

## **I - Theoretical Study of CellDesigner:**

### **I-1 Introduction**

Systems biology is characterized by the synergistic integration of theory, computational modeling, and experiments. Although software infrastructure is one of the most crucial components of research in this field, there has been no common standard or infrastructure to enable integration of computational resources. To address this issue, several approaches based on standardized technologies have been developed, including CellDesigner, a process diagram editor designed for representing gene-regulatory and biochemical networks.[1]

### **I-2 General overview of CellDesigner:**

CellDesigner is a process diagram editor for drawing gene-regulatory and biochemical networks. Networks are drawn based on the process diagram, with [graphical notation](#) system proposed by Kitano, and are stored using the [Systems Biology Markup Language \(SBML\)](#), a standard for representing models of biochemical and gene-regulatory networks. Networks are able to link with simulation and other analysis packages through [Systems Biology Workbench \(SBW\)](#). [2]

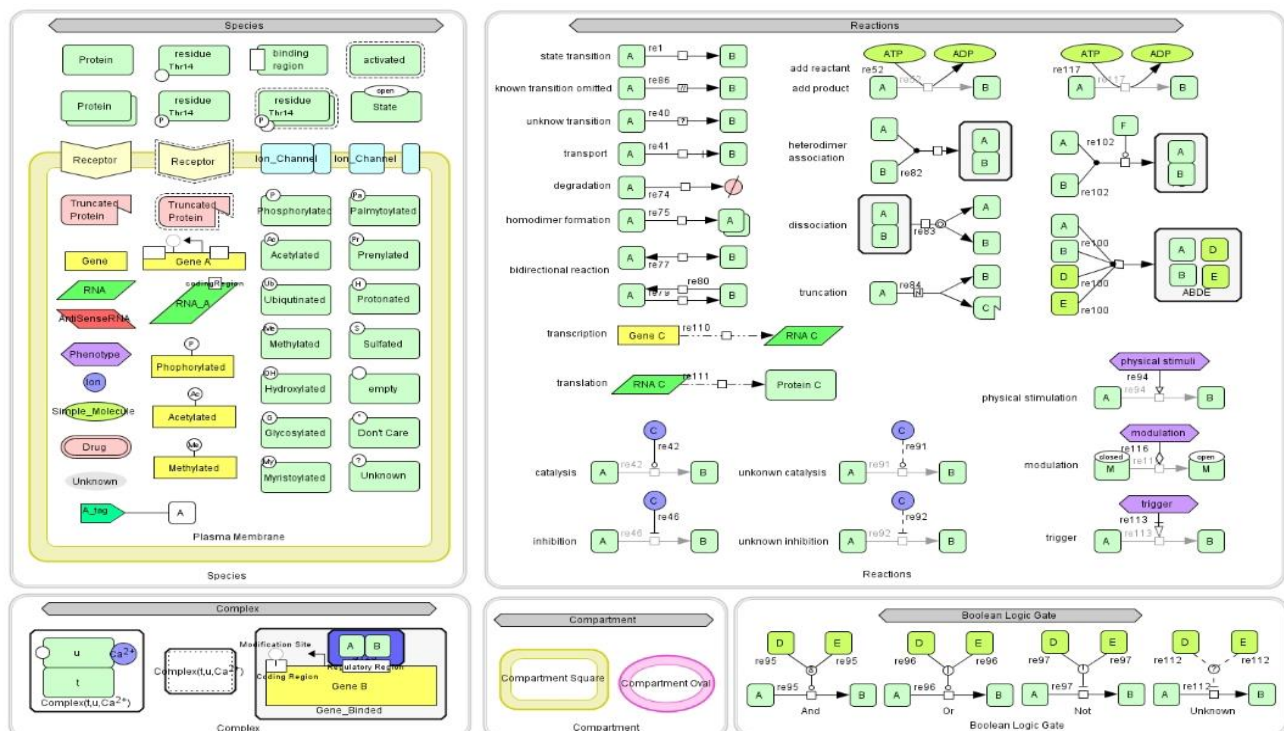


### **I- 3 Main Features: [3]**

- Representation of biochemical semantics
- Detailed description of state transition of protein
- SBML compliant (SBML Level-1, Level-2 document)
- Integration with SBW-enabled simulation modules
- Extreme portability as a Java application

#### I- 4 Technical Aspects: [4] :

- CellDesigner is a software for modeling biological networks, used to represent biochemical interactions and cellular signaling pathways.
- It uses the SBML standard to save models, allowing them to be shared with other software.
- The software also relies on the SBGN notation to ensure a clear and standardized graphical representation.



