



IKER
GAZTE
NAZIOARTEKO
IKERKETA EUSKARAZ

IV. IKERGAZTE

NAZIOARTEKO IKERKETA EUSKARAZ

2021eko ekainaren 9, 10 eta 11a
Gasteiz, Euskal Herria

ANTOLATZAILEA:
Udako Euskal Unibertsitatea (UEU)

ZIENTZIAK ETA NATURA ZIENTZIAK

**Irudi-analisi eta machine learning
bidezko itsas-triku enbrioi
biosaiaren automatizazioa**

*Iker Alvarez, Leire Mijangos,
Naroa Lopez, Harkaitz Eguiraun,
Jose Manuel Amigo,
Mathilde Monperrus eta
Nestor Etxebarria*

175-180 or.
<https://dx.doi.org/10.26876/ikergazte.iv.05.23>



Irudi-analisi eta machine learning bidezko itsas-triku enbrioi biosaioaren automatizazioa

Alvarez-Mora, I.^{1,2}, Mijangos, L.^{1,2}, Lopez-Herguedas, N.^{1,2}, Eguiraun, H.^{2,3}, Amigo, J.M.^{1,2}, Monperrus, M.⁴, Etxebarria, N.^{1,2}

1. Kimika Analitikoa Saila (Euskal Herriko Unibertsitatea, Leioa, Euskal Herria)
2. Plentziako itsas estazioa (Euskal Herriko Unibertsitatea, Plentzia, Euskal Herria)
3. Adierazpen Grafikoa eta Ingeniaritzako Proiektuak Saila (Euskal Herriko Unibertsitatea, Bilbao, Euskal Herria)
4. Institut des Sciences Analytiques et de Physico-chimie pour l'Environnement et les matériaux (Université de Pau et des Pays de l'Adour, Pau, France)

Iker.alvarez@ehu.eus

Laburpena

Zorionez, sortutako kutsatzaileengan jarritako arreta gero eta handiagoa da, era horretan, horien helmena eta eragin ditzaketen efektuak ikertzeko eta alde zuzenetik ezagutzeko. Ingurumen lagin konplexuen kausa-efektu erlazioak azaltzeko analisi kimiko eta toxikologikoaren bateratzea beharrezkoa dela ondorioztatu da. Efektuei Bideratutako Analisia (EDA, effect-directed analysis) bateratze horren adibide garbia da, kutsatzaile nahasketa konplexuetan egon daitezkeen konposatu toxikoen identifikazioa errazten duelarik. Lan honetan, EDA metodologietan aplikatu daitezkeen itsas-triku biosaioaren automatizazioa proposatzen da irudi analisi eta ikasketa automatikoko tresnak erabiliz, prozesuaren lan-karga eta denbora murrizteko. Jatorriko biosaioan neurtzen diren hazkuntza eta aldatze maila, modu automatiko batean neurtzea lortu da, irudi-irakurle baten bidez argazkiak automatikoki aterata ondoren. Prozesua MATLABen programatutako kodearen bitartez lortu da, eta aplikazio batean bildu da, programazio ezaguerarik ez dituztenek jarraitu ahal izateko.

Hitz gakoak: ekotoxikologia, itsas-triku enbrioi testa, automatizazioa, irudi-analisisa, ikasketa-automatikoa, efektuei bideratutako analisisa.

Abstract

Hopefully, the efforts done in the last years give us a better idea of the contaminant's fate and their effects. However, elucidating the cause-effect relations of complex contaminants mixtures is not an easy task. Usually, chemical analysis techniques are coupled to ecotoxicological bioassays in complex pipelines such as Effect Directed Analysis (EDA). This kind of pipelines have a clear drawback, their workload, that make the methodology hard to apply when the sample amount is too big. In this work, the automatization of a bioassay, the sea-urchin embryo test (SET), is proposed to be used in an EDA approach. The endpoints measured in the original sea-urchin embryo test were automatized using an image reader, image analysis and machine learning. All the scripts used for this task were compiled to create an app, available for any type of user.

Keywords: ecotoxicology, sea-urchin embryo test, automatization, image-analysis, machine learning, effect directed analysis.

