

Miquel-Àngel Sánchez Ferriz (cur.)

La terminologia en les ciències de la vida, en la química i en el món educatiu

Barcelona: Institut d'Estudis Catalans, 2013, p. 33-38

DOI: 10.2436/15.2503.02.28

Recursos terminològics en línia de les ciències de la vida

Ricard ROCA

Societat Catalana de Biologia. Institut d'Estudis Catalans

Resum

En aquest article es repassen algunes de les eines en línia més útils per a professionals de la llengua pel que fa a dos camps terminològics de les ciències de la vida especialment complexos des del punt de vista ortotipogràfic: la taxonomia i la nomenclatura de gens i proteïnes. S'hi esmenten unes quantes normes bàsiques i es donen una sèrie de recursos en línia per facilitar les consultes habituals que pot necessitar el professional de la llengua.

PARAULES CLAU: eines en línia, terminologia, ciències de la vida, taxonomia, nomenclatura de gens i proteïnes.

Abstract: *Terminology on-line resources of life sciences*

In this paper a review is made of some of the on-line resources more useful to language professionals as regards two terminology fields in life sciences especially complex from the orthotypographic viewpoint: taxonomy, and nomenclature of genes and proteins. Various basic rules are mentioned and a series of on-line resources are given in order to facilitate the usual queries needed by language professionals.

KEY WORDS: on-line tools, terminology, life sciences, taxonomy, nomenclature of genes and proteins.

1. INTRODUCCIÓ

Alguns camps d'aplicació de la terminologia són especialment complexos des del punt de vista ortotipogràfic. En el cas de les ciències de la vida, dos dels casos són la taxonomia i la nomenclatura de gens i proteïnes. En tots dos casos,

l'ortotipografia està regulada per codis i normes escrites complexos, que sovint introdueixen canvis. El Web (entès com a sistema de consulta i publicació de documents accessible als ordinadors en xarxa) ha representat una millora substancial pel que fa a la consulta d'aquestes normes, atès que la fa molt més àgil i ràpida, i permet que els materials i criteris nous siguin accessibles molt abans del que ho serien amb els mitjans de publicació en paper habituals.

2. TAXONOMIA

No hi ha un únic codi que reguli la taxonomia de tots els éssers vius, sinó que els organismes estudiats tradicionalment per la zoologia, la botànica, la microbiologia i la virologia tenen codis diferents, que són, respectivament, el *Codi internacional de nomenclatura zoològica* (ICZN), el *Codi internacional de nomenclatura d'algues, fongs i plantes* (ICN), el *Codi internacional de nomenclatura bacteriana* (ICNB) i el Comitè Internacional de Taxonomia de Virus (ICTV). Aquests codis estableixen normes que poden ser lleugerament diferents, tot i que es basen, tret dels dels virus, en la **nomenclatura binomial** i els **tàxons supraespecífics**.

La nomenclatura binomial va ser sistematitzada per Carl von Linné. En aquesta, el nom d'una espècie consta d'un **gènere** i un **epítet específic**, escrits en cursiva. El nom del gènere s'escriu amb majúscula inicial, i l'epítet específic en minúscules (p. ex., *Hypoderma bovis*) i, després de la primera citació, el gènere es pot abreujar amb la majúscula inicial seguida de punt (*H. bovis*).

La cursiva té funció emfàtica i, si cal, es pot substituir pel subratllat. No és obligatòria en llistes i índexs en què tot el que apareixen són noms científics, ja que no cal distingir-los del text principal. En contextos formals, com descripcions i altres treballs acadèmics, se cita l'autoritat que ha anomenat l'espècie, escrita en rodona: *Patella vulgata* (Linnaeus, 1758). Hi ha abreviacions establertes per als noms de les autoritats; per exemple, Linnaeus esdevé L. L'any que de vegades es posa després de l'autoritat només és obligatori per a zoòlegs, no per a botànics. Hi ha, també, tota una sèrie d'abreviatures (com sp., per a referir-se a espècie, quan només es vol esmentar el gènere), que s'escriuen en rodona.