#### I- 5. Strengths [5] :

- CellDesigner is characterized by its compatibility with the SBML and SBGN standards, facilitating model sharing.
- It features a simple graphical interface, accessible to users without programming skills.
- Additionally, it allows models to be exported in various formats for use in scientific reports and presentations.

#### **I-6. Limitations and Weaknesses: [6]**

- 1-The layout information is stored in the CellDesigner's SBML extension.
- 2-CellDesigner format does not define the shape of the species or reaction symbols.
- 3-The rendering of the shapes corresponding to these features is hardcoded in CellDesigner software.
- 4-For this reason, the CellDesigner encoding does not support multiple layouts.

#### **I- 7. Conclusion :**

##### **In conclusion, we can state:**

CellDesigner is a powerful and essential tool for modeling biological networks and systems biology. It enables clear and standardized representation of biochemical interactions and gene-regulatory networks based on SBML and SBGN standards, and facilitates model sharing and integration with simulation modules for computational analysis and experiments.

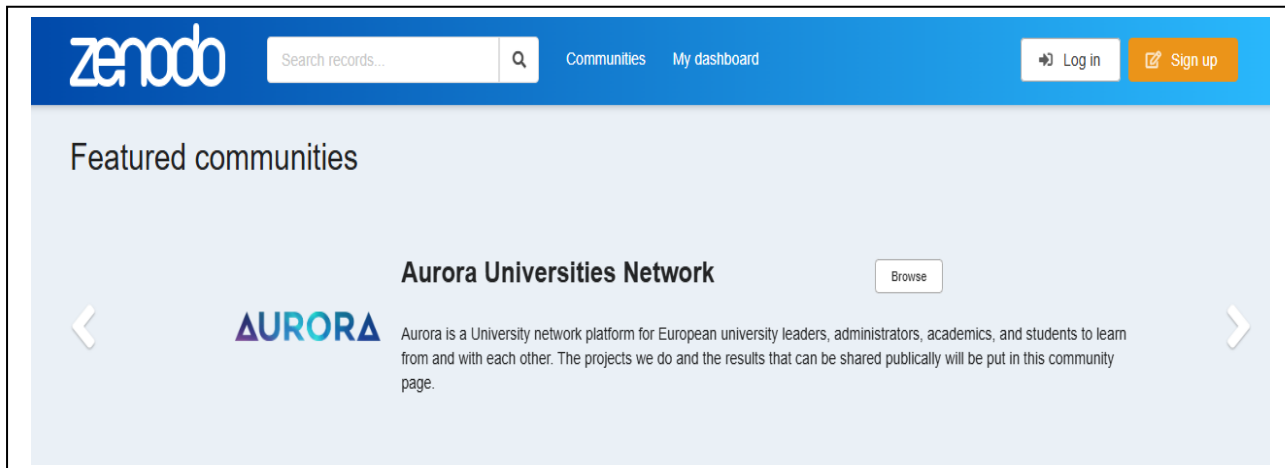
Its main advantages include ease of use without the need for programming skills, and the ability to export models in various formats suitable for scientific publications. However, some technical limitations, such as difficulty in customizing graphical layouts and lack of support for multiple layouts, may restrict the optimal visualization of complex networks.

In short, CellDesigner remains a reliable and accessible tool for systems biology research, with potential future developments to enhance graphical flexibility and interoperability with other platforms, making it a clear example of the importance of specialized software in understanding and simulating complex biological processes.

## **II- Practical Study: Exploration of Zenodo:**

### **II- 1 Presentation of Zenodo :**

Zenodo is an open science platform created by CERN and supported by the European Commission. It allows researchers to upload, share, and preserve various types of scientific outputs in a free and open manner.