1. Sarrera eta motibazioa

Azkenengo urteetan, ingurumenara isurtzen diren hondakinen karakterizazio kimikoa hobeto ulertzeko egin diren aurrerakuntzei esker, kutsatzaile azpi-talde berri bat izendatu da, ingelesez *Contaminants of Emerging Concern* edo CEC bezala ezagutzen dena. Askotan kutsatzaile emergente bezala ezagutzen badira ere, kasu askotan berria dena ez da kutsatzaile hauen presentzia ingurumenean, baizik eta hauekiko sortutako kezka. Multzo horretan sartzen diren konposatuak erregulazio edo arautik kanpo daude eta haien artean aurkitzen ditugu eguneroko bizitzan erabilitako zainketa pertsonalerako produktuetatik, medikamentuetara; industria-produktuetatik, nekazaritzan erabiltzen diren pestizidetara. Behin bere funtzioa bete ondoren,

konposatu hauek ur-araztegietan edo zuzenean ingurumenean amaitzen dute. CEC-ei buruzko hainbat ikerketen emaitzek erakutsi dute araztegiak ez direla gai kutsatzaile hauek eliminatzeko, araztegiko efluenteak kutsatzaile zopa bat direla frogatuz (Loos et al., 2013).

Erronka ekotoxikologikoak, kutsatzaile nahasketa konplexu honetan datza, hain zuzen. Kutsatzaileen arrisku-ebaluaziorako hamaika biosaio erabiltzen diren arren (Brack et al., 2016), *in-vivo* zein *in-vitro*, hauek ezer gutxi argitzen digute aurkitutako efektu toxikologikoen jatorriari buruz. Hori dela eta, *in-vivo* biosaioak askotan analisi kimikoko teknikekin batera erabiltzen dira, kausa-efektu erlazioak azaldu nahian. Estrategia hauen artean adibide nagusia, Efektuei Bideratutako Analisia (*Effect Directed Analysis*, EDA) da (Burgess et al., 2013). Erabilitako estrategia edozein izanda, kausa-efektuak ulertzeko prozesua geldoa eta lan-karga handikoa da.

Zientzia doan erritmora, denbora galtzea ez da aukera bat. Erlojuaren aurkako lehia oso presente dago zientzialariaren bizitzan. Emaizta egokiak ematea bezain garrantzitsua da gaur egun, hauek, azkar ematea, hau da, segundora optimizatzea emaitzen eskurapena; batez ere, behin baino gehiagotan errepikatzen diren prozesuez mintzatzerako orduan. Prozesu errepikakor hauen optimizazioan eta automatizazioan datza, hain zuzen, modan dauden HTS (*High-Throughput Screening*) metodoak. Metodo hauek robotikaz, datu prozesamendu softwarez, etab.-ez baliatzen dira, modu tradizional batean luze joko lukeen prozesu bat minutu gutxi batzuetan aurrera eramateko.

2. Arloko egoera eta ikerketaren helburuak

Aipatutako EDA metodologiaren barruan, hainbat biosaio erabili daitezke laginaren toxikotasun efektuak aztertzeke. Aztertu nahi den erantzunaren arabera, espezifikoak diren testak (adibidez, disrupzio endokrinoa aztertzeke *Yeast Estrogen Screen*, YES) (Rutishauser et al., 2004), organismo baten gainean hainbat 'endpoint' aztertzen dituzten test ez-espezifikoak (adibidez, zebra-arrain enbrioi testa, ZFET)(OECD, 2013), toxikotasun erreaktiboa neurtzen duten testak (adibidez, Ames testa, mutagenizitatea neurtzeke) edota haien konbinazioak aurki ditzakegu EDAn (Escher et al., 2011). Era berean, oso garrantzitsua da aukeratutako biosaioan erabiltzen den espeziea aztertzen ari den ingurumenerako adierazgarria izatea.

Lerro horretan, itsas-triku enbrioi testa edo SET (*Sea-urchin Embryo Test*) itsas-ingurumenerako biosaio esanguratsuenen artean dago. Euskal Herriko kostaldean aurkitzeko aukerak eta akuarioetan hazteko erraztasunak, gure kostaldeko ekosistema aztertzeke ezinbesteko espeziea bilakatzen dute itsas-trikua (Mijangos et al., 2020). Entsegu honetan, kutsatzaileei esposatutako enbrioiaren hazkuntza eta malformazioak aztertzen dira 48 orduz inkubatu ondoren. Alderantzizko mikroskopioaren bidez, larben tamaina banaka neurtzen da eta Carballeira et al. -ek proposatutako lau alderazio mailen artean sailkatzen dira (Carballeira et al., 2012) (0. maila garapen normaleko larbak definitzeko; 1. maila eskeletoko erpina malformatuta duten larbak; 2. maila eskeletoaren malformazio nagusiak dituztenak; eta 3. maila garapen oso murriztua izan duten larbak). Maila bakoitzari indize bat aplikatuz, larben kontaketa egiten da eta toxikotasun indizea (IT) kalkulatu daiteke ondorengo ekuazioan (1) azaltzen den bezala:

$$IT = [0 \times \% 0. maila + 1 \times \% 1. maila + 2 \times \% 2. maila + 3 \times \% 3. maila] / 100 \quad (1)$$