Els tàxons supraespecífics tenen un nom formal llatí i, en moltes llengües romàniques, un nom vulgar (p. ex., *Mollusca* i mol·luscs). En català, com en altres llengües romàniques, el nom llatí només s'usa en contextos molt formals, però en anglès, per exemple, s'usa sempre el nom llatí, tret de casos en què s'ha lexicalitzat un nom comú. El nom llatí s'escriu en rodona i amb majúscula inicial, mentre que el nom català s'escriu com un nom comú, en minúscules i rodona. Alguns tàxons supraespecífics acaben amb un sufix determinat, depenent del tipus d'organisme. Alguns exemples d'angiospermes: *-àcies*, *-als*, *-fits*, etc. I alguns exemples d'animals: *-formes*, *-ids*, *-pters*, etc.

En els quatre subapartats següents s'esmenten unes quantes adreces on es poden trobar alguns codis taxonòmics per a poder consultar-los en línia:

2.1. *Nomenclatura zoològica*

En relació amb la nomenclatura zoològica és imprescindible esmentar el *Codi internacional de nomenclatura zoològica* (<http://www.nhm.ac.uk/hosted-sites/iczn/code/>), el web de l'ICZN (<http://iczn.org/>) i la base de dades de Zoobank (<http://www.zoobank.org/>).

2.2. *Nomenclatura d'algues, fongs i plantes*

En aquest aspecte, l'adreça imprescindible es la del *Codi internacional de nomenclatura d'algues, fongs i plantes* (<http://ibot.sav.sk/icbn/main.htm>). Aquesta versió és el *Codi de Viena* de 2005, la immediatament anterior a l'actual, de 2011, coneguda com a *Codi de Melbourne*.

Com a codi complementari, hi ha el *Codi internacional de nomenclatura de plantes cultivades*.

2.3. *Nomenclatura bacteriana*

En aquest respecte, l'adreça imprescindible és la del *Codi internacional de nomenclatura bacteriana* (un resum del text publicat del qual és accessible a <http://www.bacterio.cict.fr/foreword.html>), la base de dades taxonòmica de l'LPSN o *List of Prokaryotic names with Standing in Nomenclature* (<http://www.bacterio.cict.fr/>) i l'*International Journal of Systematic and Evolutionary Microbiology* (<http://ijs.sgmjournals.org/>).

2.4. *Nomenclatura de virus*

Els noms científics d'espècies de virus no segueixen la nomenclatura binomial, com ocorre amb els organismes cel·lulars. Consten de diversos mots (normalment en anglès) seguits del terme *virus*. S'escriuen en cursiva amb majúscula inicial només en la primera paraula i en les que corresponguin a noms propis. Per exemple, *Influenza A virus*, en català esdevé *virus de la grip A*. També, la classificació de l'ICTV usa codis decimals per a una identificació més precisa de les espècies, amb els nivells de tàxons separats per punts (p. ex., 00.046.0.01.001 és el codi d'aquest virus).

Els tàxons supraespecífics dels virus no són els mateixos que per als organismes cel·lulars, i van de l'ordre a l'espècie (amb alguns tàxons infraespecífics).

Igual com passa amb la resta d'organismes, els tàxons dels virus també acaben amb sufixos determinats, segons el rang (-*virus*, -*virinae*, -*viridae* i -*virales*). Aquests noms, en usos formals, s'escriuen en cursiva i amb majúscula inicial, com per exemple, *Picornaviridae*.

I ara esmentem un parell d'adreces imprescindibles per a la nomenclatura de virus: la del web de l'ICTV (<http://www.ictvonline.org/index.asp?bhcp=1>) i la de Codi de nomenclatura de virus (http://www.ictvonline.org/codeOfVirusClassification_2002.asp).

3. NOMENCLATURA DE GENS I PROTEÏNES

Hi ha unes normes ortotipogràfiques més o menys genèriques per a escriure els diferents tipus de gens i proteïnes, encara que en molts casos per a cada organisme model que s'estudia (és a dir, els organismes en què es coneix millor el genoma i el proteoma) canvien en aspectes més específics, amb la qual cosa, dependent de l'abast de l'obra cal consultar les regles amb profunditat amb l'autor. Per regla general, els **gens** s'escriuen en cursiva i les **proteïnes** que codifiquen, en rodona. Per exemple, la proteïna que codifica el gen *shh* en *Xenopus* és *shh*. Per analogia, també s'escriuen en cursiva les denominacions dels **genotips** i en rodona les dels **fenotips**. Els gens **ortòlegs** entre espècies es denominen amb les mateixes lletres, sense afegir-ne d'altres que facin referència a l'espècie (com *Z* per al peix zebra, una pràctica antiga en aquesta espècie). Per exemple, *SHH* i *Shh* són gens ortòlegs en humans i ratolí.

