screenings

Potential targets of the new tyrosinases of ciliates were predicted using a molecular docking (virtual screening) approach. For that, binding energies (-kcal/mol) between tyrosinase 3D models and a set of 76 CECs, including pharmaceuticals, illicit drugs, chemicals and natural toxins were estimated. Two known ligands of tyrosinases, L-tyrosine (CID6057) and o-dopaquinone (CID), and a well-studied tyrosinase isolated from the basidiomycete mushroom, *Agaricus bisporus* (PDB:2y9wA) were used to serve as a reference to determine likely and unlikely targets.

Before the assay, putative cofactors and pockets were predicted using *COACH* (Yang et al., 2013, 2012) and *DoGSiteScorer* (<https://proteins.plus/>), respectively. Then, correct protonation states of these new tyrosinases were inferred using *Protein Plus* (<https://proteins.plus/>). Three-dimensional CECs structures were retrieved from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and when unavailable, 2D structures (in SMILE format) were downloaded, instead, and converted to 3D using *Obabel* (O'Boyle et al., 2011). All ligands were desalted, polar hydrogens added, and clashes and torsions resolved by minimization (steepest descent), using *Obabel* (O'Boyle et al., 2011). Next, grids of 10 Å were placed into the active type-3 copper center of the *oxy*-tyrosinases using *ADT* (*Auto Dock Tools*) (Morris et al., 2009), and then, an in-house python script was used to run the virtual screenings using *AutoDock Vina* (Trott and Olson, 2010) and to compile the results from the highest binding energies values of the best poses for each tyrosinase 3D models and ligand structures.

To further evaluate the use of these new tyrosinases to target pesticides within the environment, ~1000 pesticide structures from PubChem were subjected to a new round of virtual screening assay, but now running the analysis within the online server DockThor (<https://dockthor.lncc.br/v2/>), with default settings and following the same framework for protein and ligand preparation as mentioned above.

3. Results and Discussion

3.1. Identification and primary structure characterization

The identification of tyrosinase homologs from the inferred proteome of 23 ciliate species was done using a combined approach based in pair-wise alignments against protein sequences from the manually curated protein database, Swiss-Prot; and profile-based searches using tyrosinase HMM-profiles from PFAM and tyrosinase signatures from PROSITE (as described in details in Method section). Using this framework, we identified six new tyrosinase homologs within the genomes of four freshwater stichotrichids, *Paraurostyla* sp. (PuTYR-1), *Stylonychia lemnae* (SITYR-1 and SITYR-2), *Uroleptopsis citrina* (UcTYR-1 and UcTYR-2) and *Urostyla* sp. (UTYR-1). These data are summarized in Table 1.

These tyrosinase homologs have a sequence identity of 39% with a tyrosinase characterized from the filamentous fungus *Aspergillus orizae* (SwissProt:Q00234) and carry conserved tyrosinase domains (InterProt domains IPR002227 and IPR008922) (Table 1). Their expected molecular weight (MW) and isoelectric point (pI) are within the range of typical tyrosinases described from literature (MW 14-75 KDa; pI 4.65 and 11.9) (Agarwal et al., 2016; Ba and Vinoth Kumar, 2017) and interestingly, all of them were predicted as extracellular

enzymes, but no clear secretory signals could be identified (Table 1). This character was previously reported in some fungal tyrosinases (Mayer, 2006) and a possible route for their secretion would be through quaternary associations with “caddy-like” proteins carrying their own signal peptides, such as those described in *Streptomyces* sp. (Matoba et al., 2018). The sequence similarity between tyrosinases of ciliates and fungi is further evidenced in our phylogenetic reconstruction (Figure 1), the ciliate’s homologs emerged in a single cluster together with the tyrosinase from *A. orizae* (SwissProt:Q00234) (Ciliophora/Fungi Clade I) that is a sister group of tyrosinases from higher fungi (Fungi Clade II), apart from tyrosinase clusters of *Streptomyces* (Bacteria Clade), insects (Metazoa Clade I), and mammals (Metazoa Clade II) (Figure 1).

3.2. Tertiary structure prediction

We used the X-ray diffraction-defined structure of a tyrosinase isolated from *Burkholderia thailandensis* (PDB:5ZRD) as template to predict, through comparative modeling, the theoretical 3D models of these tyrosinase homologs of ciliates (Table 2 and Figure 2). Their overall quality are rather satisfactory, >97% of amino acid residues are in allowed positions (Ramachandran plots) and low Molprobit index values, which reflects the crystallographic resolution at which each model is expected based in combined information derived from the clash score, percentage of unfavored residues within Ramachandran plot and bad side-chain rotamers (Table 2). These 3D models are pearl-shaped structures supported by a series of 8 α -helices secondary structures (Figure 2) with the characteristic active site, in central position, consisting of two copper ions (CuA and CuB), each one linked to three conserved histidine residues (Figure 3), within a pocket of mean volume of 411.43 Å³ and surface of 523.44 Å² (Table 2). Five out six conserved histidine residues are rigid and located within α -helical structures, while the sixth is flexible and located within a loop region (Figure 3). We could also detect within the active site of these enzymes: a conserved cysteine residue, typical of type-3 eukaryotic tyrosinases (Gielens et al., 1997), which interacts with the flexible histidine residue, through covalent thioester bonding, to favor copper ions incorporation into the active site (Fujieda et al., 2013); a conserved asparagine residue, possibly involved in hydrogen-bonding interactions during monophenol deprotonation reactions (Goldfeder et al., 2014); and 2 bulk phenylalanine residues close to CuA, which may be associated with monophenolase activity hindering (Mauracher et al., 2014) (Figure 3).

