

The SysNDD Documentation

Bernt Popp, Christiane Zweier

2022-06-08

Contents

Preface	2
History of SysID and SysNDD	2
The SysNDD concept	2
Acknowledgments	3
1 Curating gene-disease relationships	3
1.1 Neurodevelopmental disorders	3
1.2 Genetic heterogeneity	3
1.3 Expert curation	4
1.4 Technical concepts	4
1.5 Outlook	4
2 Web tool	5
2.1 Landing page	5
2.2 Main navigation	6
2.3 Footer navigation	6
2.4 Table views	7
2.5 Analyses views	10
2.6 About page	10
2.7 Login page	10
3 API	11
4 Database structure	15
5 Curation criteria	15
5.1 Definitions	16
5.2 NDD Definitive entities	16
5.3 NDD Moderate and Limited entities	17

6 Re-review instructions	19
6.1 Re-review tool usage	19
6.2 Re-review curation	24

Preface

This documentation is intended to describe the SysNDD¹ project and provide instructions for regular users to show how to use the tool and for curator status users how to perform reviews and how to enter data.

History of SysID and SysNDD

SysNDD is based on its predecessor SysID² which had been published in 2016 (Kochinke et al., 2016). Christiane Zweier has been involved in establishing and updating SysID from its start in 2009. She has since performed and coordinated curation and regular updates.

The PHP based SysID web tool (Yii 2 framework) was however not further developed and maintained besides necessary bugfixes. After the maintenance agreement for the original server at the CMBI at Radboud University in Nijmegen ran out, the installation was moved to a virtual server at the Department for BioMedical Research (DBMR) at the University Bern. The former link from the initial publication is re-directed so it still works. The legacy code base was updated to allow installation and security fixes and to be uploaded to a GitHub repository (SysID)³.

In 2019 the chance arose to integrate the SysID curation effort with the Orphanet resource, supported by ERN ITHACA. In the process of aligning the curation and naming conventions for genes, diseases and phenotypes we decided to redesign the database and web tool.

The SysNDD concept

SysNDD contains a manually curated catalogue of published gene-disease-associations implicated in neurodevelopmental disorders (NDD).

To allow interoperability and mapping between gene-, phenotype- or disease-oriented databases, we center our approach around curated gene-inheritance-disease units, so called entities. These entities are classified into different confidence status (categories: “Definitive”, “Moderate”, “Limited”, “Refuted”) according to the degree of underlying scientific evidence. Furthermore, manually curated information on associated phenotypes is provided.

The entries in SysNDD are currently updated every 3-4 months and can be utilized for a broad spectrum of tasks from both research and diagnostics.

One of our goals is to incorporate the SysNDD data⁴ into other gene/ disease-relationship databases like the Orphanet ontology (first results: id-genes.orphanet.app⁵).

Bernt Popp (scientist at the Institute of Human Genetics at the University Hospital Leipzig, Germany) developed and programmed the SysNDD tool and will be integrating further functionality including variants associated with entities in future updates.

¹<https://sysnnd.dbmr.unibe.ch/>

²<https://www.sysid.dbmr.unibe.ch/>

³<https://github.com/berntpopp/SysID>

⁴<https://sysnnd.dbmr.unibe.ch/>

⁵<https://id-genes.orphanet.app/ithaca/>

Acknowledgments

The current SysNDD database development is supported by:

- DFG (Deutsche Forschungsgemeinschaft) grant PO2366/2-1 to Bernt Popp⁶.
- DFG (Deutsche Forschungsgemeinschaft) grant ZW184/6-1 to Christiane Zweier⁷.
- ERN ITHACA⁸ through Alain Verloes⁹.

The previous SysID database and data curation was supported by:

- The European Union's FP7 large scale integrated network GenCoDys (HEALTH-241995) to MA Huynen and Annette Schenck.
- VIDI and TOP grants (917-96-346, 912-12-109) from The Netherlands Organisation for Scientific Research (NWO) to Annette Schenck.
- DFG (Deutsche Forschungsgemeinschaft) grants ZW184/1-1 and -2 to Christiane Zweier¹⁰.
- the IZKF (Interdisziplinäres Zentrum für Klinische Forschung) Erlangen to Christiane Zweier¹¹.
- ZonMw grant (NWO, 907-00-365) to Tjitske Kleefstra.

1 Curating gene-disease relationships

As the name implies a rare disease affects only very few individuals. However, there are many unique causes of rare diseases, thus many individuals are affected by such a disease. Due to the rarity of each single entity, effective management, surveillance and treatment is challenging. So is finding the correct diagnosis, which is often described as the “diagnostic odyssey”.

Rare diseases often have a genetic cause, making high-throughput sequencing (next-generations sequencing; NGS) a central part of finding the molecular diagnosis.

1.1 Neurodevelopmental disorders

Neurodevelopmental disorders (NDD) affect about 2% of children. They represent a clinically and genetically extremely heterogeneous disease group comprising amongst other developmental delay (DD), intellectual disability (ID) and autism spectrum disorder (ASD) and developmental and epileptic encephalopathies (DEE).

1.2 Genetic heterogeneity

The huge genetic heterogeneity is evident when looking at the published gene-disease associations over time.

Thus the question arises:

How can we keep track of this fast development and have the information at hand when we need it in the clinic or when analyzing sequencing data?

⁶<https://orcid.org/0000-0002-3679-1081>

⁷<https://orcid.org/0000-0001-8002-2020>

⁸<https://ern-ithaca.eu/>

⁹<https://orcid.org/0000-0003-4819-0264>

¹⁰<https://orcid.org/0000-0001-8002-2020>

¹¹<https://orcid.org/0000-0001-8002-2020>

While the answer to this question is easy:

We need curated databases to catalogue and summarize the wealth of published information.

The task at hand is not only laborious but also requires expertise and consistency.

1.3 Expert curation

In our opinion, the curation of gene-disease relationships in rare disease such as NDDs requires clinical and scientific proficiency in the respective field. This implies that clinician scientists involved in counseling, diagnostics, and research of NDD are predestined for this task.

To reduce workload and dependence on single experts, a distributed effort in larger consortia and collaboration between different work groups is needed.

In the course of updating SysID we had the great chance to contribute our data to Orphanet to create a European ID/NDD specific reference list. With support from the „ITHACA Workgroup: intellectual disability“ ([id-genes.orphanet.app¹²](https://id-genes.orphanet.app/ithaca%5D)) in 2019 we started working with the Orphanet team which is part of the Gene Curation Coalition (GenCC).

Additionally, we are able to recruit expert curators from ERN ITHACA¹³ to contribute to re-curation of old data and updating new data in SysNDD.

1.4 Technical concepts

In addition to a pool of experts, the right tools are needed.

We defined “gene-inheritance-disease” units as “entities” which represent the central curation effort. The components of these entities are normalized using widely used and standardized ontology terms (e.g. HGNC identifier for genes, OMIM or MONDO for disease and inheritance from HPO). This allows interoperability and linking to other data sources.

Based on this concept we developed a new database scheme, which allows entities to be systematically and reproducibly cataloged. The database is abstracted into a JSON API, which allows structured programmatic access to the underlying data.

Finally, the API feeds the web tool which can be used to easily search, filter, download and visualize the database contents in modern webbrowsers.

1.5 Outlook

- The SysNDD database will improve the understanding and curation of rare NDD entities.
- SysNDD will enable systems biology and network analyses.
- Our long-term goal is incorporation of the high-quality, manually curated SysNDD data into European and international gene disease relationship databases,
- thus, improving diagnostics and care for individuals with rare NDDs.

¹²<https://id-genes.orphanet.app/ithaca%5D>

¹³<https://ern-ithaca.eu/>

2 Web tool

The SysNDD web tool is available from <https://sysnnd.dbmr.unibe.ch/> on a server hosted at the Department for BioMedical Research (DBMR¹⁴) (University of Bern) and the web address <https://sysnnd.org/> redirects to this server.

The web tool uses the Vue.js¹⁵ (v2.6) JavaScript framework with BootstrapVue to generate a Bootstrap v4 website frontend.

