BfR
German Federal Institute for Risk Assessment

SMAFIRA Web tool
for searching the scientific literature for alternatives

10.04.2025, FC3R

Mariana Neves, Daniel Butzke

Agenda

Introduction

Demo

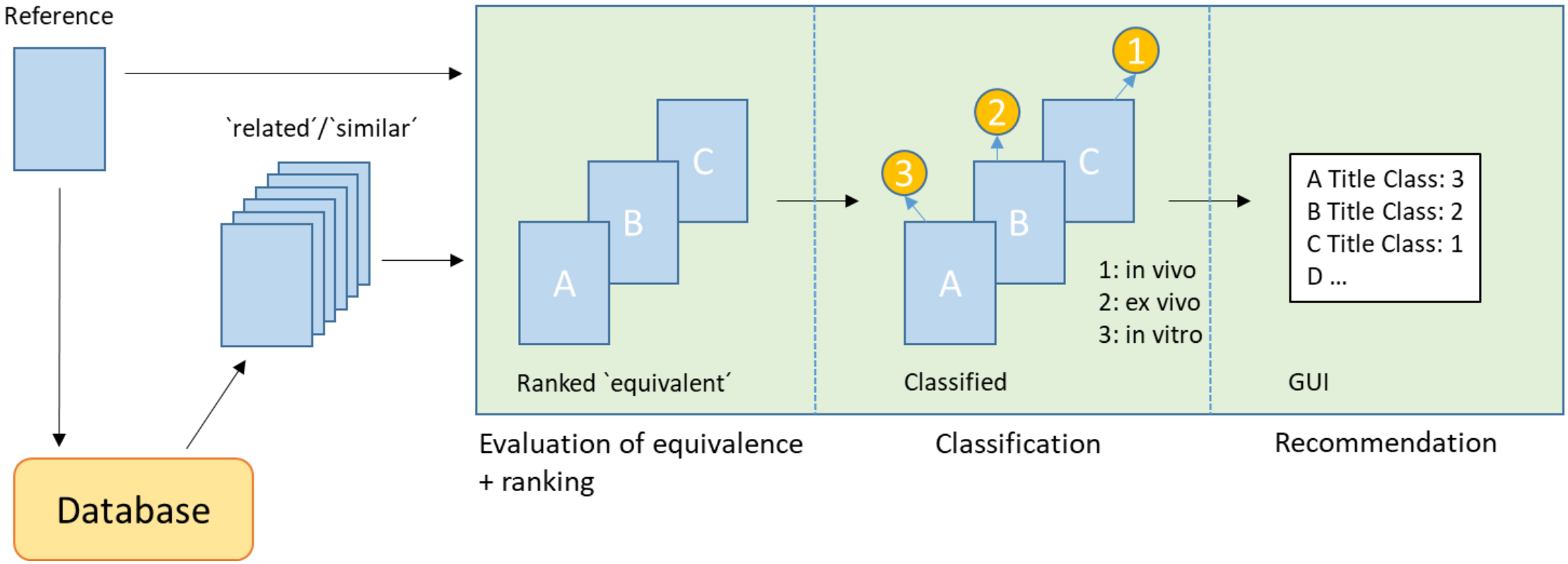
Methods

Current work

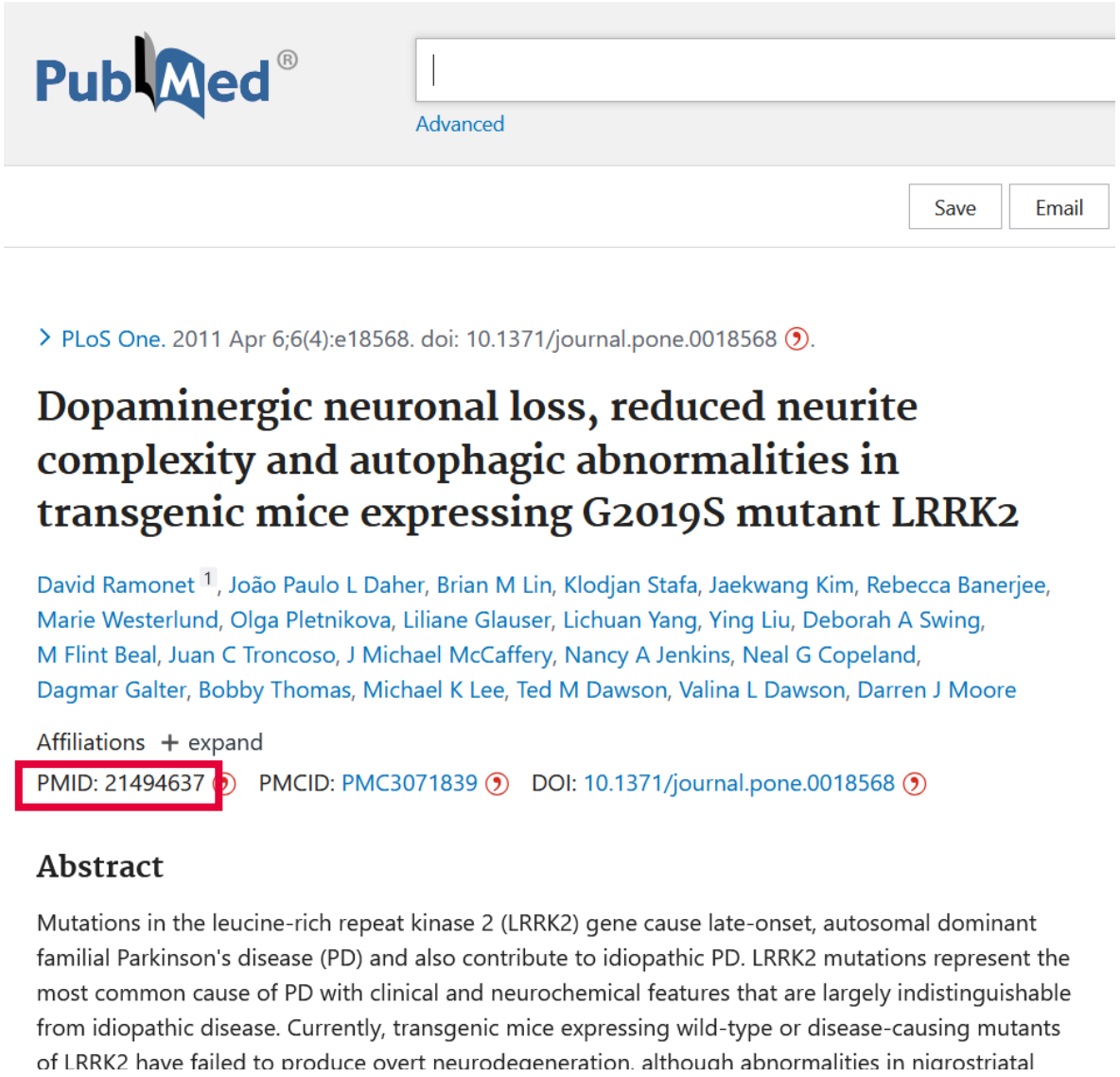
What is SMAFIRA?

- **A free online tool to screen the literature database PubMed© for candidate alternative methods**
- **Employs Artificial Intelligence and human feedback**
- **No login, no sensitive information, no search terms required**
- **No judgement on the suitability of retrieved methods as alternatives (reserved to human experts)**

Basic Workflow of SMAFIRA



Reference article



PubMed®

Advanced

Save Email

PLoS One. 2011 Apr 6;6(4):e18568. doi: 10.1371/journal.pone.0018568.