These new tyrosinases seems to be synthesized in a latent form, which requires post-translational modifications to become active. They all present a C-terminal region of ~120 aa that is absent from well-characterized mature fungi tyrosinases, such as PDB:4OUA and PDB:2Y9W (Table 2), which, when present, covers the active site of the enzyme, possibly a strategy to control enzymatic activity. Moreover, these homologs present a short N-terminal region (~4 aa), with no significant alignment with any known protein from PDB, possibly representing a 5' UTR. Both regions were. Both regions were trimmed off from the tyrosinases before modeling their 3D structures to avoid artifact problems in the subsequent docking analyses.

Before performing the molecular docking assay to predict the spectrum of action of these new tyrosinases against CECs, we subjected these theoretical models to molecular dynamics (MD) simulations of 50 ns to access their stability. Data were evaluated through root mean square deviation (RMSD) and radius of gyration plots (Figure 4), to reveal variations in the backbone and in the globularity of each one of these models, respectively, and by DSSP analyses (Suppl. Figure 1), to evidence fluctuations within their secondary structures during the MD simulation. Accordingly, the models showed, in general, low variations within backbone (Figure 4a), only slight changes into their globularity (Figure 4b), and their overall secondary structure organization remained considerably uniform (Suppl. Figure 1) along the MD run of 50 ns. The highest variations were observed for SITYR-1, which varied 4,50 nm in the RMSD and 2,17 nm in the globularity. However, such variations could be considered within acceptable ranges for the analyses of potential spectrum of action over CECs.

3.3. Spectrum of action prediction

The spectrum of action of these new tyrosinase homologs of ciliates might be much broader than previously thought (Figure 5, Table 3). To evaluate potential targets for these enzymes, we first estimated binding energies between the generated 3D models and a set of 76 CECs structures, including antibiotics, industrial chemicals, illicit drugs, pesticides, pharmaceuticals, and natural toxins. Our rationale was that ligands with high binding energies, above the positive controls (bottom part of Figure 5), may represent positive interactions. As another second reference, we included the well-studied fungus tyrosinase PDB:2Y9WA (left side of Figure 5) in the analysis, and its general behavior was greatly similar to the tyrosinases from ciliates. As shown in Figure 5, binding energies vary considerably within tyrosinases and within ligands, ranging from -3,8 kcal/mol, between acephate and UcTYR-1, to -9,0 kcal/mol between microcystin and SITYR-1, but in general, high binding energy values were obtained

for most of the tested classes, especially for antibiotics, cyanotoxins, drugs, and hormones, where the highest values were -7.6 kcal/mol between UTYR-1 and the antibiotic ofloxacin (CID 4583); -9.0 kcal/mol between SITYR-1 and the cyanotoxin Anatoxin-a (CID 3034748), -7.4 kcal/mol UTYR-1 and the pharmaceutical drugs Carbamazepine (CID 2554) and Diazepam (CID 3016); and -7.8 kcal/mol between UTYR-1 and the hormone estrone (CID 5870). On the other hand, the class pesticides, as whole, presented only low binding energy values, suggesting the group are unlikely targets for tyrosinases. However, given the environmental relevance of pesticides and to test whether the general low binding energies was specifically related with the selected pesticides (Figure 5), we decided to repeat the assay, now using ~1000 pesticide structures. The top 10 ligands with the highest binding energies are summarized in Table 3. In fact, high binding energies, way above the control, could be detected against a number of pesticides of different chemical groups, including widespread ones, such as the coumarinic rodenticide brodifacoum (CID 54680676); the organophosphorus herbicide falone (CID 7209), and the trifluoromethyl aminohydrazone insecticide hydramethylnon (CID 5281875) among many others (Table 3).

Interestingly, in both virtual screenings performed (Figure 5 and Table 3) binding energies varied considerably between tyrosinases and CECs, which could be explained by their high degree of polymorphism. In fact, pairwise sequence distances among PuTYR-1, SITYR-1, SITYR-2, UcTYR-1, UcTYR-2, and UTYR-1 are relatively high (0.46 in average), indicating numerous inter- and intraspecific polymorphisms (Suppl. Figure 2), which may have impact on their affinity to ligands. Contrasting binding energies were found, for example, for UTYR-1 (-7.5 kcal/mol) and UcTYR-2 (-5.5 kcal/mol) and Imipramine (CID 3696), an antidepressant drug that is associated with reduction in neuron numbers within the frontal lobe of fetal rats (Swerts et al., 2010); and for UTYR-1 (-7.6 kcal/mol) and PuTYR-1 (-5.8 kcal/mol) and Bisphenol B (CID 66166), a widely used plasticizer agent and known endocrine disruptor associated with different reproductive disorders in rats and fishes (Serra et al., 2019). Future studies aiming to map the influence of these polymorphisms in the binding energies against CECs will greatly contribute to the rational design of more efficient and ligand-specific tyrosinases.

Despite of our data indicate tyrosinases may have a broad spectrum of action against CECs, expanding the potential environmental applications of these versatile enzymes, all data generated require further experimental validation. Nevertheless, it is known from the literature that tyrosinases are effective for removal of the plasticizer agents and endocrine disruptors bisphenols A and B (Nicolucci et al., 2011), the methylphenol p-cresol (Bayramoglu et al.,

2013) and Acetaminophen (Ba et al., 2014) from water matrices, which were also tested here; therefore, their binding energies would serve as references to guide future *in vitro* works toward the most likely targets for tyrosinases.

