

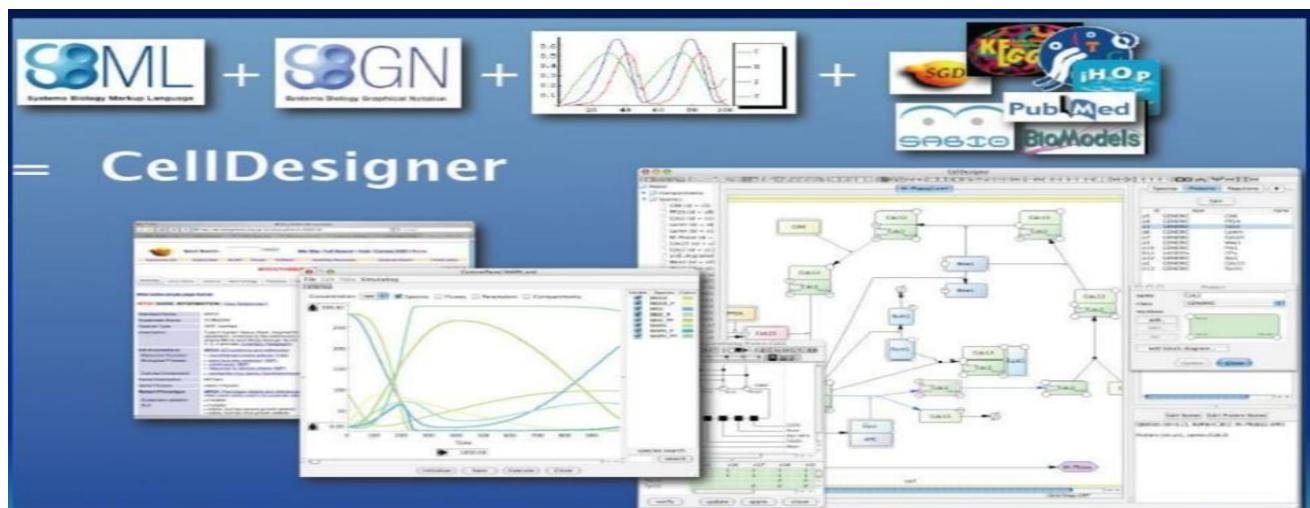
## **I - Theoretical Study of CellDesigner:**