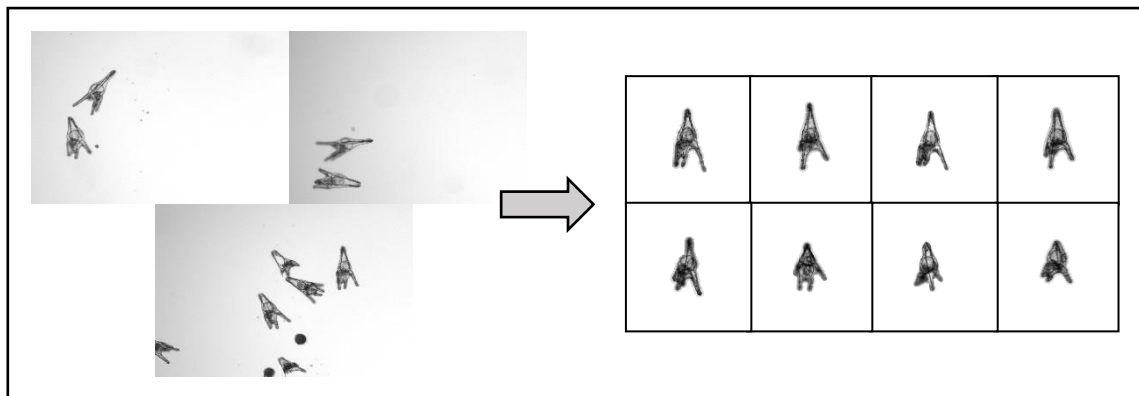
Lagin kopuruaren arabera egunak iraungo lituzkeen prozesu hau arintzea eta automatizatzea da, hain zuzen ere, lan honen helburu nagusia; irudi eskuratze automatikoaz, irudi-analisisaz eta ikasketa automatikoaz (*machine learning*) baliatuz.

3. Ikerketaren muina

3.1. Esposizio esperimentuak eta irudi prozesatzea

Aipatutako helburua lortzeko, esposizio esperimentuak gauzatu dira kutsatzaile desberdinekin, alterazio sorta zabal bat erakusten dituzten larben irudiak lortzeko. Albendazola, amitriptilina, kafeina eta kobre (II) kloruroa erabili dira esposizioarako, larbak 48 orduz eta 20 °C-tan inkubagailuan inkubatu direlarik. Grisen-eskalako irudiak lortzeko Cytation™ 5 (4x handipen objektiboa, BIOTEK) irudi irakurgailua erabili da, plaka osoan zehar 16 argazkiko ekorketa eginez aurretiaz, larbak finkatzeko, formalina tanta bat gehitu den putzu bakoitzeko. Larben tamaina neurtzeko eta malformazio egoeraren arabera sailkatu ahal izateko irudiak prozesatu dira, hainbat larba agertzen diren irudi gordinetatik, larba bakoitza orientatuta eta zentratuta izan arte (ikus 1. Irudia).

1. irudia. Irudi prozesatzea irudi gordinetik irudi tratatuetera



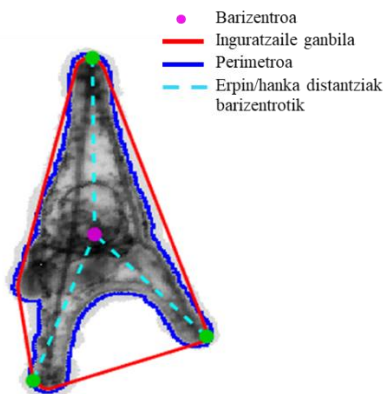
Irudi gordinetik larbak ateratzeko, jatorrizko irudietan pixel-intentsitate atalase bat ezarriz lortu diren irudi binarioetan dilatazioa aplikatu da, eta artefaktu txikiak eta ertzetako elementuak kendu dira. Gainerako objektuak, larbei dagokienak alegia, erabili dira irudi gordinetik hauek erauzteko, orientatzeko eta 500x400 pixeletako irudi berrietan normalizatzeko (dimentsio zehaztutako irudiak lortuz). Honela, 270 larben irudi-bilduma lortu da. Bilduma hori bitan banatu da, 240 irudirekin sailkapenerako eredua eta metodologiaren optimizazioa egiteko entrenamenduko sorta prestatuz; eta gainontzeko 30-ekin berriz, eredua berresteko sorta. Irudi prozesamendurako *MATLAB*-en garatutako kodigoa erabili da *Image Processing Toolbox*-eko zenbait funtzioz baliatuz.