According to Ba et al. (Ba and Vinoth Kumar, 2017), despite the great biotechnological potential foreseen for tyrosinases in remediation and monitoring processes, high costs and low yield of purified enzymes are the main limitations for the widespread use of these enzymes. The six tyrosinases were identified four *in vitro* culturable and widespread freshwater ciliates, *Paraurostyla* sp. (PuTYR-1), *Stylonychia lemnae* (SITYR-1 and SITYR-2), *Uroleptopsis citrina* (UcTYR-1 and UcTYR-2), and *Urostyla* sp. (UTYR-1). Similarly, many ciliate species can be efficiently cultured under *in vitro* conditions, they have small sizes and fast generation times, and some of them (especially those from genera *Paramecium* and *Tetrahymena*) have a wide array of well-established molecular tools available. Moreover, numerous molecules of industrial relevance have been identified from a wide variety of ciliate species (Elguero et al., 2019), including tyrosinases (this work), antimicrobials (Petrelli et al., 2012), antitumoral (Carpi et al., 2018), and bioactive molecules (Anesi et al., 2016). With that, we predict a great potential in ciliates to reduce production costs and to serve as sources of novel molecules of industrial and pharmaceutical interest, but also of environmental relevance.

4. Conclusion

Tyrosinases (EC1.14.18.1) are bifunctional copper monooxygenases of prokaryotes and eukaryotes that have a great environmental biotechnological potential to monitor and transform a variety of phenolic and non-phenolic aromatic compounds. Here, six new tyrosinases were identified from the genome of four freshwater stichtrichid ciliates, *Paraurostyla* sp. (PuTYR-1), *Stylonychia lemnae* (SITYR-1 and SITYR-2), *Uroleptopsis citrina* (UcTYR-1 and UcTYR-2) and *Urostyla* sp. (UTYR-1) present conserved domains and the characteristic active type-3 copper center, of well-characterized tyrosinases, representing the first report of tyrosinases in protists. Based on the measured binding energies in the virtual screening assays, these new enzymes may bind, or even be active against antibiotics, anti-oxidants, anti-inflammatory drugs, industrial chemicals, illicit drugs, hormones, pharmaceutical drugs, insect repellents, cyanotoxins, UV filters, and pesticides from different classes, indicating that the spectrum of action and biotechnological potential of this family of enzymes against contaminants of emerging concern might be considerably broader than previously expected. Many ciliates, including these caring tyrosinase genes, have short generation times, can be easily *in*

vitro cultivated, are amenable to genetic modifications and heterologous gene expression, and are known to produce a variety of secondary metabolites of biotechnological relevance. Therefore, we would like to stress the great biotechnological potential of ciliates as sources of novel molecules and to reduce costs of industrial-scale production of relevant enzymes. Although this is a pure bioinformatic work and requires proper biochemical data validation, it highlights the importance of a more intimate integration between bioinformatics and environmental biotechnology, which can greatly contribute to accelerate and reduce costs of new enzymes discovery/production steps for uses in remediation and monitoring processes of contaminants of emerging contaminants within natural environments.

Acknowledgement

The authors acknowledge the “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior” (CAPES) for the post-doctoral fellowship (PNPD) (88882.317976/2019-01) conferred to MS.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindah, E., Lindahl, E., 2015. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2, 19–25.
<https://doi.org/10.1016/j.softx.2015.06.001>
- Agarwal, P., Gupta, R., Agarwal, N., 2016. A Review on Enzymatic Treatment of Phenols in Wastewater. *J. Biotechnol. Biomater.* 06. <https://doi.org/10.4172/2155-952x.1000249>
- Agarwal, P., Singh, M., Singh, J., Singh, R.P., 2019. Microbial Tyrosinases: A Novel Enzyme, Structural Features, and Applications. *Appl. Microbiol. Bioeng.* 3–19.
<https://doi.org/10.1016/B978-0-12-815407-6.00001-0>
- Ahlawat, S., Singh, D., Viridi, J.S., Sharma, K.K., 2019. Molecular modeling and MD-simulation studies: Fast and reliable tool to study the role of low-redox bacterial laccases in the decolorization of various commercial dyes. *Environ. Pollut.* 253, 1056–1065.
<https://doi.org/10.1016/j.envpol.2019.07.083>

- Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O.,
Brunak, S., von Heijne, G., Nielsen, H., 2019. SignalP 5.0 improves signal peptide
predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423.
<https://doi.org/10.1038/s41587-019-0036-z>
- Anesi, A., Buonanno, F., Di Giuseppe, G., Ortenzi, C., Guella, G., 2016. Metabolites from
the Euryhaline Ciliate *Pseudokeronopsis erythrina*. *European J. Org. Chem.* 2016, 1330–
1336. <https://doi.org/10.1002/ejoc.201501424>
- Asgher, M., Shahid, M., Kamal, S., Iqbal, H.M.N., 2014. Recent trends and valorization of
immobilization strategies and ligninolytic enzymes by industrial biotechnology. *J. Mol.*
Catal. B Enzym. 101, 56–66. <https://doi.org/10.1016/j.molcatb.2013.12.016>
- Ba, S., Haroune, L., Cruz-Morató, C., Jacquet, C., Touahar, I.E., Bellenger, J.P., Legault,
C.Y., Jones, J.P., Cabana, H., 2014. Synthesis and characterization of combined cross-
linked laccase and tyrosinase aggregates transforming acetaminophen as a model
phenolic compound in wastewaters. *Sci. Total Environ.* 487, 748–755.
<https://doi.org/10.1016/j.scitotenv.2013.10.004>
- Ba, S., Vinoth Kumar, V., 2017. Recent developments in the use of tyrosinase and laccase in
environmental applications. *Crit. Rev. Biotechnol.* 37, 819–832.
<https://doi.org/10.1080/07388551.2016.1261081>
- Bayramoglu, G., Akbulut, A., Yakup Arica, M., 2013. Immobilization of tyrosinase on
modified diatom biosilica: Enzymatic removal of phenolic compounds from aqueous
solution. *J. Hazard. Mater.* 244–245, 528–536.
<https://doi.org/10.1016/j.jhazmat.2012.10.041>
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov,
I.N., Bourne, P.E., 2000. The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242.
<https://doi.org/10.1093/nar/28.1.235>
- Bilal, M., Adeel, M., Rasheed, T., Zhao, Y., Iqbal, H.M.N., Iqbal, M.N., 2019. Emerging
contaminants of high concern and their enzyme-assisted biodegradation – A review.
Environ. Int. 124, 336–353. <https://doi.org/10.1016/j.envint.2019.01.011>
- Bolong, N., Ismail, A.F., Salim, M.R., Matsuura, T., 2009. A review of the effects of
emerging contaminants in wastewater and options for their removal. *Desalination* 239,
229–246. <https://doi.org/10.1016/j.desal.2008.03.020>
- Brühl, C.A., Zaller, J.G., 2019. Biodiversity Decline as a Consequence of an Inappropriate
Environmental Risk Assessment of Pesticides. *Front. Environ. Sci.* 7, 177.
<https://doi.org/10.3389/fenvs.2019.00177>

- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. <https://doi.org/10.1186/1471-2105-10-421>
- Carpi, S., Polini, B., Poli, G., Barata, G.A., Fogli, S., Romanini, A., Tuccinardi, T., Guella, G., Paolo Frontini, F., Nieri, P., Di Giuseppe, G., 2018. Anticancer activity of euplotin C, isolated from the marine ciliate euplotes crassus, against human melanoma cells. Mar. Drugs 16. <https://doi.org/10.3390/md16050166>
- Chen, V.B.V.B., Arendall, W.B.B., Headd, J.J.J., Keedy, D.A.D.A., Immormino, R.M.R.M., Kapral, G.J.G.J., Murray, L.W.L.W., Richardson, J.S.J.S., Richardson, D.C.D.C., IUCr, 2010. MolProbity: All-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 12–21. <https://doi.org/10.1107/S0907444909042073>
- Dutilh, B.E., Jurgelenaite, R., Szklarczyk, R., van Hijum, S.A.F.T.F.T., Harhangi, H.R., Schmid, M., de Wild, B., François, K.J., Stunnenberg, H.G., Strous, M., Jetten, M.S.M.M., Op den Camp, H.J.M.M., Huynen, M.A., 2011. FACIL: Fast and accurate genetic code inference and logo. Bioinformatics 27, 1929–1933. <https://doi.org/10.1093/bioinformatics/btr316>
- Eddy, S.R., 2011. Accelerated Profile HMM Searches. PLoS Comput. Biol. 7, e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., Finn, R.D., 2019. The Pfam protein families database in 2019. Nucleic Acids Res. 47, D427–D432. <https://doi.org/10.1093/nar/gky995>
- Elguero, M.E., Nudel, C.B., Nusblat, A.D., 2019. Biotechnology in ciliates: an overview. Crit. Rev. Biotechnol. 39, 220–234. <https://doi.org/10.1080/07388551.2018.1530188>
- Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., 1995. A smooth particle mesh Ewald method. J. Chem. Phys. 103, 8577–8593. <https://doi.org/10.1063/1.470117>
- Fernandes, N.M., Schrago, C.G., 2019. A multigene timescale and diversification dynamics of Ciliophora evolution. Mol. Phylogenet. Evol. 139, 106521. <https://doi.org/10.1016/j.ympev.2019.106521>
- Foissner, W., Berger, H., 1996. A user-friendly guide to the ciliates (Protozoa, Ciliophora) commonly used by hydrobiologists as bioindicators in rivers, lakes, and waste waters, with notes on their ecology. Freshw. Biol. 35, 375–388. <https://doi.org/10.1111/j.1365->

2427.1996.tb01775.x

Fujieda, N., Yabuta, S., Ikeda, T., Oyama, T., Muraki, N., Kurisu, G., Itoh, S., 2013. Crystal structures of copper-depleted and copper-bound fungal pro-tyrosinase: Insights into endogenous cysteine-dependent copper incorporation. *J. Biol. Chem.* 288, 22128–22140. <https://doi.org/10.1074/jbc.M113.477612>

Galindo-Miranda, J.M., Guízar-González, C., Becerril-Bravo, E.J., Moeller-Chávez, G., León-Becerril, E., Vallejo-Rodríguez, R., 2019. Occurrence of emerging contaminants in environmental surface waters and their analytical methodology - A review. *Water Sci. Technol. Water Supply*. <https://doi.org/10.2166/ws.2019.087>

Gielens, C., De Geest, N., Xin, X.Q., Devreese, B., Van Beeumen, J., Préaux, G., 1997. Evidence for a cysteine-histidine thioether bridge in functional units of molluscan haemocyanins and location of the disulfide bridges in functional units d and g of the β (C)-haemocyanin of *Helix pomatia*. *Eur. J. Biochem.* 248, 879–888. <https://doi.org/10.1111/j.1432-1033.1997.00879.x>