Com hem dit, cada **organisme model** segueix les seves regles ortogràfiques pel que fa a les majúscules, els superíndexs i altres símbols que s'utilitzen en la nomenclatura. Aquestes regles es mantenen actualitzades en línia mitjançant dues eines: les **normes de nomenclatura** i les **bases de dades** de gens i proteïnes. En les normes de nomenclatura hi ha les especificacions generals, i en les bases de dades es poden consultar els noms de símbols i proteïnes, que poden estar actualitzats després de la publicació de les normes. Els investigadors que fan recerca amb un organisme determinat són els que proposen el nom per als gens, d'acord amb les regles establertes.

Dependent de l'organisme, es diferencia entre el **nom complet** del gen i el **símbol** del gen. Aquesta distinció es pot estendre a les proteïnes. Per exemple, *shh* és un símbol i *sonic hedgehog* és el nom complet que hi correspon. De vegades, però, alguns noms de gens establerts abans de la codificació actual es mantenen per tradició encara que no segueixin les regles establertes.

A continuació esmentem algunes institucions que s'encarreguen d'establir la nomenclatura per a diferents organismes model, juntament amb algunes regles concretes per a cada cas, i adreces amb normes en línia i bases de dades.

Pel que fa a la nostra espècie, *Homo sapiens*, l'HGNC (HUGO Gene Nomenclature Committee) estableix la nomenclatura per als gens. Els gens humans s'escriuen amb totes les lletres en majúscula, en cursiva, i les proteïnes s'escriuen igual, però en rodona.

Normes de l'HGNC: <http://www.genenames.org/guidelines.html>.

Base de dades de l'HGNC: <http://www.genenames.org/>.

Pel que fa a rates i ratolins, els gens s'escriuen amb la primera lletra majúscula i la resta en minúscula, i les proteïnes s'escriuen amb totes les lletres en majúscula. Per exemple, el gen és *Shh* i la proteïna, SHH.

Normes de l'MGI (o *Mouse Genome Informatics* del The Jackson Laboratory, EUA): <http://www.informatics.jax.org/mgihome/nomen/gene.shtml>.

Base de dades de l'MGI: <http://www.informatics.jax.org/>.

Per a les espècies següents no esmentarem cap detall de les normes, a causa de la seva complexitat, però oferim les adreces d'aquestes normes mateixes i de les bases de dades en línia.

Drosophila sp.

Normes de FlyBase (base de dades vinculada a les universitats nord-americanes de Harvard i Indiana i a la britànica de Cambridge): http://flybase.bio.indiana.edu/static_pages/docs/nomenclature/nomenclature3.html.

Base de dades de FlyBase: <http://flybase.bio.indiana.edu/>.

Arabidopsis thaliana

Normes de TAIR (o *The Arabidopsis Information Resource*, vinculada a la Carnegie Institution, EUA): <http://www.arabidopsis.org/portals/nomenclature/guidelines.jsp>.

Base de dades de TAIR: <http://www.arabidopsis.org/>.

Escherichia coli i altres bacteris

Normes al web de la Universitat de l'Estat de San Diego (SDSU, en les sigles angleses): <http://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/mutations/nomenclature-v3.pdf>.

Bases de dades del CGSC (*Coli Genetic Stock Center*, de la Universitat nord-americana de Yale): <http://cgsc.biology.yale.edu/index.php>.

Finalment, cal esmentar un grup d'eines en línia molt útils, que són les bases de dades conjuntes, que agrupen informació sobre diversos organismes model, i faciliten molt la feina de cerca de gens compartits, sobretot per a veure les analogies ortotipogràfiques entre els gens de les diverses espècies, i per a consultar en general molta informació que trobaríem dispersa entre diversos llocs. No cal dir que aquestes eines funcionen molt bé com a enllaç per a accedir a bases de dades més especialitzades. Un exemple n'és la base de dades Gene de l'NCBI (o *National Center for Biotechnology Information*, del govern dels EUA): <http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene>.

4. CONCLUSIÓ

El Web ha representat un canvi important pel que fa a la manera de consultar la normativa terminològica de les ciències de la vida, però també hem de veure que encara queda camí per recórrer. Per exemple, hauria de ser important que les darreres versions dels codis de taxonomia fossin també accessibles en línia. Així mateix, amb el pas del temps cal esperar més convergència en la nomenclatura de gens i proteïnes d'espècies diferents, la qual cosa es reflectirà en el camp dinàmic de les bases de dades en línia.