3.2. Itsas-triku enbrioi testaren automatizazioa.

Esan bezala, lan honen muina SET biosaioa eguneratzean datza. Horretarako, kimiometriak, larba irudien sailkapena egiteko behar ditugun tresna guztiez hornitzen gaitu. Lan honetan *Partial Least Squares Discriminant Analysis* (PLS-DA) algoritmoaren bi moldaera konparatu dira, larbak Carballeirak proposatutako taldeen artean banatzeko. Irudi prozesamenduaren mugak direla eta, erdiko bi egoerak (1 eta 2 mailak) bateratu dira.

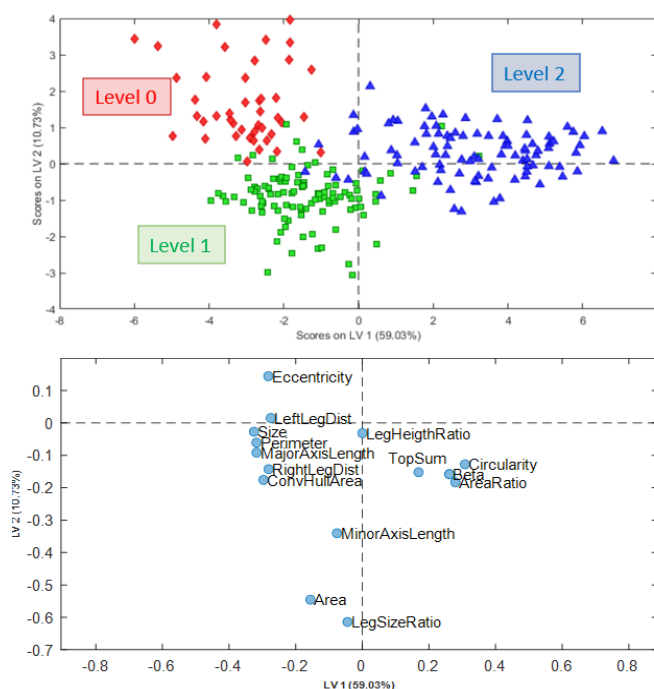
Sailkapen erduekin sartu aurretik, hauetan erabiliko den datu sarrera lortu da. Irudien informazioarekin modu eraginkor batean lan egiteko, argazkiak parametrizatu dira. Larben egoera definitzen duten 17 parametro aukeratu dira, adibidez, larbaren luzera, zabalera, azalera, biribiltasuna, perimetra, etab. (ikus 2. irudia). Entrenamendurako larba bakoitza parametrizatu ostean, sailkapen erduek eraikitzeke erabiliko den 270x17-ko datu matrizea lortu da.

2. irudia. Larba irudien parametroen adibidiea



Larbak alterazio egoeraren arabera sailkatzeko gai den eredu bat eratzeko, lehenik, garrantzitsua da, adituen laguntzaz, kalibraziorako irudi sortaren sailkapena egitea. 240 argazkietatik 37, 0.mailan sailkatu dira, 107, 1.mailan eta 98, 2.mailan. Arestian aipatutako bi PLS-DA moldaera konparatu dira: lau aldagai latenterekin (59%, 11%, 9% eta 5%-ko bariantza azalduz, hurrenez hurren) eraikitako 3 taldeko PLS-DA (ikus 3. irudia) eta bata bestearen atzean eraikitako PLS-DA hierarkikoa (HPLS-DA). HPLS-DA-ren kasuan, lehenengo 0. eta 1. mailak 2.-etik bereizi dira, eta ondoren, 0. eta 1. mailak bata bestearekiko.

3. irudia. Hiru talderekin eraikitako PLS-DA ereduaren Scores (goian) eta Loadings (behean) irudikapenak.