Goldfeder, M., Kanteev, M., Isaschar-Ovdat, S., Adir, N., Fishman, A., 2014. Determination of tyrosinase substrate-binding modes reveals mechanistic differences between type-3 copper proteins. *Nat. Commun.* 5, 4505. <https://doi.org/10.1038/ncomms5505>

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. <https://doi.org/10.1093/sysbio/syq010>

Harms, H., Schlosser, D., Wick, L.Y., 2011. Untapped potential: Exploiting fungi in bioremediation of hazardous chemicals. *Nat. Rev. Microbiol.* 9, 177–192. <https://doi.org/10.1038/nrmicro2519>

Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M., 1997. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18, 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)

Jones, P., Binns, D., Chang, H.Y.H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y.S.Y., Lopez, R., Hunter, S., 2014. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>

Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J., 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic

liquids. *J. Am. Chem. Soc.* 118, 11225–11236. <https://doi.org/10.1021/ja9621760>

Kaintz, C., Mauracher, S.G., Rompel, A., 2014. Type-3 Copper Proteins, in: *Advances in Protein Chemistry and Structural Biology*. Elsevier Inc., pp. 1–35.

<https://doi.org/10.1016/bs.apcsb.2014.07.001>

Käll, L., Krogh, A., Sonnhammer, E.L., 2004. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* 338, 1027–1036.

<https://doi.org/10.1016/j.jmb.2004.03.016>

Lapworth, D.J., Baran, N., Stuart, M.E., Ward, R.S., 2012. Emerging organic contaminants in groundwater: A review of sources, fate and occurrence. *Environ. Pollut.* 163, 287–303.

<https://doi.org/10.1016/j.envpol.2011.12.034>

Lefort, V., Longueville, J.-E., Gascuel, O., 2017. SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* 34, 2422–2424. <https://doi.org/10.1093/molbev/msx149>

Lynn, D., 2008. *The Ciliated Protozoa*, 3rd ed. Pergamon Press, New York, NY.

Mark, P., Nilsson, L., 2001. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* 105, 9954–9960. <https://doi.org/10.1021/jp003020w>

Matoba, Y., Kihara, S., Bando, N., Yoshitsu, H., Sakaguchi, M., Kayama, K., Yanagisawa, S., Ogura, T., Sugiyama, M., 2018. Catalytic mechanism of the tyrosinase reaction toward the Tyr 98 residue in the caddie protein. *PLoS Biol.* 16, 1–22.

<https://doi.org/10.1371/journal.pbio.3000077>

Mauracher, S.G., Molitor, C., Al-Oweini, R., Kortz, U., Rompel, A., 2014. Latent and active abPPO4 mushroom tyrosinase cocrystallized with hexatungstotellurate(VI) in a single crystal. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70, 2301–2315.

<https://doi.org/10.1107/S1399004714013777>

Mayer, A.M., 2006. Polyphenol oxidases in plants and fungi: Going places? A review. *Phytochemistry*. <https://doi.org/10.1016/j.phytochem.2006.08.006>

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–91. <https://doi.org/10.1002/jcc.21256>

Nicolucci, C., Rossi, S., Menale, C., Godjevargova, T., Ivanov, Y., Bianco, M., Mita, L., Bencivenga, U., Mita, D.G., Diano, N., 2011. Biodegradation of bisphenols with immobilized laccase or tyrosinase on polyacrylonitrile beads. *Biodegradation* 22, 673–683. <https://doi.org/10.1007/s10532-010-9440-2>

O’Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., Hutchison, G.R., 2011. Open Babel: An open chemical toolbox. *J. Cheminform.* 3, 33.

<https://doi.org/10.1186/1758-2946-3-33>

Petrelli, D., Buonanno, F., Vitali, L.A., Ortenzi, C., 2012. Antimicrobial activity of the protozoan toxin climacostol and its derivatives. *Biologia (Bratisl)*. 67, 525–529.

<https://doi.org/10.2478/s11756-012-0030-0>

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera. A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. <https://doi.org/10.1002/jcc.20084>

Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276–7. [https://doi.org/10.1016/s0168-9525\(00\)02024-2](https://doi.org/10.1016/s0168-9525(00)02024-2)

Rozewicki, J., Li, S., Amada, K.M., Standley, D.M., Katoh, K., 2019. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* 47, W5–W10. <https://doi.org/10.1093/nar/gkz342>

Saaristo, M., Brodin, T., Balshine, S., Bertram, M.G., Brooks, B.W., Ehlman, S.M., McCallum, E.S., Sih, A., Sundin, J., Wong, B.B.M., Arnold, K.E., 2018. Direct and indirect effects of chemical contaminants on the behaviour, ecology and evolution of wildlife. *Proc. R. Soc. B Biol. Sci.* 285, 20181297. <https://doi.org/10.1098/rspb.2018.1297>

Sanchez, W., Egea, E., 2018. Health and environmental risks associated with emerging pollutants and novel green processes. *Environ. Sci. Pollut. Res.* <https://doi.org/10.1007/s11356-018-1372-0>

Santangeli, S., Consales, C., Pacchierotti, F., Habibi, H.R., Carnevali, O., 2019. Transgenerational effects of BPA on female reproduction. *Sci. Total Environ.* 685, 1294–1305. <https://doi.org/10.1016/j.scitotenv.2019.06.029>

Sauvé, S., Desrosiers, M., Mélanie Desrosiers, 2014. A review of what is an emerging contaminant. *Chem. Cent. J.* 8, 1–7. <https://doi.org/10.1186/1752-153X-8-15>