Dopaminergic neuronal loss, reduced neurite complexity and autophagic abnormalities in transgenic mice expressing G2019S mutant LRRK2

David Ramonet [1], João Paulo L Daher, Brian M Lin, Klodjan Stafa, Jaekwang Kim, Rebecca Banerjee, Marie Westerlund, Olga Pletnikova, Liliane Glauser, Lichuan Yang, Ying Liu, Deborah A Swing, M Flint Beal, Juan C Troncoso, J Michael McCaffery, Nancy A Jenkins, Neal G Copeland, Dagmar Galter, Bobby Thomas, Michael K Lee, Ted M Dawson, Valina L Dawson, Darren J Moore

Affiliations + expand

PMID: 21494637 PMCID: PMC3071839 DOI: 10.1371/journal.pone.0018568

Abstract

Mutations in the leucine-rich repeat kinase 2 (LRRK2) gene cause late-onset, autosomal dominant familial Parkinson's disease (PD) and also contribute to idiopathic PD. LRRK2 mutations represent the most common cause of PD with clinical and neurochemical features that are largely indistinguishable from idiopathic disease. Currently, transgenic mice expressing wild-type or disease-causing mutants of LRRK2 have failed to produce overt neurodegeneration, although abnormalities in nigrostriatal

BfR

SMAFIRA Demo

https://smafira.bf3r.de

BfR

SMAFIRA Search FAQ About Data Protection Declaration Imprint Contact

SMAFIRA is a free online tool that is intended to support biomedical researchers and animal welfare officers in their efforts to screen the literature database PubMed© for candidate alternative methods to animal experiments. SMAFIRA DOES NOT judge on the suitability of retrieved methods as alternatives. This judgement is reserved to human experts. Anyway, SMAFIRA decreases the time needed to search the literature for candidate alternatives. All you need is a PubMed® identifier (PMID).

Quick Tutorial

Search for a PubMed® identifier that describes an animal experiment with the research goal of your interest. Insert the PMID in the field. By clicking on the "Search for PUBMED similar articles" button, SMAFIRA will first let you confirm if it is the correct reference article.

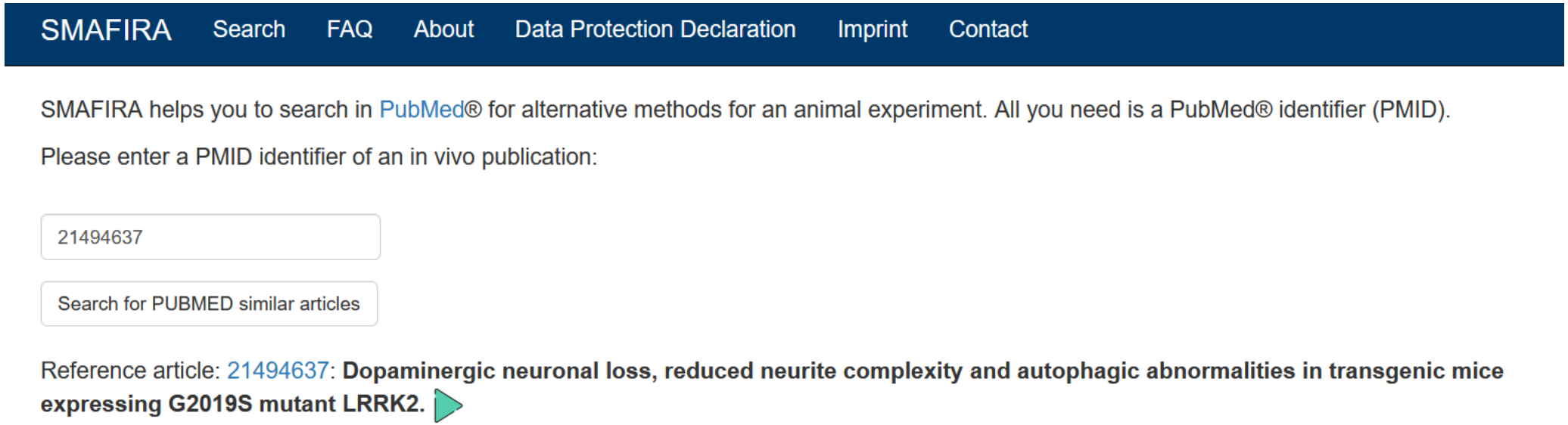
Please enter a PMID identifier of an in vivo publication:

Search for PUBMED similar articles

BfR German Federal Institute for Risk Assessment

GERMAN CENTRE FOR THE PROTECTION OF LABORATORY ANIMALS

Input of a PMID, confirmation of the reference article



List of results

SMAFIRA Search FAQ About Data Protection Declaration Imprint Contact

149 similar articles found for 21494637 Save the URL of your session: /artresults/?session=E4yHwi1ojAjPpMFA

Filters for the used experimental models

Vertebrate (other than human): in vivo, tissue/biopsy, primary cell, immortal cell | Invertebrate: all | Human: all | Virtual Model: all | Others: all

	PMID	SMAFIRA Rank	PubMed® Rank	My Rank	Title	Labels
1.	25174890	[1]	[1]		Conditional expression of Parkinson's disease-related R1441C LRRK2 in midbrain dopaminergic neurons of mice causes nuclear abnormalities without neurodegeneration.	
2.	21698001	[2]	[7]		Temporal expression of mutant LRRK2 in adult rats impairs dopamine reuptake.	
3.	25174649	[3]	[3]		Behavioral, neurochemical, and pathologic alterations in bacterial artificial chromosome transgenic G2019S leucine-rich repeated kinase 2 rats.	
4.	20382224	[4]	[10]		LRRK2-mediated neurodegeneration and dysfunction of dopaminergic neurons in a Caenorhabditis elegans model of Parkinson's disease.	
5.	20130188	[5]	[11]		Enhanced striatal dopamine transmission and motor performance with LRRK2 overexpression in mice is eliminated by familial Parkinson's disease mutation G2019S.	
9.	26123485	[9]	[2]		Selective expression of Parkinson's disease-related Leucine-rich repeat kinase 2 G2019S missense mutation in midbrain dopaminergic neurons impairs dopamine release and dopaminergic gene expression.	
10.	26363496	[10]	[27]		(R1441C) LRRK2 induces the degeneration of SN dopaminergic neurons and alters the expression of genes regulating neuronal survival in a transgenic mouse model.	

Results per page [10] [25] [50]

« 1 2 3 4 5 … 14 15 »

BfR

List of results

SMAFIRA Search FAQ About Data Protection Declaration Imprint Contact

149 similar articles found for 21494637 Save the URL of your session: /artresults/?session=E4yHwi1ojAjPpMFA

Filters for the used experimental models

Vertebrate (other than human): in vivo, tissue/biopsy, primary cell, immortal cell | Invertebrate: all | Human: all | Virtual Model: all | Others: all