**Figure 1** : The homepage of the Zenodo platform.

#### **II.1.1 Objectives of the platform :**

The main objective of Zenodo is to facilitate the sharing and dissemination of research data, ensure their long-term preservation, and promote open access to scientific results.

#### **II .1.2 Types of hosted content**

Zenodo hosts various types of scientific content, such as datasets, articles, theses, reports, software, presentations, and scientific posters.

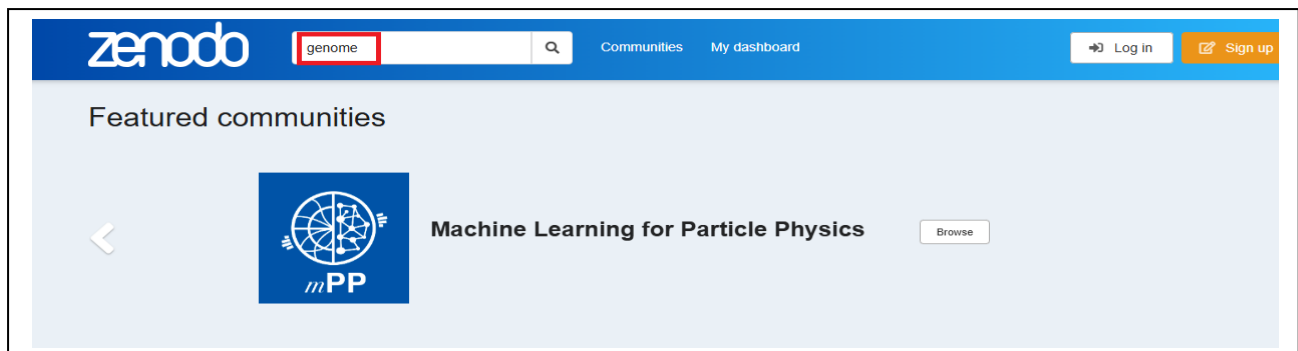
#### **II .1.3 Importance of Zenodo for Open Science and Life and Natural Sciences :**

Zenodo plays an important role in open science by providing free access to scientific data. It is particularly useful for the life and natural sciences, as it facilitates the sharing of biological, genomic, and environmental data among researchers worldwide.

## II .2. Description of the steps carried out :

### II .2.1 Search performed (query used) :

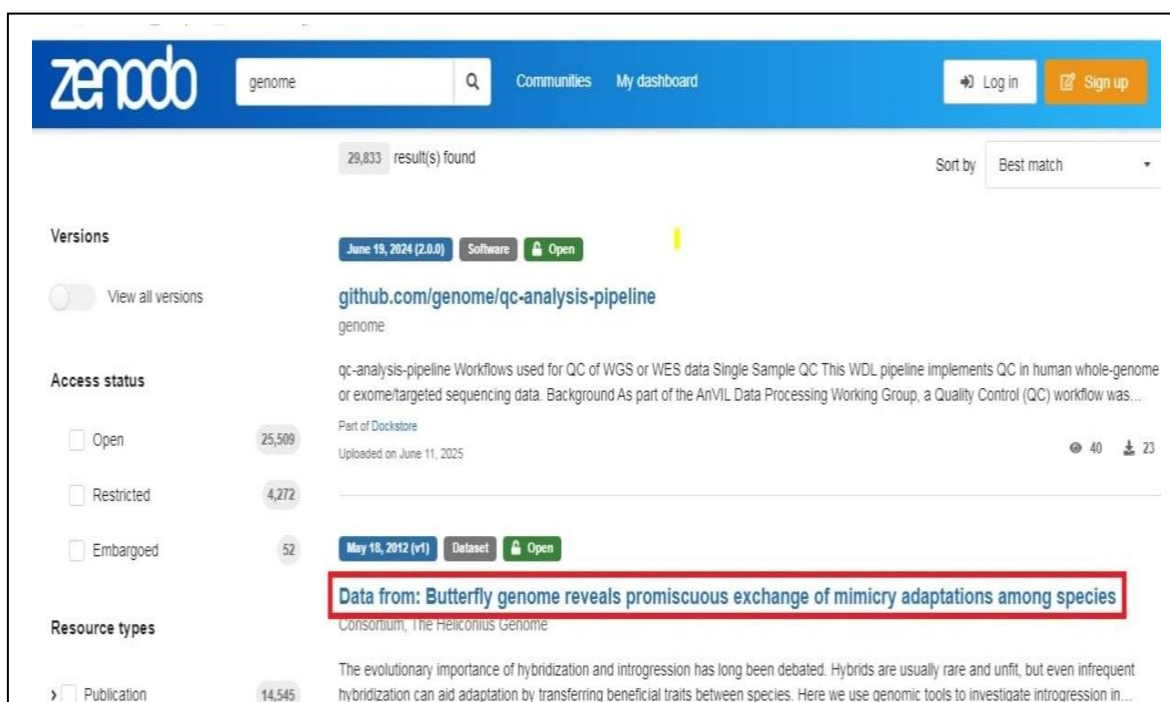
The search was conducted on the Zenodo platform using the keyword “genome” in the search bar to find datasets related to the life and natural sciences.