Serra, H., Beausoleil, C., Habert, R., Minier, C., Picard-Hagen, N., Michel, C., 2019. Evidence for Bisphenol B Endocrine Properties: Scientific and Regulatory Perspectives. *Environ. Health Perspect.* 127, 106001. <https://doi.org/10.1289/EHP5200>

Sigrist, C.J.A., de Castro, E., Cerutti, L., Cuche, B.A., Hulo, N., Bridge, A., Bougueleret, L., Xenarios, I., 2012. New and continuing developments at PROSITE. *Nucleic Acids Res.* 41, D344–D347. <https://doi.org/10.1093/nar/gks1067>

Srinivasan, S., Sadasivam, S.K., Gunalan, S., Shanmugam, G., Kothandan, G., 2019. Application of docking and active site analysis for enzyme linked biodegradation of

textile dyes. *Environ. Pollut.* 248, 599–608.
<https://doi.org/10.1016/j.envpol.2019.02.080>

Swerts, C.A.S., Costa, A.M.D.D., Esteves, A., Borato, C.E.S., Swerts, M.S.O., 2010. Effects of fluoxetine and imipramine in rat fetuses treated during a critical gestational period: A macro and microscopic study. *Rev. Bras. Psiquiatr.* 32, 152–158.
<https://doi.org/10.1590/S1516-44462009005000015>

Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–61. <https://doi.org/10.1002/jcc.21334>

Vanacek, P., Sebestova, E., Babkova, P., Bidmanova, S., Daniel, L., Dvorak, P., Stepankova, V., Chaloupkova, R., Brezovsky, J., Prokop, Z., Damborsky, J., 2018. Exploration of Enzyme Diversity by Integrating Bioinformatics with Expression Analysis and Biochemical Characterization. *ACS Catal.* 8, 2402–2412.
<https://doi.org/10.1021/acscatal.7b03523>

Visanji, Z., Sadr, S.M.K., Johns, M.B., Savic, D., Memon, F.A., 2020. Optimising wastewater treatment solutions for the removal of contaminants of emerging concern (CECs): a case study for application in India. *J. Hydroinformatics* 22, 93–110.
<https://doi.org/10.2166/hydro.2019.031>

Webb, B., Sali, A., 2016. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* 2016, 5.6.1–5.6.37. <https://doi.org/10.1002/cpbi.3>

Wilkinson, J., Hooda, P.S., Barker, J., Barton, S., Swinden, J., 2017. Occurrence, fate and transformation of emerging contaminants in water: An overarching review of the field. *Environ. Pollut.* <https://doi.org/10.1016/j.envpol.2017.08.032>

Xia, X., 2017. Bioinformatics and Drug Discovery. *Curr. Top. Med. Chem.* 17, 1709–1726.
<https://doi.org/10.2174/1568026617666161116143440>

Yang, J., Roy, A., Zhang, Y., 2013. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics* 29, 2588–2595. <https://doi.org/10.1093/bioinformatics/btt447>

Yang, J., Roy, A., Zhang, Y., 2012. BioLiP: a semi-manually curated database for biologically relevant ligand–protein interactions. *Nucleic Acids Res.* 41, D1096–D1103.
<https://doi.org/10.1093/nar/gks966>

Zaidi, K.U., Ali, A.S., Ali, S.A., Naaz, I., 2014. Microbial tyrosinases: Promising enzymes for pharmaceutical, food bioprocessing, and environmental industry. *Biochem. Res. Int.* 2014. <https://doi.org/10.1155/2014/854687>

Zimmermann, L., Stephens, A., Nam, S.-Z.Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F.,
Söding, J., Lupas, A.N., Alva, V., 2017. A Completely Reimplemented MPI
Bioinformatics Toolkit with a New HHpred Server at its Core. J. Mol. Biol. 430, 2237–
2243. <https://doi.org/10.1016/j.jmb.2017.12.007>

Figures and Supplementary material captions

Figure 1 Phylogenetic reconstruction of tyrosinases based on amino acid sequences and
inferred using maximum likelihood framework. Tyrosinases from ciliates are in bold. Barr
represent the number of substitutions per 100 aa.

Figure 2 Three-dimensional structures of the six tyrosinases from ciliates (PuTYR-1,
SITYR-1, SITYR-2, UcTYR-1, UcTYR-2 and UTYR-1). The two golden spheres represent
copper atoms and the red rod, a dioxygen molecule.

Figure 3 Representative active site of tyrosinases of ciliates based on the structure of
PuTYR-1. Copper center formed by CuA and CuB (golden spheres), dioxygen (red rod), and
conserved asparagine, histidine, phenylalanine, cysteine residues are presented.

Figure 4 Molecular dynamic evaluation of PuTYR-1, SITYR-1, SITYR-2, UcTYR-1,
UcTYR-2, and UTYR-1 theoretical structures. (A) a root mean square deviation (RMSD)
plot, showing the mean backbone variation over the 50 ns of run; and (B) a radius of gyration
plot, showing the mean variation within the globularity of each structure over the 50 ns of
run.

Figure 5 Binding energies (score) values from virtual screening of 76 CEC structures against
six tyrosinases from ciliates (PuTYR-1, SITYR-1, SITYR-2, UcTYR-1, UcTYR-2 and
UTYR-1) and one from fungi (2y9wA).

Suppl. Figure 1 Secondary structure analyses (DSSP) of the models PuTYR-1, SITYR-1,
SITYR-2, UcTYR-1, UcTYR-2 and UTYR-1 over the molecular dynamics run of 50 ns.