2.1 Landing page

The landing page is designed as simple Bootstrap v4 website with:

- 1) a navigation menu at the top,
- 2) the main site content, which changes with navigation to other routes, and
- 3) a footer navigation bar at the bottom

Screenshot of the landing page with elements marked:

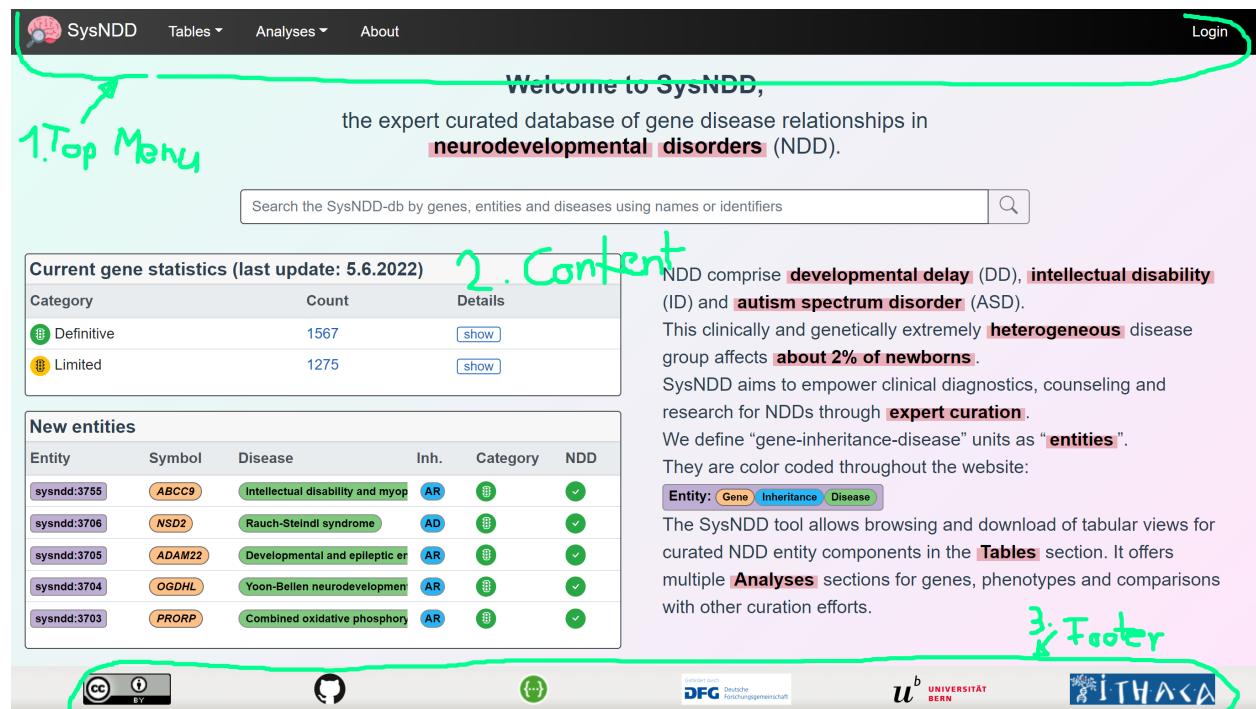


Figure 1: Landing page

The landing page content displays different elements to give a quick overview and allow fast navigation:

- a centered search input at the top,

¹⁴<https://www.dbmr.unibe.ch/>

¹⁵<https://vuejs.org/>

- a box (left side top) with current gene statistics divided by association category and inheritance patterns (Details),
- a box (left side bottom) showing a table of the 5 last entities entered into the database,
- an explanatory text on the right.

2.2 Main navigation

The main navigation allows quick access to all subpages.

The “Tables” button triggers a dropdown menu with links to: - “Entities” table view - “Genes” table view - “Phenotypes” table view - “Panels table” view

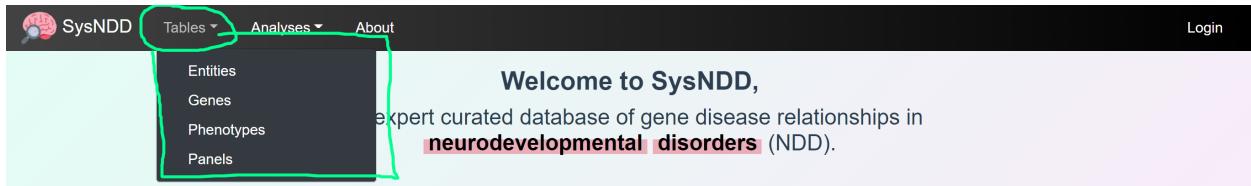


Figure 2: Navigation menu tables

The “Analyses” button triggers a dropdown menu with links to: - “Compare curations” view - “Correlate phenotypes” view - “Entries over time” view - “NDD Publications” view - “Gene networks” view

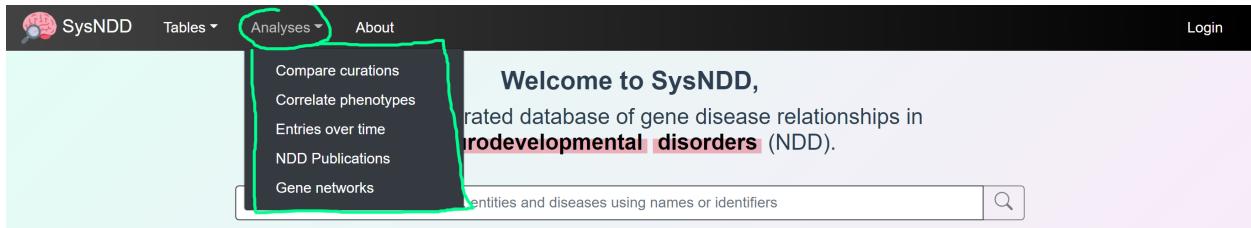


Figure 3: Navigation menu analyses

If not on the landing page the navigation bar also contains a “Search” button which can show a search field on any page.



Figure 4: Navigation menu search

If not logged in the right side of the menu shows a button which directs to the Login page. When logged in as a registered user the menu shows your username and additional links to page views depending on your user rights:

2.3 Footer navigation

The footer navigation shows pictures/ logos with links to:

- 1) the license applied to SysNDD

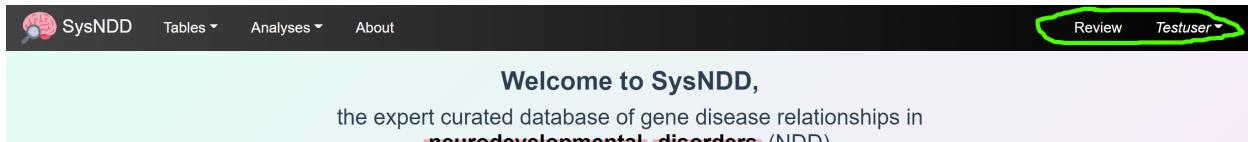


Figure 5: Navigation menu login

- 2) our GitHub repository
- 3) the SysNDD API view
- 4) the DFG funder website
- 5) the website of the University of Bern hosting our server
- 6) the ERN-ITHACA website



Figure 6: Footer navigation

2.4 Table views

We provide tabular representations with search, filtering, sorting and pagination functionality for different aspects of the entity concept.

2.4.1 Entities table

The Entities table is intended to provide an overview centered on the entity concept.

2.4.2 Genes table

The Genes table is intended to provide gene centered overview.

2.4.3 Phenotypes table

The Phenotypes table provides the possibility to filter for phenotype combinations annotated to the entities. The “AND/ OR” switch allows the user to change the logic how phenotype combinations are requested:

- AND: only entities having all selected phenotypes annotated are shown.
- OR: all entities having any of the selected phenotypes annotated are shown.

2.4.4 Panels table

The Panels table is intended for users to be able to create lists of NDD-associated genes. Additionally, the columns in the lists can be configured. Finally, the configuration can be downloaded as Excel file with information on the exact query in the meta sheet and the requested information in the “data” sheet. These files can then be used as “virtual panels” to filter genetic variants derived from high-throughput sequencing in external analysis tools.