**Figure 2: Search results for the keyword “genome” on Zenodo.**

### II .2.2. Dataset selection criteria :

The dataset titled *Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species* was selected based on several criteria: its relevance to the field of biological sciences, the clarity of its description, the availability of metadata, and open access to the files.



**Figures 3 :** Page of the selected dataset on the Zenodo platform.

### II.2.3 Navigation on the platform :

Navigation on the Zenodo platform is simple and intuitive. It allows users to search for datasets, access their detailed descriptions, view the associated metadata, and download the available files.

Files (365.8 kB)		
Name	Size	Download all
heliconius_genome_scaffolding_scripts.tar.gz	28.8 kB	Download
md5:6b2048c51764ad85e045f0dfa1c4a78		
HGC_Nature_2012_CSP_OBP_OR_Gene_Predictions.xls	335.4 kB	Download
md5:b2092d62ed925eebbf081b7b16ab612e		
README_for_heliconius_genome_scaffolding_scripts.tar.txt	1.7 kB	Preview Download
md5:2042e713bd0686d28bb10ca6f9e32fc0		

Figure 4: Downloading the dataset files.

### Dataset download results :

	C	D	E	F	G	H	I	L	M	N	O	P	Q
1	GeneID	Exon	Scaffold ID	1-1	Chromosome	bp	CDS Sequence	AAs	Protein sequence	orientat	Top hit	tBlast	E-value
2	HMELO10992	2	HE671155.61017-62329		chromosome 1	366	ATGAAGTGTG	122	MKCVLLGLVFSI	reverse	AB24374	46%	#####
3	HMELO20916	2			N/A	380	ATGAAGTTTCG	126	MKFVILCVFALIAVA	forward	HQ4363	56%	#####
4	HMELO10990	4	HE671155.49149-57892		chromosome 1	375	ATGAAGACCG	125	MKTVLCLFALVAVA	forward	AB2601	75%	#####
5	HMELO18122	2	HE671155.68082-68941		chromosome 1	390	ATGGTTTCCA	130	MVSKTFVVLCCFVA	reverse	AB26012	60%	#####
6	HMELO10991	2			chromosome 1	348	TGTGCTATCA	116	CAITVNARPAEHTT	forward	AY1015	60%	#####
7	HMELO03114	2	HE688707.5532-7154		chromosome 1	360	ATGAAATTCTT	120	MKFLIVLSCLVLSAF	reverse	AJ97346	99%	#####
8	HMELO03116	2	HE688707.13464-17708		chromosome 1	360	ATGAAGTGCC	120	MKCLLVLSGLLVLA	reverse	AB24375	63%	#####
9	HMELO21490	2			N/A	372	ATGAAGACTA	124	MKTILVLATVLA	reverse	AB26012	63%	#####
10	HMELO22515	1	HE683854.10191-10400.H		chromosome 1	205	>HMELO07604-	68	MKVIVIMLCLCVTVL	reverse	AJ97345	73%	#####
11	HMELO20809	1			N/A	381	ATGAAGTCCG	127	MKSIVLVALMCLVA	forward	AJ97346	96%	#####
12	HMELO20810	2			N/A	360	ATGAAGATTTT	120	MKIFIAFIALPSVLT	reverse	NM_001	63%	#####
13	HMELO10988	2	HE671155.32518-38297		chromosome 1	315	ATGCACATCA	105	MHIQQSVLILLVYH	reverse	EF18675	78%	#####
14	HMELO20812	2			N/A	366	ATGAAGATTA	122	MKIILLFCAIGIALVF	reverse	AJ97345	60%	#####
15	HMELO20574	2			N/A	369	ATGAAGATTG	123	MKIVIVLALIAVVAAR	reverse	AB24375	54%	#####