Suppl. Figure 2 Pair-wise genetic distances between tyrosinases of ciliates.

Table 1. Summarized data and general features of the identified tyrosinases of ciliates.

Putative tyrosinase	Species	NCBI accession number	Genomic position (nt)	BLASTP (SwissProt)			General features					
				*HSP	Identity (%)	Host	Length (aa)	Tyrosinase conserved domain (InterPro)	Phobius	SignalP	Molecular weight (kDa)	Isoelectric point
PuTYR-1	<i>Paraurostyla</i> sp. PUJRC_G6	LASR02004360	Contig 4360 (127 – 1713)	Tyrosinase (Q00234)	39,69	<i>Aspergillus oryzae</i>	529	IPR002227 IPR008922	Non-cytoplasmatic	-	59,87	5,05
SITYR-1	<i>Stylonychia lemnae</i>	CCKQ01010535	Contig 0171 (2469 – 4043)	Tyrosinase (Q00234)	39,21	<i>Aspergillus oryzae</i>	525	IPR002227 IPR008922	Non-cytoplasmatic	-	59,66	5,08
SITYR-2	<i>Stylonychia lemnae</i>	CCKQ01013503	Contig 0533 (1717 – 137)	Tyrosinase (Q00234)	39,42	<i>Aspergillus oryzae</i>	527	IPR002227 IPR008922	Non-cytoplasmatic	-	59,8	4,99
UcTYR-1	<i>Uroleptopsis citrina</i>	LXJT01005355	Contig 3255 (1651 – 107)	Tyrosinase (Q00234)	39,47	<i>Aspergillus oryzae</i>	515	IPR002227 IPR008922	Non-cytoplasmatic	-	58,24	5,99
UcTYR-2	<i>Uroleptopsis citrina</i>	LXJT01021807	Contig 3230 (116 – 1657)	Tyrosinase (Q00234)	39,56	<i>Aspergillus oryzae</i>	514	IPR002227 IPR008922	Non-cytoplasmatic	-	58,08	5,95
UTYR-1	<i>Urostyla</i> sp. PUJRC_G1	LASQ02000317	Contig 0317 (2405 – 72)	Tyrosinase (Q00234)	38,84	<i>Aspergillus oryzae</i>	523	IPR002227 IPR008922	Non-cytoplasmatic	-	60,5	6,33

ϕ HSP - High Scoring Pair

Table 2. Structural analyses of ciliate's tyrosinases.

Putative tyrosinase	Template	ID. (%)	Cov.	Ramachandran plot			MolProbity score†	COACH		DoGSiteScorer		5'UTR (aa)	C-Terminal Propeptide (aa)
				Favorable positions (%)	Allowed position (%)	Outliers		Cof.	C-score	Pocket Volume (Å ³)	Pocket Surface (Å ²)		
PuTYR-1	5ZRD C	52,24	0,95	95,7	99	4	2,64	Cu ₂ O ₂	0.24	172.93	250.73	1-4	405-529
SITYR-1	5ZRD C	48,4	0,96	96,4	99	4	0,79	Cu ₂ O ₂	0.19	319.30	359.05	1-4	396-525
SITYR-2	5ZRD C	53,05	0,96	93,1	97,4	10	1	Cu ₂ O ₂	0.27	202.42	270.30	1-4	398-527
UcTYR-1	5ZRD C	51,73	0,96	95,9	99,5	2	2,46	Cu ₂ O ₂	0.26	262.85	389.62	1-4	396-515
UcTYR-2	5ZRD C	50,4	0,97	94,3	98,4	6	1,4	Cu ₂ O ₂	0.23	854.21	1160.99	1-4	395-514
UTYR-1	5ZRD C	52,52	0,95	94,9	97,7	9	2,53	Cu ₂ O ₂	0.23	656.90	709.97	1-4	399-523

†MolProbity score is a single score which combines the clashscore, rotamer, and Ramachandran evaluations. C-score is the confidence score of the COACH prediction. ID. - Identity. Cov. - Coverage. Cof. - cofactor

Table 3. Top 10 binding energies (DockThor score) from virtual screening of ~1000 pesticides and six tyrosinases from ciliates (PuTYR-1, SITYR-1, SITYR-2, UcTYR-1, UcTYR-2 and UTYR-1) and one from fungi (2y9wA).