SysNDD Tables ▾ Analyses ▾ About Search ▾ Login

Entities table Entities: 3629

Search any field by typing here Per page 10

Entity id	Symbol	Disease ontology na...	Hpo mode of inherit...	Category	Ndd phenotype word	Details
.. Entity id Symbol Disease ontolog...	.. Hpo mode of in...	.. Category Ndd phenotype w ..	
sysndd:1	ABCC9	Hypertrichotic osteochondro...	AD	█	✓	Show
sysndd:2	ABCC9	Cardiomyopathy, dilated, 10	AD	█	✗	Show
sysndd:3	ABCC9	?Atrial fibrillation, familial, 12	AD	█	✗	Show
sysndd:4	ABCD1	Adrenoleukodystrophy	XR	█	✓	Show
sysndd:5	ABCD4	Methylmalonic aciduria and h...	AR	█	✓	Show
sysndd:6	ABHD5	Chanarin-Dorfman syndrome	AR	█	✓	Show
sysndd:7	ACAD9	Mitochondrial complex I defic...	AR	█	✓	Show
sysndd:8	ACO2	Infantile cerebellar-retinal deg...	AR	█	✓	Show
sysndd:9	ACOX1	Peroxisomal acyl-CoA oxidase	AR	█	✓	Show
sysndd:10	ACSF3	Combined malonic and methylmalon...	AR	█	✓	Show



Figure 7: Entities view

SysNDD Tables ▾ Analyses ▾ About Search ▾ Login

Genes table Genes: 2842

any field by typing here Per page 10

Symbol	Category	Hpo mode of inherit...	Ndd phenotype word	Entities count	Details
.. Symbol Category Hpo mode of inherit....	.. Ndd phenotype word Entities count ..	
A2ML1	█	AD	✓	1	Show
AAAS	█	AR	✓	1	Show
AACS	█	AR	✓	1	Show
AARS1	█ █	AR AD	✓ ✗	2	Show
AARS2	█ █	AR AR	✓ ✓	2	Show
AASS	█	AR	✓	1	Show
ABAT	█	AR	✓	1	Show
ABCAT3	█	AD	✓	1	Show
ABCAT2	█	AR	✓	1	Show
ABCA5	█	AR	✓	1	Show



Figure 8: Genes view

SysNDD Tables ▾ Analyses ▾ About Search ▾ Login

Phenotype search [Associated entities: 3185]

Intellectual disability AND Per page 10

Entity id	Symbol	Disease ontology na...	Hpo mode of inherit...	Category	Ndd phenotype word	Details
... Entity id Symbol Disease ontolog...	... Hpo mode of in...	... Category Ndd phenotype w ...	
sysndd:1	ABCC9	Hypertrichotic osteochondro...	AD	...	✓	Show
sysndd:4	ABCD1	Adrenoleukodystrophy	XR	...	✓	Show
sysndd:5	ABCD4	Methylmalonic aciduria and h...	AR	...	✓	Show
sysndd:6	ABHD5	Chanarin-Dorfman syndrome	AR	...	✓	Show
sysndd:7	ACAD9	Mitochondrial complex I defic...	AR	...	✓	Show
sysndd:8	ACO2	Infantile cerebellar-retinal deg...	AR	...	✓	Show
sysndd:9	ACOX1	Peroxisomal acyl-CoA oxidase	AR	...	✓	Show
sysndd:10	ACSF3	Combined malonic and methy...	AR	...	✓	Show
sysndd:11	ACSL4	Intellectual developmental dis...	XR	...	✓	Show
sysndd:12	ACTB	Bartsier-Winter syndrome 1	AD	...	✓	Show



Figure 9: Phenotypes view

SysNDD Tables ▾ Analyses ▾ About Search ▾ Login

Panel compilation and download [Genes: 2842]

Category All Inheritance All Sort symbol

Columns category inheritance symbol

Per page 10

Category	Inheritance	Symbol	Hgnc id	Entrez id	Ensembl gene id	Ucsc id	Bed hg19	Bed hg38
Limited	Autosomal domi...	A2ML1	HGNC:23336	144568	ENSG00000166...	uc001quz.6	chr12:8975068...	chr12:8822621...
Limited	Autosomal reces...	AAAS	HGNC:13666	8086	ENSG00000094...	uc001scr.5	chr12:53701240...	chr12:53307456...
Limited	Autosomal reces...	AACS	HGNC:21298	65985	ENSG00000081...	uc001uhc.4	chr12:12554992...	chr12:12506543...
Definitive	Autosomal reces...	AARS1	HGNC:20	16	ENSG00000090...	uc002eyn.2	chr16:70286198...	chr16:70251983...
Limited	Autosomal reces...	AARS2	HGNC:21022	57505	ENSG00000124...	uc010jza.2	chr6:44267391...	chr6:44298731...
Limited	Autosomal reces...	AASS	HGNC:17366	10157	ENSG00000008...	uc003vkb.4	chr7:121715701...	chr7:122064583...
Limited	Autosomal reces...	ABAT	HGNC:23	18	ENSG00000183...	uc002czc.5	chr16:8768422...	chr16:8674596...
Limited	Autosomal domi...	ABCA13	HGNC:14638	154664	ENSG00000179...	uc003toq.2	chr7:48211055...	chr7:48171458...
Definitive	Autosomal reces...	ABCA2	HGNC:32	20	ENSG00000107...	uc064xhf.1	chr9:139901686...	chr9:137007234...
Limited	Autosomal reces...	ABCA5	HGNC:35	23461	ENSG00000154...	uc002jig.3	chr17:67240452...	chr17:69244311...

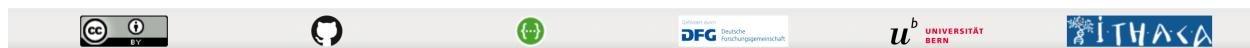


Figure 10: Panels view

2.5 Analyses views

2.6 About page

The About page contains general information about the project, its creators, funding, updates, and how to find help.

2.7 Login page

The Login page shows a simple form with inputs for the (1) user name and their (2) password, (3) buttons to reset the form and (4) links to registration and password reset.

Figure 11: Login modal

2.7.1 Register user page

This page can be used to apply for a SysNDD account by entering following information:

- 1) desired username
- 2) institutional E-mail
- 3) ORCID identifier
- 4) First name
- 5) Family name
- 6) description of your interest in SysNDD and why you want to participate in the curation effort

and (7) accepting the terms of use.

The (8) buttons allow resetting or submitting the form.

Upon submission the curator status users will receive a mail to review your application. After your application has been confirmed you will receive a mail with your login information and instructions.

2.7.2 Reset password page

The form on this page allows users who forgot their password to reset this by entering the E-mail they registered with.

Upon submission the E-mail account will receive a message with a one-time link allowing the user to enter a new password.

Figure 12: Register modal

Figure 13: Reset modal

3 API

The SysNDD api (application programming interface) is available from <https://sysnnd.dbmr.unibe.ch/API>.

The api is written in R using the plumber package¹⁶.

We intend to follow the Swagger/ OpenAPI¹⁷ and JSON:API¹⁸ specifications.

The api scripts run in a Docker container using the official “rocker/tidyverse” image (version 4.2.0).

As R is single threaded, we deploy multiple instances of the api container. These are bundled together using HAProxy¹⁹ load balancer.

The api is rate limited through our NGINX²⁰ web server configuration with a rate limit of 10 requests per second (10r/s; equals 1 request every 100 milliseconds) per requesting ip. The configuration allows bursts of up to 30r/s but introduces a delay after 10 requests to enforce the rate limit.

3.0.1 Endpoints

The SysNDD api currently contains all endpoints for external and internal usage in one api script. This may change with future releases.