16	HMELO03115	2	HE688707.10377-11579		chromosome 1	366	ATGAGAAAAG	122	MRKVSIIILLISCVLT	reverse	NM_001	74%	#####
17	HMELO22528	2	HE683454.31134-31853		chromosome 4	276	TGTATGGCGC	92	CMAQTQRPEVSDT	reverse	AB26012	91%	#####
18	HMELO20813	2			N/A	381	ATGAATAACTC	127	MNNWLLYLCAITVA	reverse	AB43071	79%	#####
19	HMELO03113	2	HE688707.2491-4220		chromosome 1	369	ATGCTGGTGC	123	MLVPYIMILLVTVA	forward	AB43071	46%	#####
20	HMELO05594	2	HE671039.1046-1800		chromosome 1	372	ATGAAGACAA	124	MKTILVLAATLVA	forward	AB26012	62%	#####
21	HMELO20491	2			N/A	369	ATGCAACTCC	123	MQLLAAATLVASL	forward	AB26012	57%	#####
22	HMELO22550	2	HE683454.37610-39283		chromosome 4	288	ATTATAATCTC	96	IIISICVTVMAAPQMS	reverse	AB26012	56%	#####
23	HMELO22572	2	HE683825.2800-3636		N/A	390	GTGGCTAAAA	130	VAKKSIVVICWFLAS	reverse	DQ2856	71%	#####
24	HMELO18123	2	HE671155.71535-73032		chromosome 1	390	ATGGTTGCTA	130	MVAKSIVLVCCFLA	reverse	DQ2856	69%	#####
25	HMELO18110	2	HE683825.7982-8802		N/A	390	ATGGCTTCTA	130	MASKLIVLMCCFVA	reverse	DQ2856	62%	#####
26	HMELO10987	2	HE671155.30506-31629			441	ATGGCCGTAC	147	MAVLIFTILFISSISS	forward	EF20282	28%	0.43
27	HMELO18109	2	HE683825.4617-5419		N/A	390	ATGGCTTCTA	130	MASKLIVLMCCFVA	reverse	DQ2856	63%	#####





Figure 5: Preview of the downloaded file containing the genomic

### II. 3 Dataset metadata :

#### Metadata table (Dublin Core):

<u>Element (Dublin core)</u>	<u>Value</u>
<u>Title</u>	Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species
<u>Creator</u>	The Heliconius Genome Consortium
<u>Data</u>	18/05/2012
<u>Subject</u>	genome, butterfly, Heliconius melpomene
<u>Description</u>	Dataset containing detailed genomic data on the butterfly <i>Heliconius melpomene</i> , including tables related to the CSP, OBP, and OR genes. These data are used for studying evolution, mimicry, and genetic mechanisms in this species.
<u>Publisher</u>	<b>Zenodo</b>
<u>Type</u>	Dataset
<u>Format</u>	<b>XLS</b>
<u>Identifier</u>	DOI: 10.5061/dryad.m27qq
<u>Rights</u>	Creative Commons Zero v1.0



<p><b>Details</b></p> <p><b>DOI</b>  <div>DOI 10.5061/dryad.m27qq</div> </p> <p><b>Resource type</b>  Dataset</p> <p><b>Publisher</b>  Zenodo</p>	<p><b>Citation</b></p> <p>Consortium, T. H. G. (2012). Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species [Data set]. Zenodo. <a href="https://doi.org/10.5061/dryad.m27qq">https://doi.org/10.5061/dryad.m27qq</a></p> <p>Style <div>APA</div> </p>
<p><b>Rights</b></p> <p><b>License</b></p> <div>  Creative Commons Zero v1.0 Universal </div>	<p><b>Export</b></p> <div> <div>JSON</div> <div>Export</div>  </div> <p><b>Technical metadata</b>  Created June 16, 2021</p>