### **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]

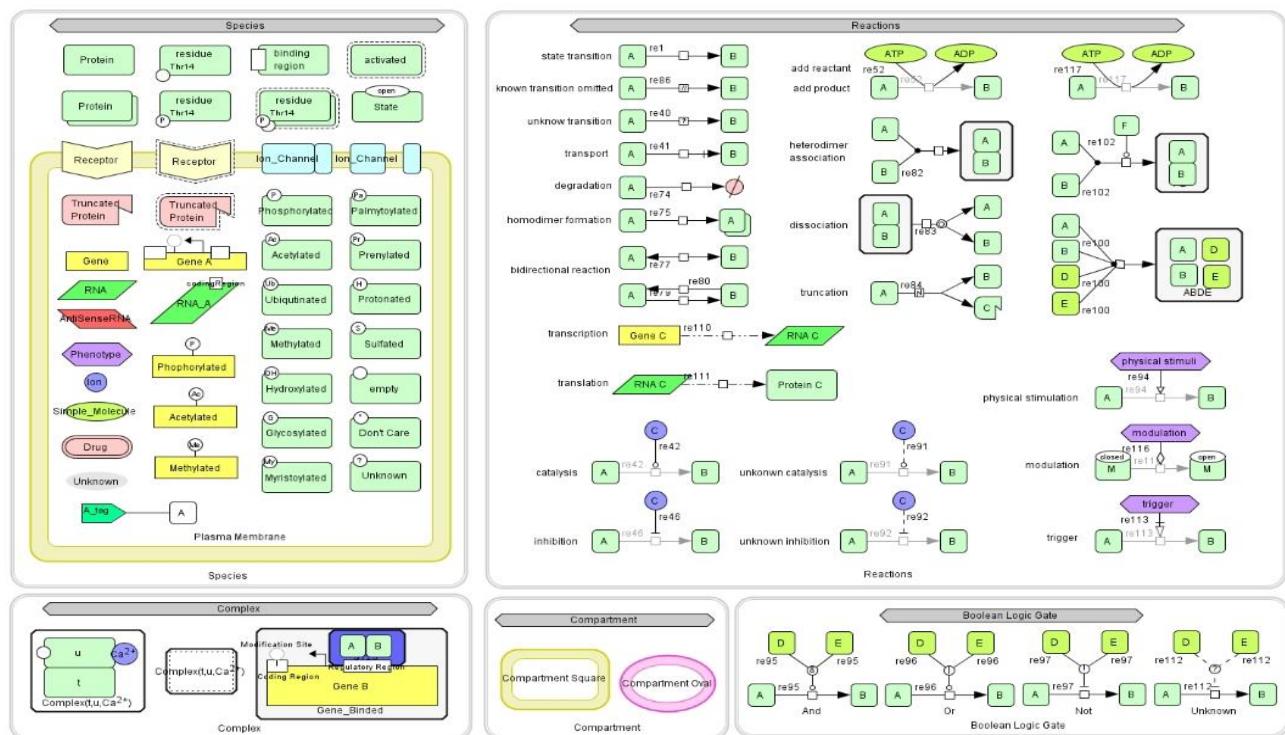


### **I-3 Main Features: [3]**

- 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.

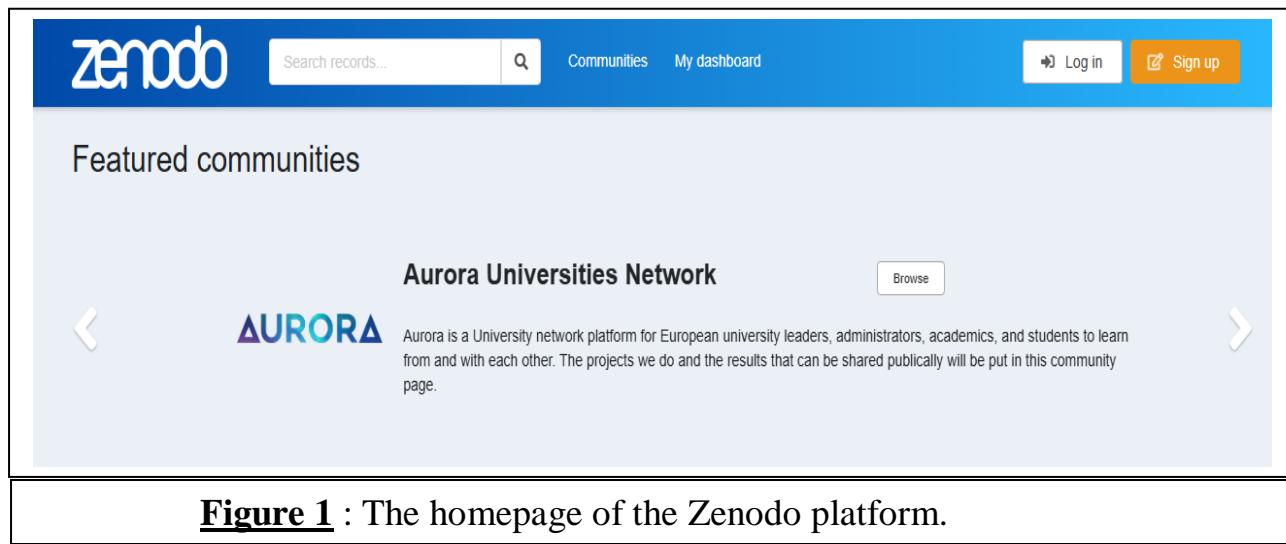
Its main advantages include ease of use without the need for programming skills, and the ability to export models in various formats suitable for scientific publications. However, some technical limitations, such as difficulty in customizing graphical layouts and lack of support for multiple layouts,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

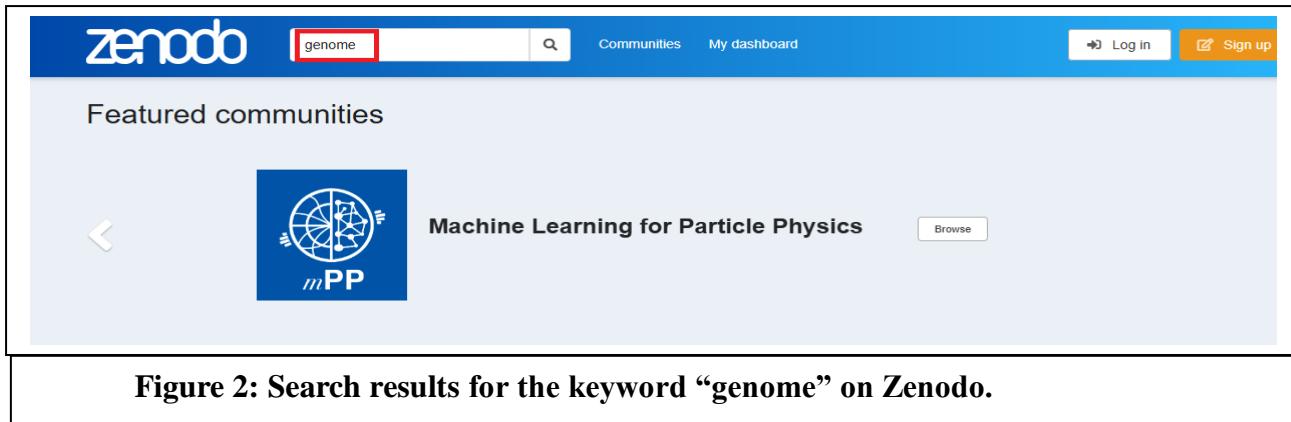
#### **II .1.3 Importance of Zenodo for Open Science and Life and Natural Sciences :**