The api is structured into different components based on the SysND concept:

¹⁶<https://www.rplumber.io/>

¹⁷<https://swagger.io/specification/>

¹⁸<https://jsonapi.org/>

¹⁹<http://www.haproxy.org/>

²⁰<https://www.nginx.com/>

- entity: Entity related endpoints
- review: Reviews related endpoints
- status: Status related endpoints
- re_review: Re-review related endpoints
- publication: Publication related endpoints
- gene: Gene related endpoints
- ontology: Ontology related endpoints
- inheritance: Inheritance related endpoints
- phenotype: Phenoptype related endpoints
- panels: Gene panel related endpoints
- comparisons: NDD gene list comparisons related endpoints
- search: Database search related endpoints
- list: Database list related endpoints
- statistics: Database statistics
- user: User account related endpoints
- authentication: Authentication related endpoints

The endpoints are documented and can be tested using the Swagger/ OpenAPI user interface at <https://sysnnd.dbmr.unibe.ch/API>. Here one can generate cURL requests to use in external software.

3.0.2 Usage policy

The SysNDD api powers the web tool for everyday users. We also provide the SysNDD api free to allow users to use the SysNDD data and build on it by creating software or services that connect to our platform.

Usage requirements: - optimize your requests to stay in the above described limits - be sensible about re-using data (e.g., store your requests until data is updated on our server) - use pagination where possible instead of requesting large data chunks (e.g., restrict usage of “all” option in large, potentially blocking list endpoints like “entity” and “gene”) - if you require more api ressources please get in contact

Updates and disclaimer: - We provide the SysNDD api as-is. - Due to the current development status (version 0.X.Y) we may update or modify the api any time. These changes may affect your use of the api or the way your integration interacts with the api.

3.0.3 Authentication and authorization

The SysNDD api uses JSON Web Tokens (JWT²¹) to implement stateless authentication and authorization.

The api user can manually (test purposes) request a token by entering their login credentials in the ionput form provided at the “api/auth/authenticate” endpoint:

This endpoint will generate and respond with and JWT token:

This Bearer token can then be copied and entered in the OpenAPI/ Swagger authorize modal which opens after clicking the “Authorize” modal button at the upper right corner:

After entereing the token in the respective field (1) and cklicking the “Authorize” submission button the modal will change and show the login status. This field can be closed now:

The user is now fully authenticated and can access the endpoints requiring user rights:

The token is valid for 60 minutes. It can be refreshed using the endpoint “api/auth/refresh”.

²¹<https://jwt.io/>

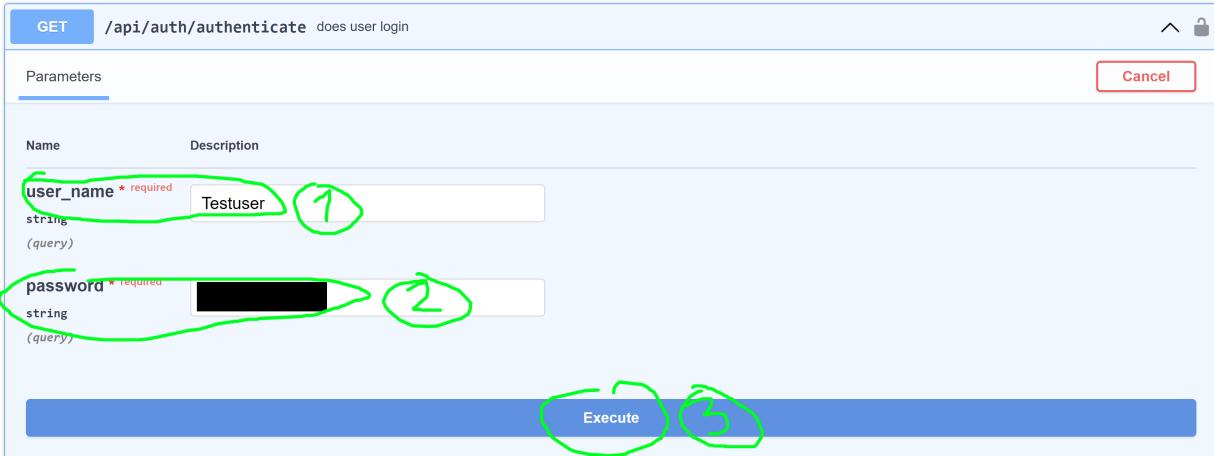


Figure 14: Authenticate endpoint

```
Response body
[{"eyJ0eXAiOiJKV1QiLCJhbGciOiJIUzI1NiJ9.eyJleHAiOiE2NTQ2NzcXMDo..."}]
  Copy Download
```

Figure 15: JWT token

SysNDD API 0.1.0 OAS3

<https://sysnnd.dbmr.unibe.ch/alb/openapi.json>

This is the API powering the SysNDD website and allowing programmatic access to the database contents.

[Terms of service](#)

[API Support - Website](#)

[Send email to API Support](#)

[CC BY 4.0](#)

Servers
<https://sysnnd.dbmr.unibe.ch/alb> ▾

[Authorize](#)