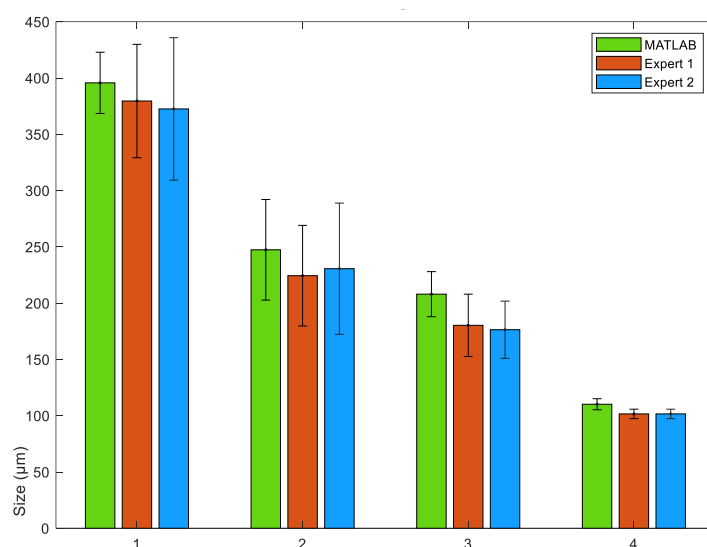
Eredu biak berretsi dira berrespenerako sortarekin, haien egokitasuna aztertu ondoren. 1. taulan ikus dezakegun bezala HPLS-DA ereduak sailkapenerako gaitasun hobea erakusten du, bai positibo eta negatibo faltsuak baztertzeko gaitasunari (zehaztasuna eta sentikortasuna, hurrenez hurren), bai iragarpen erroreari erreparatuz. Berrespenaren iragarpen erroretik ereduaren eta adituen arteko adostasun neurria lortu dezakegu, %84-koa.

1. taula. Konparatu diren bi sailkapenerako ereduaren balidaziorako parametroak.

	HPLS-DA		PLS-DA	
	Kalibrazioa	Berrespena	Kalibrazioa	Berrespena
Sentikortasuna	0.916	0.844	0.908	0.806
Zehaztasuna	0.954	0.952	0.920	0.902
Iragarpenaren errorea	0.084	0.156	0.092	0.194

Itsas-triku enbrioi testaren automatizazioarekin amaitzeko, larben tamainaren neurketa automatizatu da. Prozesatutako irudien gainean aldenduen dauden bi pixelen arteko distantzia neurtu da. Ondoren, kalibratzeko erabili den eskaladun irudi baten laguntzaz, distantzia mikrometroara pasatu da. Metodo hau berresteko, bigarren esperimentu bat egin da, hiru larba talde kobre (II) kloruro kontzentrazio desberdinekin esposatuz. Gero, talde bakoitzetik hartutako larbak, bi adituen laguntzaz eta garatutako metodoarekin neurtu dira aldi berean, euren artean konparatuz (ikus 4.irudia). Gainera, ernaldutako bezain laster formalinarekin finkatutako arrautzen neurketak ere egin dira, hazkuntza kalkulatu ahal izateko (hazkuntza, larbaren tamainari arrautzaren tamaina kenduaz ematen da). Konparatutako hiru bideek ez dute desberdintasun adierazgarriarik erakusten, era horretan, metodoaren egokitasuna erakutsiz.

4. irudia. Bi adituek eskuz neurtutako eta MATLAB-eko kodea aplikatuz neurtutako itsas-triku larben arteko konparaketa. 1. Taldea kobre (II) kloruro 0,01 mg·L⁻¹; 2. Taldea 0,1 mg·L⁻¹; 3. Taldea 0,5 mg·L⁻¹ eta 4. Taldea arrautzei dagokie.



Automatizatutako biosaio hau benetan erabilgarria izateko eta programazioko ezagutzarik ez duen edozeinek erabili ahal izateko, aplikazio bat eraiki da. Argazki gordinak atera ondoren, aplikazioan irudi hauen jatorrizko karpeta aukeratzea nahikoa da, botoi bat sakatzuz, lan honetan zehar azaldutako prozesu osoa martxan jartzeko. Behin irudien prozesatzea amaituta, bai neurketen bai sailkapenaren emaitzak grafiko zein taula moduan azaltzen dira.