Tyrosinase	Pesticide (PubChem-CID)	DockThor score	Compound class	Tyrosinase	Pesticide (PubChem-CID)	DockThor score	Compound class
2y9wA					Falone (7209)	-8.930	Organophosphorus
	Dihydrorotenone (73068)	-9.474	Rotenoid		Antimycin A (14957)	-8.840	Bis-lactone
	Deltamethrin (40585)	-9.378	Pyrethroid ester		Ethion (3286)	-8.786	Organophosphate
	Synthetic-pyrethrins (2982)	-9.356	Synthetic pyrethrin		Dihydrorotenone (73068)	-8.757	Rotenoid
	Cypermethrin (2912)	-9.205	Carboxylic ester		Tralomethrin (48132)	-8.718	Carboxylic ester
	Brodifacoum (54680676)	-9.201	Benzene		Tyrosine (6057)	-6.641	Aromatic amino acid
	Lactofen (62276)	-8.987	Diphenyl ether	UcTYR-1	Falone (7209)	-9.635	Organophosphorus
	Cis-permethrin (40463)	-8.956	Pyrethroid ester		Brodifacoum (54680676)	-9.202	Benzene
	Bromadiolone (54680085)	-8.942	Diarylheptanoid		EDO (161756)	-8.983	Nonchlorinated oxetane
	Fenoxycarb (51605)	-8.891	Carbamate ester		Phenoxathiin (9217)	-8.920	Ether
	Cyfluthrin (104926)	-8.865	Carboxylic ester		Isocycloheximide (2900)	-8.793	Piperidone
	Tyrosine (6057)	-6.779	Aromatic amino acid		Hydramethylnon (5281875)	-8.656	Hydrazone
PuTYR-1	Brodifacoum (54680676)	-9.388	Benzene		Chinomethionat (17109)	-8.634	dithioquinoxaline
	Hydramethylnon (5281875)	-9.056	Hydrazone		Bifenthrin (6442842)	-8.631	Carboxylic ester
	Bromadiolone (54680085)	-8.906	Diarylheptanoid		2,4,5-T 2-ethylhexyl ester (15998)	-8.628	Ester
	Caswell No. 277 (62011)	-8.860	Polyoxyethylene		Ethion (3286)	-8.624	Organophosphate
	Flucythrinate (50980)	-8.847	Pyrethroid		Tyrosine (6057)	-6.583	Aromatic amino acid
	Rotenone (6758)	-8.733	Rotenoid	UcTYR-2	Brodifacoum (54680676)	-8.953	Benzene
	Deltamethrin (40585)	-8.680	Pyrethroid ester		Falone (7209)	-8.940	Organophosphorus
	Reserpine (5770)	-8.663	Alkaloid		Hydramethylnon (5281875)	-8.885	Hydrazone
	Fluazifop-p-butyl (3033674)	-8.642	Butyl ester		Rotenone (6758)	-8.538	Rotenoid
	Dienochlor (16686)	-8.606	Organochlorine		Dihydrorotenone (73068)	-8.529	Rotenoid
	Tyrosine (6057)	-6.672	Aromatic amino acid		Hexaflumuron (91741)	-8.480	N-acylurea
SITYR-1	Falone (7209)	-8.698	Organophosphorus		Propyl isome (6750)	-8.469	Naphthoic acid
	Brodifacoum (54680676)	-8.670	Benzene		Synthetic-pyrethrins (2982)	-8.458	Synthetic pyrethrin
	Reserpine (5770)	-8.607	Alkaloid		Bifenthrin (6442842)	-8.454	Carboxylic ester
	Flusilazole (73675)	-8.605	Organosilicon		Flucythrinate (50980)	-8.413	Pyrethroid
	Methoxyfenozide (105010)	-8.562	Carbohydrazide	UTYR-1	Tyrosine (6057)	-6.947	Aromatic amino acid
	Triarathene (6450825)	-8.514	Thiophene		Brodifacoum (54680676)	-8.919	Benzene
	Flucythrinate (50980)	-8.487	Pyrethroid		Falone (7209)	-8.538	Organophosphorus
	Bromadiolone (54680085)	-8.468	Diarylheptanoid		Flucythrinate (50980)	-8.453	Pyrethroid
	Synthetic-pyrethrins (2982)	-8.467	Synthetic pyrethrin		Oxadiazon (29732)	-8.328	Aromatic ether
	Clofenotane (3036)	-8.455	Chlorinated hydrocarbon		Mestranol (6291)	-8.310	Semisynthetic estrogen
	Tyrosine (6057)	-6.498	Aromatic amino acid		Triarathene (6450825)	-8.252	Thiophene
SITYR-2	Hydramethylnon (5281875)	-9.352	Hydrazone		1-Hexadecanol (2682)	-8.241	Alcohol
	Reserpine (5770)	-9.165	Alkaloid		Dieldrin (969491)	-8.225	Organochlorine
	Brodifacoum (54680676)	-9.048	Benzene		Reserpine (5770)	-8.198	Alkaloid
	Bromadiolone (54680085)	-8.974	Diarylheptanoid		Kadethrin (12940701)	-8.183	Synthetic pyrethroid
	Bromophos-ethyl (20965)	-8.962	Organic thiophosphate		Tyrosine (6057)	-6.878	Aromatic amino acid

Tyrosine, a known tyrosinase ligand is highlighted in bold for reference.

Figure 1

Figure 2

Model	Conserved residues within the active type-3 copper center
PuTYR-1	H50, C69, ^ψ H71, F76, H80, H265, H269, N299, F308, H312
SITYR-1	H50, C70, ^ψ H72, F77, H81, H258, H262, N292, F301, H305
SITYR-2	H51, C71, ^ψ H73, F78, H82, H259, H263, N294, F303, H307
UcTYR-1	H50, C69, ^ψ H71, F76, H80, H260, H264, N291, F300, H304
UcTYR-2	H50, C69, ^ψ H71, F76, H80, H260, H264, N291, F300, H304
UTYR-1	H50, C69, ^ψ H71, F76, H80, H262, H266, N296, F305, H309

H - Histidine; C - cysteine; F - phenylalanine; N - asparagine. ^ψFlexible Histidines within loop regions. CuA linked residues are in blue and CuB, in green.

Figure 3

Figure 4

Figure S1

	PuTYR-1	SITYR-1	SITRY-2	UcTYR-1	UcTYR-2	UTYR-1
PuTYR-1	-					
SITYR-1	0,493014	-				
SITRY-2	0,473054	0,205589	-			
UcTYR-1	0,530938	0,50499	0,499002	-		
UcTYR-2	0,526946	0,518962	0,500998	0,145709	-	
UTYR-1	0,500998	0,477046	0,46507	0,48503	0,483034	-

Figure S2