Figure 16: Authorize button

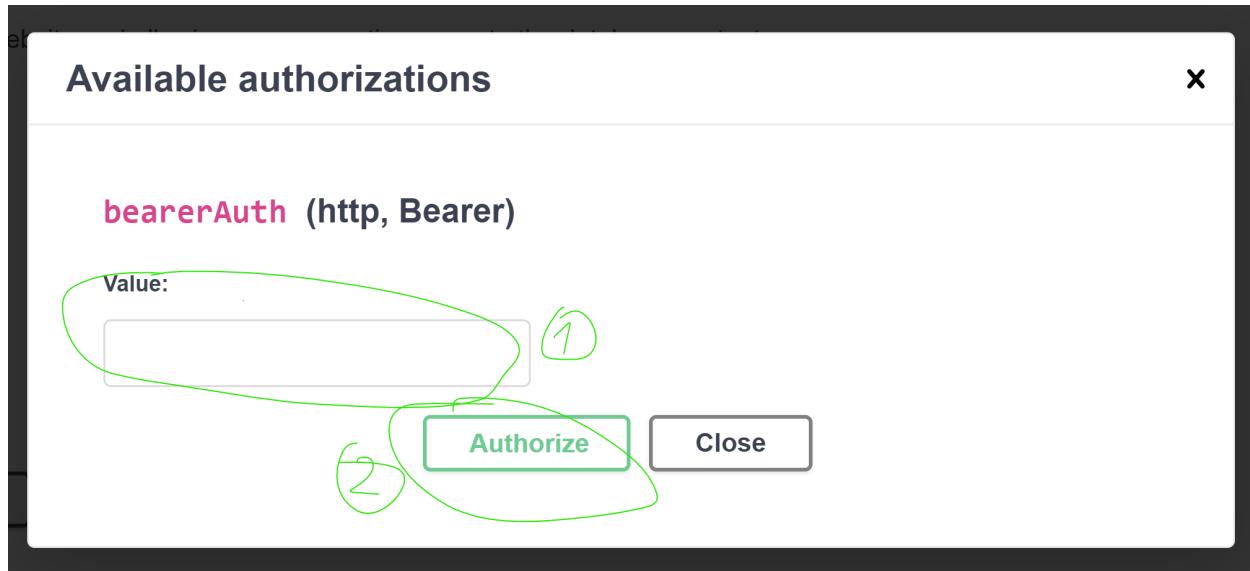


Figure 17: Authorize modal prompt

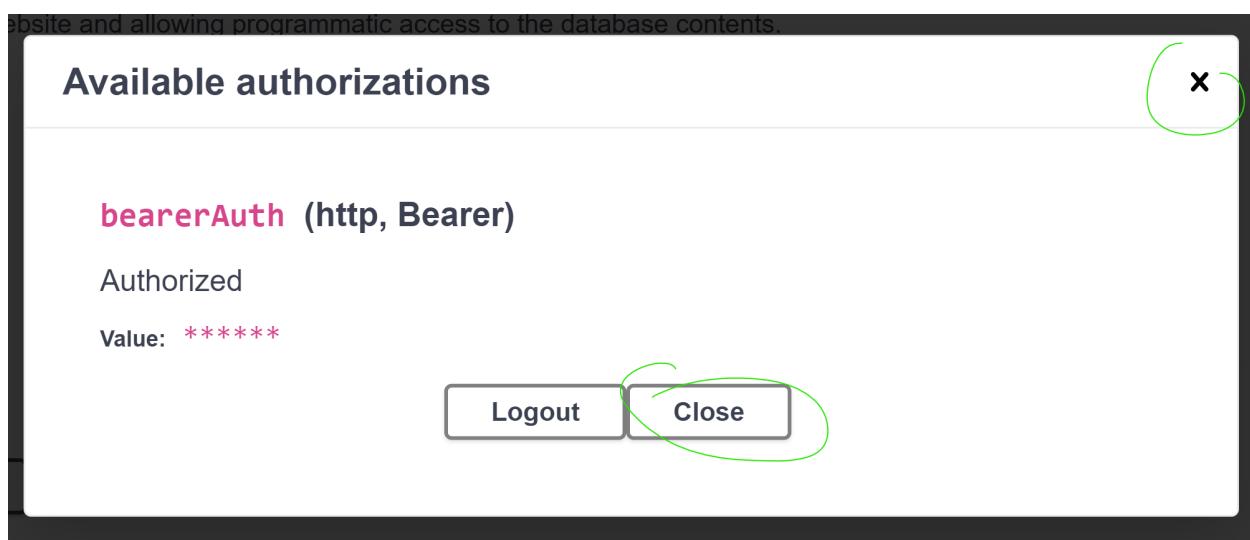


Figure 18: Authorize modal logged in

A screenshot of a web-based API endpoint list. At the top left is a "entity Entity related endpoints" header. Below it is a table of endpoints:

Method	Endpoint	Description
GET	/api/entity	allows filtering and field selection from all entities
POST	/api/entity/create	creates a new entity
POST	/api/entity/rename	renames an entity

To the right of the table are three small icons enclosed in a green circle, each consisting of a downward arrow and a lock symbol.

Figure 19: API logged in

4 Database structure

SysNDD currently uses the open-source MySQL 8.0²² relational database management system (RDBMS).

The design of our DB schema can be viewed in DB DESIGNER²³:

SysNDD DB schema²⁴

As of 2022-06-07 the database schema looks like this:

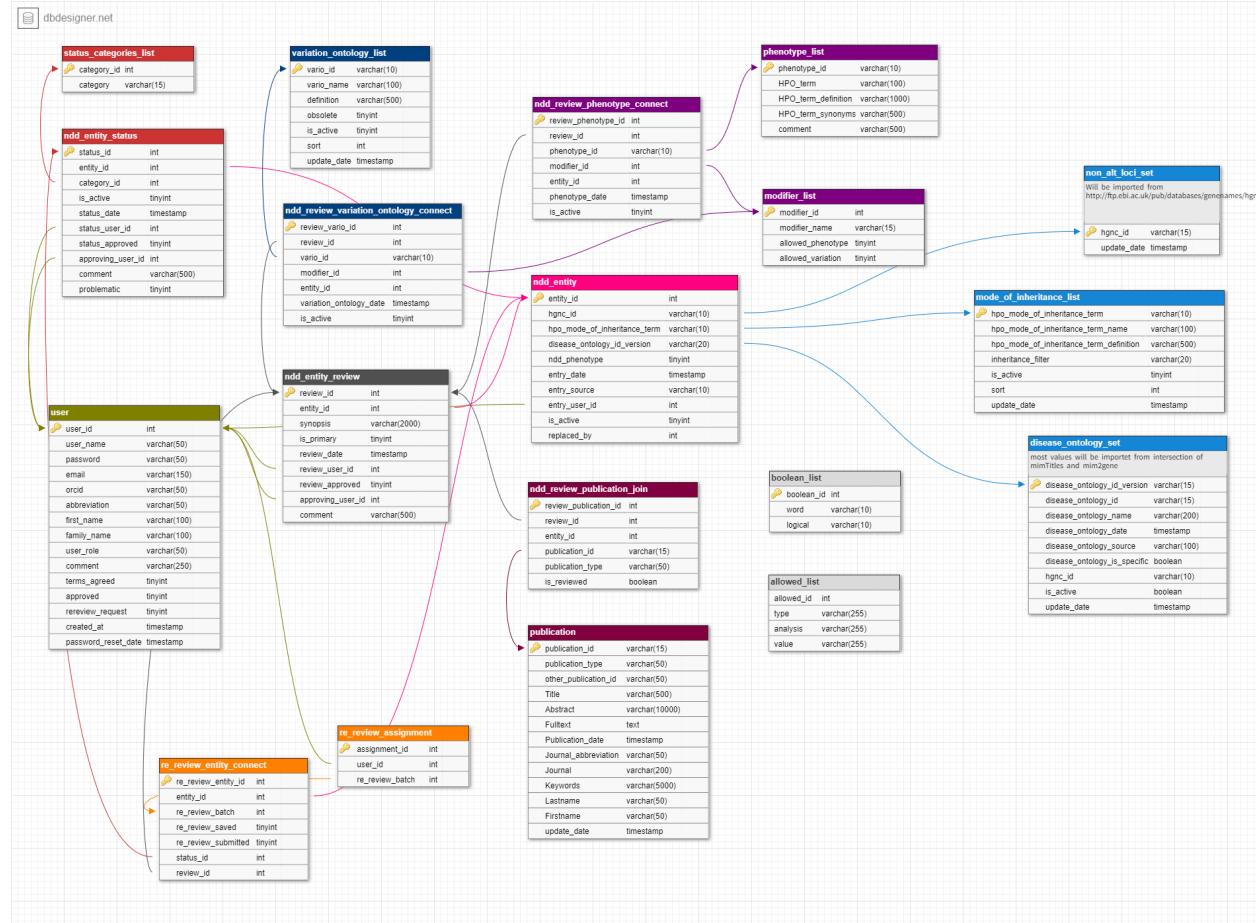


Figure 20: SysNDD MySQL database

The database runs in a docker container using the official mysql docker image²⁵ (version 8.0.29).

5 Curation criteria

²²<https://dev.mysql.com/doc/relnotes/mysql/8.0/en/>

²³<https://www.dbdesigner.net/>

²⁴<https://dbdesigner.page.link/3Morx9HZxzqt4R379>

²⁵https://hub.docker.com/_/mysql

5.1 Definitions

Intellectual disability (ID) and neurodevelopmental disorders (NDD) are defined in the scope of SysNDD as follows:

- Early onset neurodevelopmental delay and cognitive impairment (severe ID to learning difficulties)
- Regression/ neurodegeneration in the first years of life with or without prior developmental delay
- Disorders with cognitive impairment in a significant (ca. >10%) fraction of individuals

5.2 NDD Definitive entities

Inclusion criteria for Category 1 (“Definitive”):

1. Publication required (no grey literature like conference abstracts or personal communication; manuscripts on preprint servers can be considered individually but only entered through their DOI in the comment field until published with PMID, when they should be updated)

AND

2. Clear-cut frequency (no further criteria needed)

- ≥ 10 cases with *de novo* variants
- ≥ 5 autosomal-recessive families
- ≥ 3 families with X-chromosomal variants

OR

3. Cumulative evidence

- 1 strong frequency criterium

PLUS

- 1 strong genetic or 1 strong clinical criterium

OR

- 2 further strong (genetic and/or clinical) criteria in case of only 2 families with recessive inheritance

OR

- \geq two moderate criteria

5.2.1 Strong criteria

Strong frequency criteria:

- ≥ 3 patients with *de novo* variant
- ≥ 2 families with bi-allelic truncating variants
- $\geq (2\text{-})3$ families with bi-allelic missense variants
- ≥ 2 families with X-chromosomal variants

Strong genetic criteria:

- recurrence of a variant
- clustering of variants
- *de novo* truncating variants in a gene intolerant to loss-of-function variants (gnomAD constraint score)

Strong clinical criteria:

- Homogeneous phenotype
- Presence of specific/distinct clinical aspects (e.g., recognizable facial gestalt; rare specific malformations; pattern of multiple malformation; characteristic MRI anomalies; specific metabolic/enzymatic anomalies)

5.2.2 Moderate criteria

- Multigenerational segregation of variants
- Functional tests
- Gene involved in a pathway/complex where variants in other subunits are associated with a similar phenotype
- *De novo* missense variants in a gene intolerant to missense variants (gnomAD constraint scores)

5.2.3 Possible negative criteria

These should be included into consideration in borderline cases.