4. Ondorioak

Gaur egun, ezinbestekoa da zenbait metodologia zaharberritzea. Lan honetan, lehen, 10 laginen toxikotasuna aztertzeko egunak behar zituen prozesua ordu gutxitan eta modu automatiko batean egiteko aukera erakutsi da. Emaitzei erreparatuz, hazkuntzan ez da inolako desberdintasunik aurkitu modu tradizionalarekin konparatuz. Sailkapenean aldiz, ikusi da %84-ko bat-etortzea dagoela adituen sailkapenarekin. Hala ere, kontuan hartu behar da, modu automatizatu honetan aditu desberdinen subjektibotasuna desagertzen dela, tresna errepikakorragoa bilakatuz.

5. Etorkizunerako planteatzen den norabidea

Metodologia garatu ondoren, lan honen hurrengo urratsa argi dago, kasu erreal batean aplikatzea. Eraikitako ereduaren sendotasuna baieztatzeko honen aplikagarritasuna aztertu behar da, batez ere, honelako teknologiarekin trebea ez den erabiltzailearentzat. Gainera, entsegu berriek, larba irudi bilduma handitzeko aukera ematen digute, eredia hobetuz.

6. Erreferentziak

- Burgess, R. M., Ho, K. T., Brack, W., & Lamoree, M. (2013). Effects-directed analysis (EDA) and toxicity identification evaluation (TIE): Complementary but different approaches for diagnosing causes of environmental toxicity. *Environ. Toxicol. Chem.*, 32(9), 1935–1945.
- Brack, W., Ait-Aissa, S., Burgess, R. M., Busch, W., Creusot, N., Di Paolo, C., Escher, B. I., Mark Hewitt, L., Hilscherova, K., Hollender, J., Hollert, H., Jonker, W., Kool, J., Lamoree, M., Muschket, M., Neumann, S., Rostkowski, P., Ruttkies, C., Schollee, J., ... Krauss, M. (2016). Effect-directed analysis supporting monitoring of aquatic environments — An in-depth overview. *Sci. Total Environ*, 544, 1073–1118.
- Carballeira, C., Ramos-Gómez, J., Martín-Díaz, L., & DelValls, T. A. (2012). Identification of specific malformations of sea urchin larvae for toxicity assessment: Application to marine pisciculture effluents. *Marine Environmental Research*, 77, 12–22.
- Escher, B., & Leusch, F. (2011). *Bioanalytical Tools in Water Quality Assessment*. <https://doi.org/10.2166/9781780400778>
- Loos, R., Carvalho, R., António, D. C., Comero, S., Locoro, G., Tavazzi, S., Paracchini, B., Ghiani, M., Lettieri, T., Blaha, L., Jarosova, B., Voorspoels, S., Servaes, K., Haglund, P., Fick, J., Lindberg, R. H., Schwesig, D., & Gawlik, B. M. (2013). EU-wide monitoring survey on emerging polar organic contaminants in wastewater treatment plant effluents. *Water Research*, 47(17), 6475–6487.
- Mijangos, L., Krauss, M., de Miguel, L., Ziarrusta, H., Olivares, M., Zuloaga, O., Izagirre, U., Schulze, T., Brack, W., Prieto, A., & Etxebarria, N. (2020). Application of the Sea Urchin Embryo Test in Toxicity Evaluation and Effect-Directed Analysis of Wastewater Treatment Plant Effluents. *Environ Sci. & Technol*, 54(14), 8890–8899.
- OECD (2013), Test No. 236: Fish Embryo Acute Toxicity (FET) Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris,
- Rutishauser, B. V.; Pesonen, M.; Escher, B. I.; Ackermann, G. E.; Aerni, H.-R.; Suter, M. J.-F.; Eggen, R. I. L. (2004) : Comparative Analysis of Estrogenic Activity in Sewage Treatment Plant Effluents Involving Three In Vitro Assays and Chemical Analysis of Steroids. *Environ. Toxicol. Chem.* 23 (4), 857-864.

7. Eskerrak eta Oharrak

Lan hau Espainiako Ekonomia eta Lehiakortasun Ministeritzak diruz lagundu du CTM2017-84763-C3_1_R proiektuaren bitartez. Bestetik, Autoreak eskerrak eman nahi dizkie Euskal Herriko Unibertsitateari (UPV/EHU) eta Université de Pau et des Pays de l'Adour-i (UPPA) doktoretza-aurreko diru laguntzagatik.