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.



### **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.

A screenshot of the Zenodo dataset page for the selected dataset. The search bar at the top shows "genome". The main search results summary is "29,833 result(s) found". The dataset "Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species" is highlighted with a red border. This dataset has a version "June 19, 2024 (2.0.0)" and is categorized as "Software". It has 25,509 files and is marked as "Open". The dataset description states: "qc-analysis-pipeline Workflows used for QC of WGS or WES data Single Sample QC This WDL pipeline implements QC in human whole-genome or exome/targeted sequencing data. Background As part of the AnVIL Data Processing Working Group, a Quality Control (QC) workflow was...". The dataset was uploaded on June 11, 2025, and has 40 citations and 23 downloads. The dataset is associated with "github.com/genome/qc-analysis-pipeline" and "genome". The dataset is also linked to "qc-analysis-pipeline". The dataset is part of the "Consortium, The Heliconius Genome". The dataset is described as: "The evolutionary importance of hybridization and introgression has long been debated. Hybrids are usually rare and unfit, but even infrequent hybridization can aid adaptation by transferring beneficial traits between species. Here we use genomic tools to investigate introgression in...". Other sections visible include "Versions", "Access status", "Resource types", and "Publication".

**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.

The screenshot shows a list of files for download from a Zenodo dataset. The files listed are:

- heliconius\_genome\_scaffolding\_scripts.tar.gz** (28.8 kB) - Download link
- HGC\_Nature\_2012\_CSP\_OBP\_OR\_Gene\_Predictions.xls** (335.4 kB) - Download link
- README\_for\_heliconius\_genome\_scaffolding\_scripts.tar.txt** (1.7 kB) - Preview and Download links