- Age of first publication(s) without further confirmatory reports in the meantime
- Publication quality and journal or genetics expertise “doubtful”
- New evidence against gene and/or variants: e.g., constraint scores, frequencies in gnomAD

5.3 NDD Moderate and Limited entities

These categories include the previous category of “candidate genes” and are now split into criteria for entity categories 2 (“Moderate”) and 3 (“Limited”):

1. Must be published (no private, in-house candidate lists)

AND

2. ID indicated, but criteria not sufficient for category 1, examples:

a. Limited genetic evidence

- < 3 cases with *de novo*, different variants and non-specific NDD phenotype
- 1 recessive family with truncating variant or <= 2 recessive families with missense variants (category 2 or 3 depending on number of affected and tested individuals per family, functional evidence and homogeneity of phenotype etc.)
- candidate gene from translocation or larger deletion
- reports of enzymatically confirmed patients with specific metabolic disorders but without genetic mutation confirmed

b. Limited clinical evidence

- not much evidence for ID, e.g. reported as ADHD or ASD or neurological disorder without clearly reported low IQ and ID
- known disorder, but only single patients reported with ID
- motor developmental delay without evidence for cognitive impairment
- clear neurodegenerative course without ID or cognitive delay present in the first years
- lethal before ID might be evident, although e.g. brain malformations or metabolic abnormalities might point to ID
- ID reported in other, similar disorders caused by mutations in the same pathway/complex but not (yet) in association with this particular gene (e.g. Fanconi anemia)

c. Limited combined genetic and clinical evidence

- Gene enriched for *de novo* or rare deleterious variants in large NDD cohorts or meta-studies, no further details

5.3.1 Exclusion criteria

1. Published as candidate gene only based on function or experimental results but without variants reported in humans

AND/OR

2. Only 1 *de novo* case from longer ago without further evidence and gene tolerant towards missense and/or loss-of-function variants according to gnomAD constraint scores

AND/OR

3. Only 1 sporadic case with bi-allelic variants and without any further supporting evidence such as segregation in other family members, functional tests, similar phenotypes in other patients with variants in genes from the same pathway, etc.

5.3.2 When to choose category 2 (“Moderate”)?

Too good for category 3 (“Limited”) but not good enough for category 1 (“Definitive”)

Examples:

- Recurrent *de novo* variant in 2 individuals with a similar phenotype
- Bi-allelic or X-chromosomal truncating variant segregating in \geq two generations of a large family
- Convincing functional evidence
- 1-2 patients with convincing variants in a gene which is in the same complex/pathway with other known disease genes and phenotype fits (e.g. CDG syndrome)

5.3.3 Special case: non-NDD entities

Some genes are associated with multiple entities. Among these entities there might be some without ID as a clinical feature. These non-NDD entities will be included in SysNDD but they will not be classified to any of the categories. Instead, they are tagged with “n.a.” (not applicable).

6 Re-review instructions

The goal of the SysNDD “Re-Review” effort is to update and standardize the SysID entities collected during the past years to enable better integration into and interoperability international with gene curations.

6.1 Re-review tool usage

We created Reviewer status accounts for participating scientists.

6.1.1 Login

You can log into your account by pointing your browser to <https://sysnnd.dbmr.unibe.ch/> and then clicking the “Login“ link on the right side of the menu:

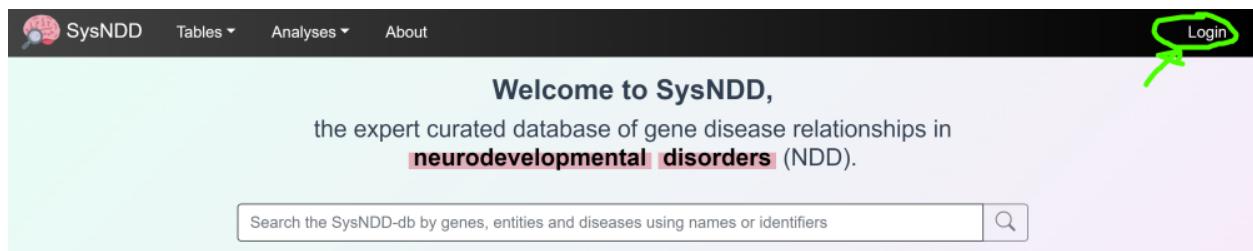


Figure 21: Login menu

On the Login page enter your credentials and press the Login button:

After successful login, you will be redirected to the start page and the navigation bar will show new links depending on your account privileges:

Your login token (JWT; JSON Web Token) is valid for 1 hour, after which you will be logged out. You can however always refresh the time by clicking the link in the user menu. The website will warn you at 5, 3 and 1 minutes before log out.

6.1.2 Review page

Click the “Review” link to your personal “Re-Review” site:

The “Re-Review” page is structured as a table enriched with information and controls.

These show you the number of entities assigned to your account

- (1) your account information status specific controls (e.g. switching to “Curator” mode, applying for a new batch of entities)
- (2) menu items to filter/ navigate the table
- (3) and finally, the table with the entity information and
- (4) controls to review and change the information:

By clicking the action buttons, you can open 3 different windows to change the entities review:

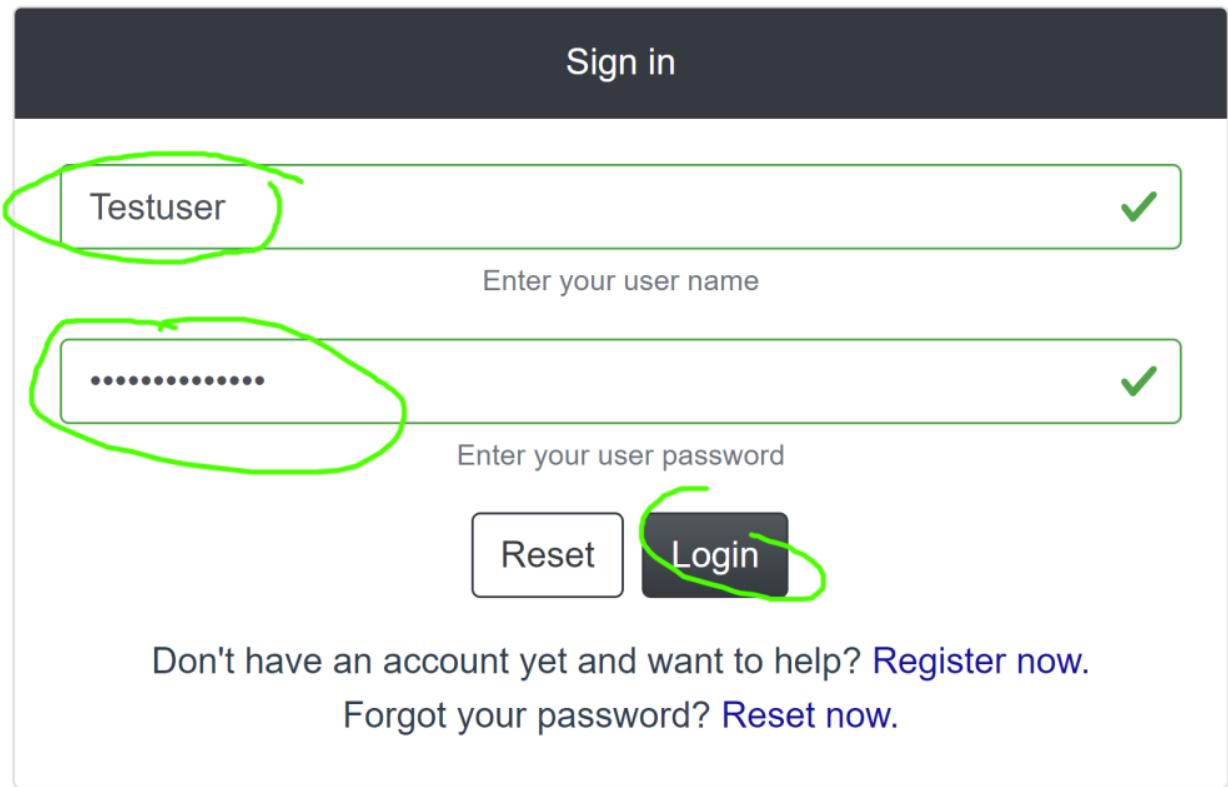


Figure 22: Login page

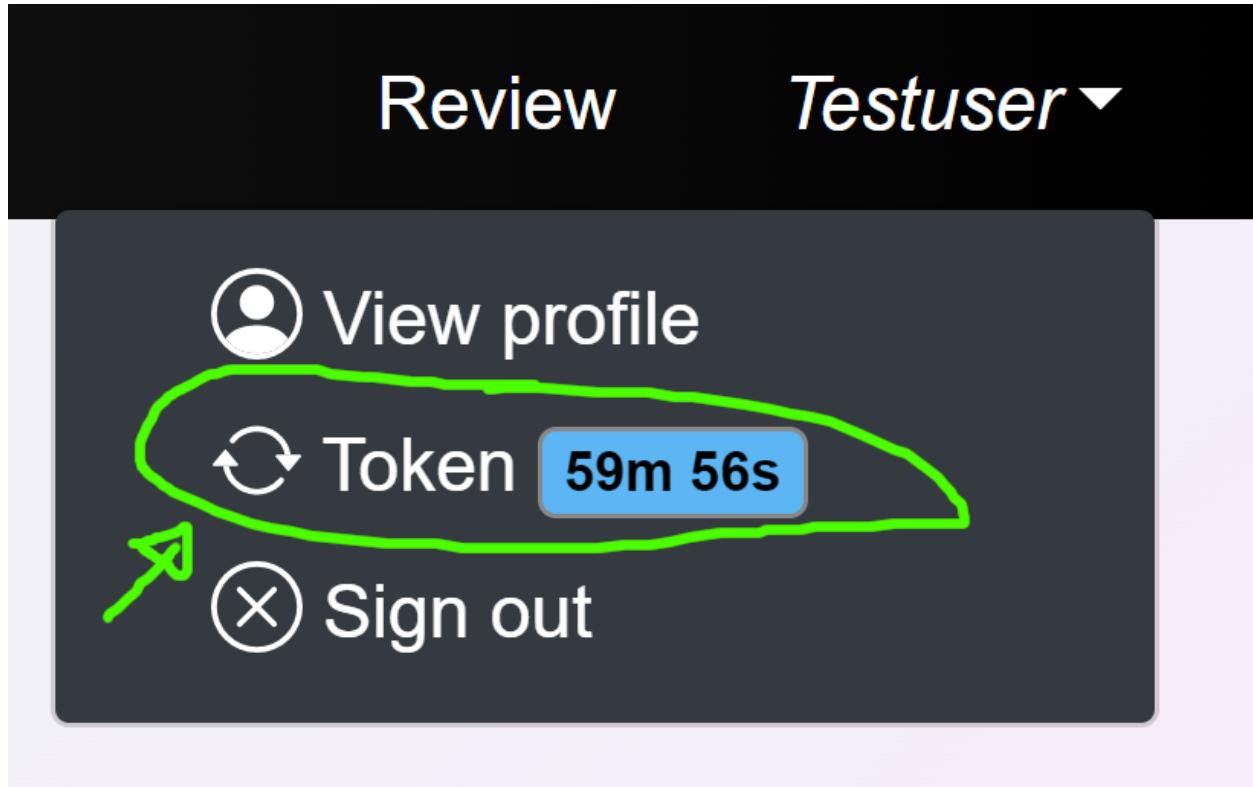


Figure 23: Login token menu

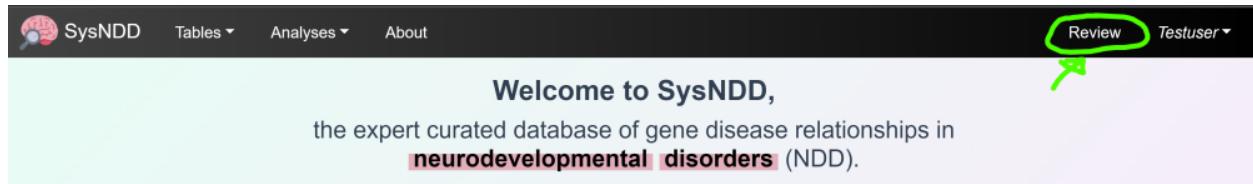
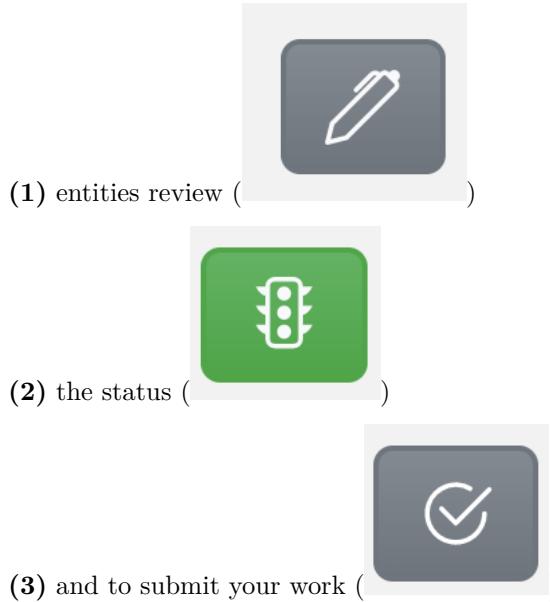


Figure 24: Review page menu

Re-review table Entities: 29					
Entity	Gene	Disease	Inheritance	NDD	Actions
sysndd:49	AP3B1	Hermansky-Pudlak syndrome 2	AR	✓	
sysndd:50	AP4B1	Spastic paraplegia 47, autosomal	AR	✓	
sysndd:51	AP4E1	Spastic paraplegia 51, autosomal	AR	✓	
sysndd:52	AP4S1	Spastic paraplegia 52, autosomal	AR	✓	
sysndd:53	APTX	Ataxia, early-onset, with oculomotor apraxia	AR	✓	
sysndd:54	ARFGEF2	Periventricular heterotopia with microcephaly	AR	✓	
sysndd:55	ARHGEF6	intellectual disability	XR	✓	
sysndd:56	ARID1A	Coffin-Siris syndrome 2	AD	✓	
sysndd:57	ARID1B	Coffin-Siris syndrome 1	AD	✓	
sysndd:58	ARL13B	Joubert syndrome 8	AR	✓	

Figure 25: Review page



6.1.3 New Review edit

In this window you have:

- the possibility to change/adapt or completely rewrite the current synopsis (1),
- add, or remove phenotype associations (2),
- add or remove publications from the review by PMID (3)
- and add/ edit fitting GeneReviews articles by PMID (4).
- Finally, you can add a comment to your review for the Curator later approving this entities changes (5) and
- save your review (6).

By clicking on the little question marks you can show help messages for each item:

The screenshot shows a 'Modify review for entity' interface for OMIM ID sysnnd:49. The page is divided into several sections with numbered green annotations:

- 1 Synopsis**: multi system disorder, albinism, bleeding diathesis, pulmonary fibrosis, granulomatous colitis, platelet and T-lymphocyte dysfunction and neutropenia, mild ID in some patients
- 2 Phenotypes**: present: Abnormality of the eye, present: Intellectual disability, rare: Intellectual disability, mild, present: Abnormality of blood and blood-forming tissues
- 3 Variation ontology**: present: variation
- 4 Publications**: Enter PMID separated by comma or semicolon
- 5 Genereviews**: Enter PMID separated by comma or semicolon, with PMID:20301464 listed
- 6 Comment**: Additional comments to this entity relevant for the curator.

At the bottom, it says 'Review by: Christiane Curator' and has 'Cancel' and 'Save review' buttons. The 'Save review' button is circled in green.

Figure 26: Review page

These help instructions are:

Synopsis: Short summary for this disease entity. Please include information on: a) approximate number of patients described in literature, b) nature of reported variants, b) severity of intellectual disability, c) further phenotypic aspects (if possible with frequencies) d) any valuable further information (e.g. genotype-phenotype correlations).

Examples:

de novo truncating or missense variants in > 20 individuals: variable ID (mild to severe), 50% short stature and microcephaly, 30% seizures, non-specific facial dysmorphisms, variable cardiac and renal anomalies in some

bi-allelic truncating variants in 7 individuals from 3 families: severe ID, microcephaly, seizures in 3/7, MRI anomalies

Phenotypes: Add or remove associated phenotypes. Only phenotypes that occur in 20% or more of affected individuals should be included. Please also include information on severity of ID where available and applicable.

Publications: No complete catalogue of entity-related literature required! If information in the clinical synopsis is not only based on OMIM entries, please include PMID of the article(s) used as a source for the clinical synopsis.

GeneReviews: Please add PMID for GeneReview article if available for this entity.

Comment: Additionally add information about your review potentially helpful to the curator approving the entity later.

6.1.4 New Status edit

In this window you can propose

- to change the entities association confidence category (1),
- suggest it's overall removal (2),
- add a comment for your change suggestions for the Curators to better understand the proposal (3) and
- save your work (4):

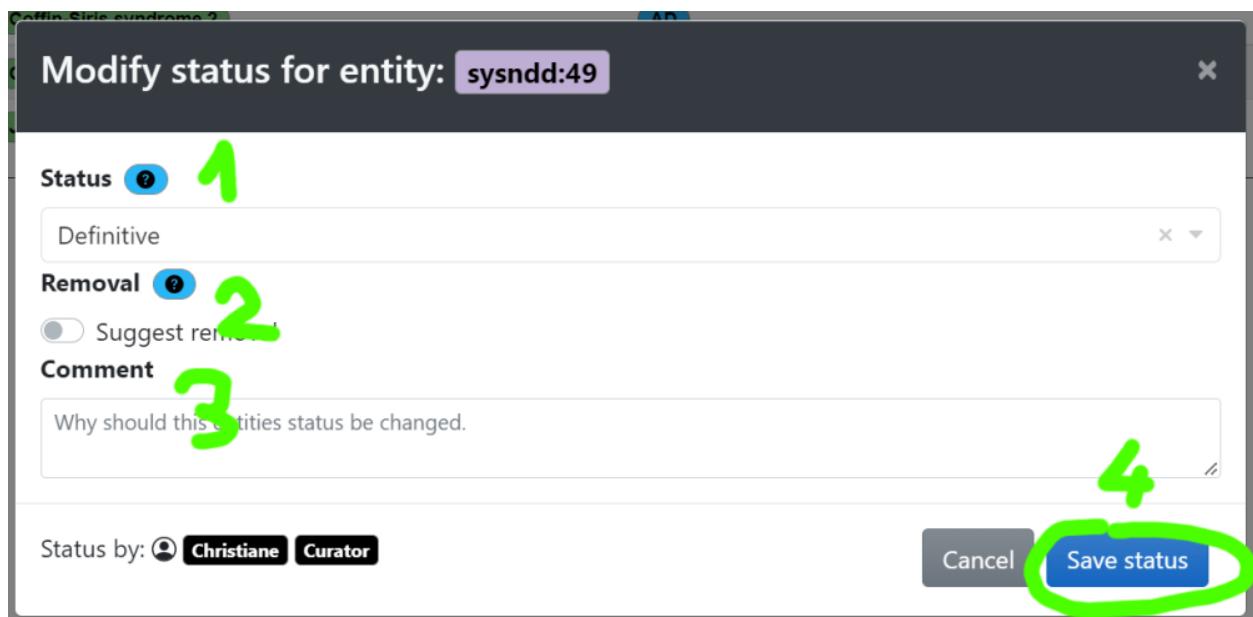


Figure 27: Submit re-review modal

6.1.5 Submit Re-review

The last action window is just to confirm that you are satisfied with your work and would like to submit it for curation:

After clicking this button, the entity will disappear from your list. And you can proceed with the remaining entries until no entity is left in your list.

6.2 Re-review curation

6.2.1 Definitive association status

1. Check if category 1 ("Definitive") is correct or shift status to category 2 ("Moderate") or 3 ("Limited"), where appropriate

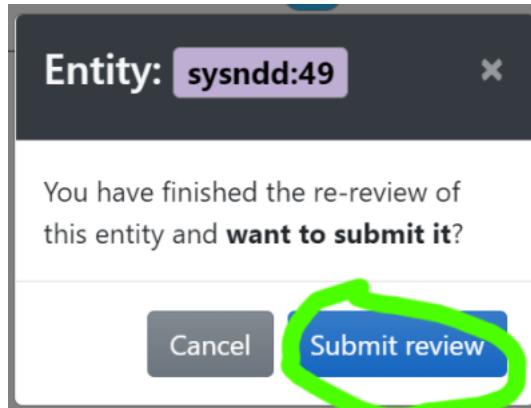


Figure 28: Submit re-review modal

2. Check and revise gene-related entities regarding diseases/inheritance patterns (ID and non-ID disorders) -> non-ID disorders will not go into any of the categories but will be tagged with “n.a.” (not applicable)
3. Check and revise associated phenotypes: select HPO terms from the list, only use HPO term if this specific aspect is present in approximately $\geq 20\%$ of patients. Please also check and revise severity of ID using HPO terms. If ID is very variable, select all appropriate ID terms (e.g. severe, moderate, mild, borderline)
4. Check references (OMIM, PMID, GeneReviews). References do not have to be complete but should be sufficient to give a good impression on the mutational and clinical spectrum. Add references where it would add to the picture.
5. Check and revise clinical synopsis: it does not have to contain everything that is known but should give a short and comprehensive picture on:
 - which data the gene and disease category were chosen on and
 - the molecular and clinical picture.

Please include information on:

- a) approximate number of patients described in literature,
- b) nature of reported variants,
- c) severity of intellectual disability,
- d) further phenotypic aspects (if possible with frequencies),
- e) any valuable further information (e.g. genotype-phenotype correlations)

Examples:

de novo truncating or missense variants in > 20 individuals: variable ID (mild to severe), 50% short stature and microcephaly, 30% seizures, non-specific facial dysmorphism, variable cardiac and renal anomalies in some

bi-allelic truncating variants in 7 individuals from 3 families: severe ID, microcephaly, seizures in 3/7, MRI anomalies

6.2.2 Moderate and Limited association status

- Check if inclusion criteria for candidate genes are still fulfilled or if it should be deleted from the list (“Refuted”)

- Check if candidate status is still correct and sort it into Category 2 (“Moderate”) and 3 (“Limited”) (or reclassify to 1 (“Definitive”), if applicable)
- Check, if associated phenotype still fits
- Check, if references are correct, if there is any new published information and modify clinical synopsis where appropriate
- Clinical synopsis can be very short for candidate genes
- no associated phenotypes (HPO terms) and frequencies are needed for candidate genes, but could be helpful

Examples:

de novo missense variants in 2 individuals: autism, ID in 50%

bi-allelic missense variant in 2 affected individuals from 1 family: moderate ID, MRI anomalies

6.2.3 Refuted association status

- Check if there is current evidence against this gene association (e.g. few truncating variants described in old publications before gnomAD constrain scores and the gene now has a pLI of 0; genes reported in a family with later report of another cause etc.)

References

Kochinke, K., Zweier, C., Nijhof, B., Fenckova, M., Cizek, P., Honti, F., Keerthikumar, S., Oortveld, M. A. W., Kleefstra, T., Kramer, J. M., Webber, C., Huynen, M. A., and Schenck, A. (2016). Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *American Journal of Human Genetics*, 98(1):149–164.