**Figure 4: Downloading the dataset files.**

### **Dataset download results :**

C	D	E	Scaffold ID	1-1	F	G	H	I	L	M	N	O	P	Q
1	GeneID				Chromosome	bp	CDS Sequence	AAs	Protein sequence	orientat	Top hit	tBlastn	E-value	
2	HMEL010992	2	<a href="#">HE671155.61017-62329</a>		chromosome 1	366	ATGAAGTGTG	122	MKCVLLLGFVFSI	/ reverse	AB2437	46%	#####	
3	HMEL020916	2			N/A		380	ATGAAGTTCG	126	MKFVLCLVFALIAVA	/ forward	HQ4363	56%	#####
4	HMEL010990	4	<a href="#">HE671155.43149-57892</a>		chromosome 1	375	ATGAAGACCG	125	MKTVLCLFALVAVA	/ forward	AB2601	75%	#####	
5	HMEL018122	2	<a href="#">HE671155.68082-68341</a>		chromosome 1	390	ATGGTTCCA	130	MVKSTFVVLCCFVA	/ reverse	AB26012	60%	#####	
6	HMEL010991	2			chromosome 1	348	TGTGCTATCA	116	CAITVNARPAEHYT	/ forward	AY1015	60%	#####	
7	HMEL003114	2	<a href="#">HE668707.5592-7154</a>		chromosome 1	360	ATGAAATTCTT	120	MKFVLVLSCLVLSA	/ reverse	AJ97346	99%	#####	
8	HMEL003116	2	<a href="#">HE668707.13464-17708</a>		chromosome 1	360	ATGAAGTGCC	120	MKCLLVLSGLLVLA	/ reverse	AB2437	63%	#####	
9	HMEL021490	2			N/A		372	ATGAAGACTA	124	MKTILVLATLVAAAV	/ reverse	AB26012	63%	#####
10	HMEL022515,	1	<a href="#">HE669854.10191-10400-H</a>		chromosome 1	205	>HMEL007604-	68	MKVIVIMLCLCVTVL	/ reverse	AJ97345	73%	#####	
11	HMEL020809	1			N/A		381	ATGAAGTCG	127	MKSVLVALMCLVA	/ forward	AJ97346	96%	#####
12	HMEL020810	2			N/A		360	ATGAAGATT	120	MKIFIIFIAPSVLT	/ reverse	NM_0011	63%	#####
13	HMEL010988	2	<a href="#">HE671155.32518-38297</a>		chromosome 1	315	ATGACACATCA	105	MHIIQQSVLILLVY	/ reverse	EF18679	78%	#####	
14	HMEL020812	2			N/A		366	ATGAAGATTA	122	MKIIIILFCIA/GIALVF	/ reverse	AJ97345	60%	#####
15	HMEL020574	2			N/A		369	ATGAAGATTG	123	MKVIV/LALIAVVAAR	/ reverse	AB2437	54%	#####
16	HMEL003115	2	<a href="#">HE668707.10377-11573</a>		chromosome 1	366	ATGAGAAAAG	122	MRKVSIIILLSCVLT	/ reverse	NM_0011	74%	#####	
17	HMEL022528	2	<a href="#">HE669454.31134-31853</a>		chromosome 4	276	TGTATGGCGC	92	CMAQTQRPEVSDT	/ reverse	AB26012	91%	#####	
18	HMEL020813	2			N/A		381	ATGATAAACTC	127	MNNWLLYLCAUTV	/ reverse	AB43077	79%	#####
19	HMEL003113	2	<a href="#">HE668707.3491-4220</a>		chromosome 1	369	ATGCTGGTGC	123	MLVPY MILLTVTVA	/ forward	AB43077	46%	#####	
20	HMEL005594	2	<a href="#">HE670139.1046-11800</a>		chromosome 1	372	ATGAAGACAA	124	MKTILVLAAATLVAL	/ forward	AB26012	62%	#####	
21	HMEL020491	2			N/A		369	ATGCAACTCC	123	MQLLLAAATLVASL	/ forward	AB26012	57%	#####
22	HMEL022550	2	<a href="#">HE669454.37610-39283</a>		chromosome 4	288	ATTATAATCTC	96	IIISICTVMAAPQMS	/ reverse	AB26012	56%	#####	
23	HMEL022572	2	<a href="#">HE669825.2800-3636</a>		N/A		390	GTGGCTAAA	130	VAKKSIVVICWFLAS	/ reverse	DQ2856	71%	#####
24	HMEL018123	2	<a href="#">HE671155.71595-73032</a>		chromosome 1	390	ATGGTTGCTA	130	MVKASIVLVCCFLA	/ reverse	DQ2856	69%	#####	
25	HMEL018110	2	<a href="#">HE669825.7982-8802</a>		N/A		390	ATGGCTTCTA	130	MASKLIVLMCCFVA	/ reverse	DQ2856	62%	#####
26	HMEL010987	2	<a href="#">HE671155.30506-31629</a>				441	ATGCCGTCAC	147	MAVLIFTLFISSISS	/ forward	EF2028	28%	0.43
27	HMEL018109	2	<a href="#">HE669825.4617-5419</a>		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):**

<b><u>Element (Dublin core)</u></b>	<b><u>Value</u></b>
<b><u>Title</u></b>	Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species
<b><u>Creator</u></b>	The Heliconius Genome Consortium
<b><u>Date</u></b>	18/05/2012
<b><u>Subject</u></b>	genome, butterfly, <i>Heliconius melpomene</i>
<b><u>Description</u></b>	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.
<b><u>Publisher</u></b>	Zenodo
<b><u>Type</u></b>	Dataset
<b><u>Format</u></b>	XLS
<b><u>Identifier</u></b>	DOI: 10.5061/dryad.m27qq
<b><u>Rights</u></b>	Creative Commons Zero v1.0

**Details**

**DOI**  
DOI [10.5061/dryad.m27qq](https://doi.org/10.5061/dryad.m27qq)

**Resource type**  
Dataset

**Publisher**  
Zenodo

**Rights**

**License**

 Creative Commons Zero v1.0  
Universal

**Citation**

Consortium, T. H. G. (2012). Data from: Butterfly genome reveals promiscuous exchange of mimicry adaptations among species [Data set]. Zenodo. <https://doi.org/10.5061/dryad.m27qq>

Style [APA](#) 

**Export**

JSON  Export 

Technical metadata  
Created June 16, 2021