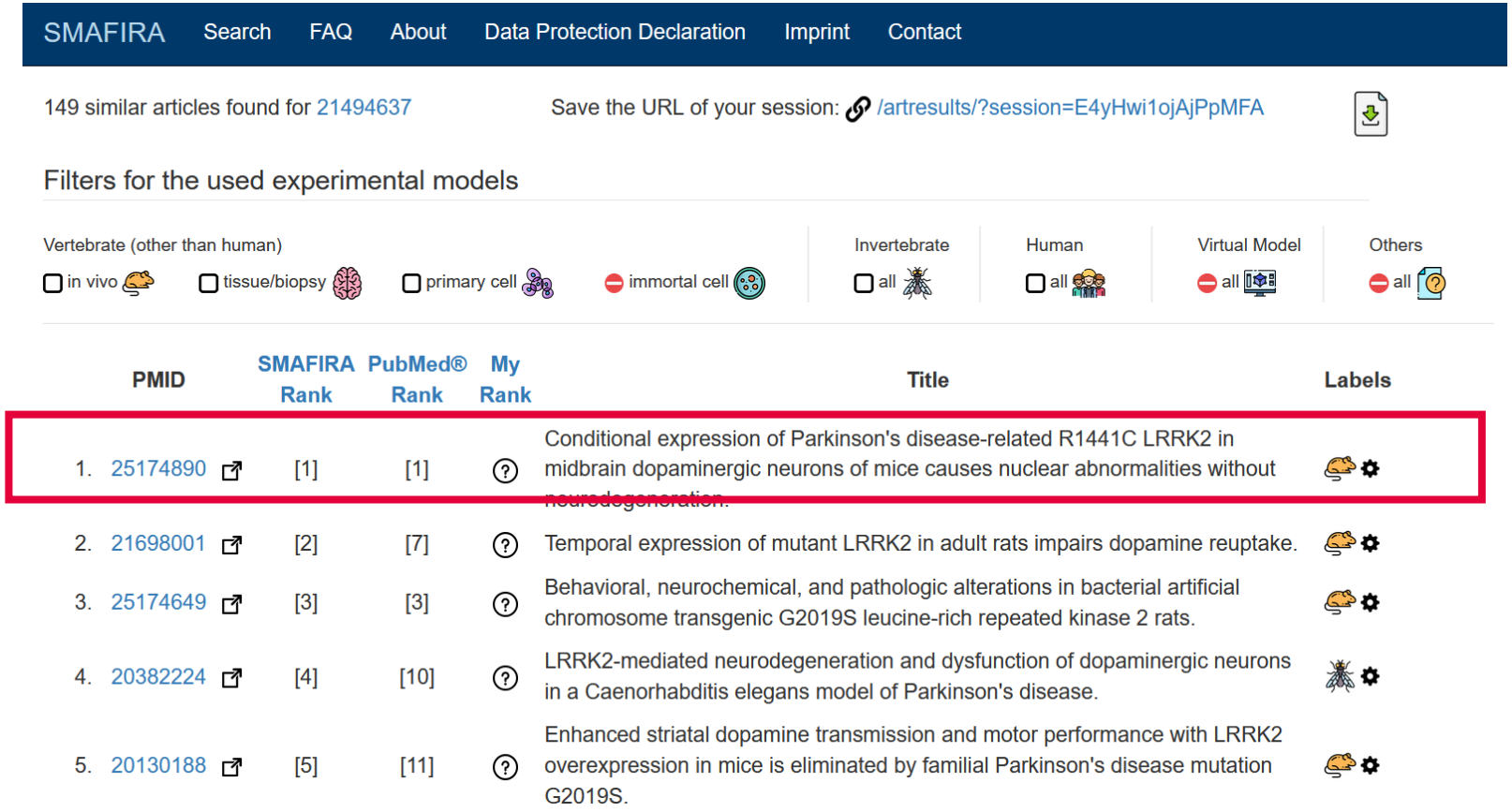
	PMID	SMAFIRA Rank	PubMed® Rank	My Rank	Title	Labels
1.	25174890	[1]	[1]		Conditional expression of Parkinson's disease-related R1441C LRRK2 in midbrain dopaminergic neurons of mice causes nuclear abnormalities without neurodegeneration.	
2.	21698001	[2]	[7]		Temporal expression of mutant LRRK2 in adult rats impairs dopamine reuptake.	
3.	25174649	[3]	[3]		Behavioral, neurochemical, and pathologic alterations in bacterial artificial chromosome transgenic G2019S leucine-rich repeated kinase 2 rats.	
4.	20382224	[4]	[10]		LRRK2-mediated neurodegeneration and dysfunction of dopaminergic neurons in a Caenorhabditis elegans model of Parkinson's disease.	
5.	20130188	[5]	[11]		Enhanced striatal dopamine transmission and motor performance with LRRK2 overexpression in mice is eliminated by familial Parkinson's disease mutation G2019S.	
9.	26123485	[9]	[2]		Selective expression of Parkinson's disease-related Leucine-rich repeat kinase 2 G2019S missense mutation in midbrain dopaminergic neurons impairs dopamine release and dopaminergic gene expression.	
10.	26363496	[10]	[27]		(R1441C) LRRK2 induces the degeneration of SN dopaminergic neurons and alters the expression of genes regulating neuronal survival in a transgenic mouse model.	

Results per page [10] [25] [50]

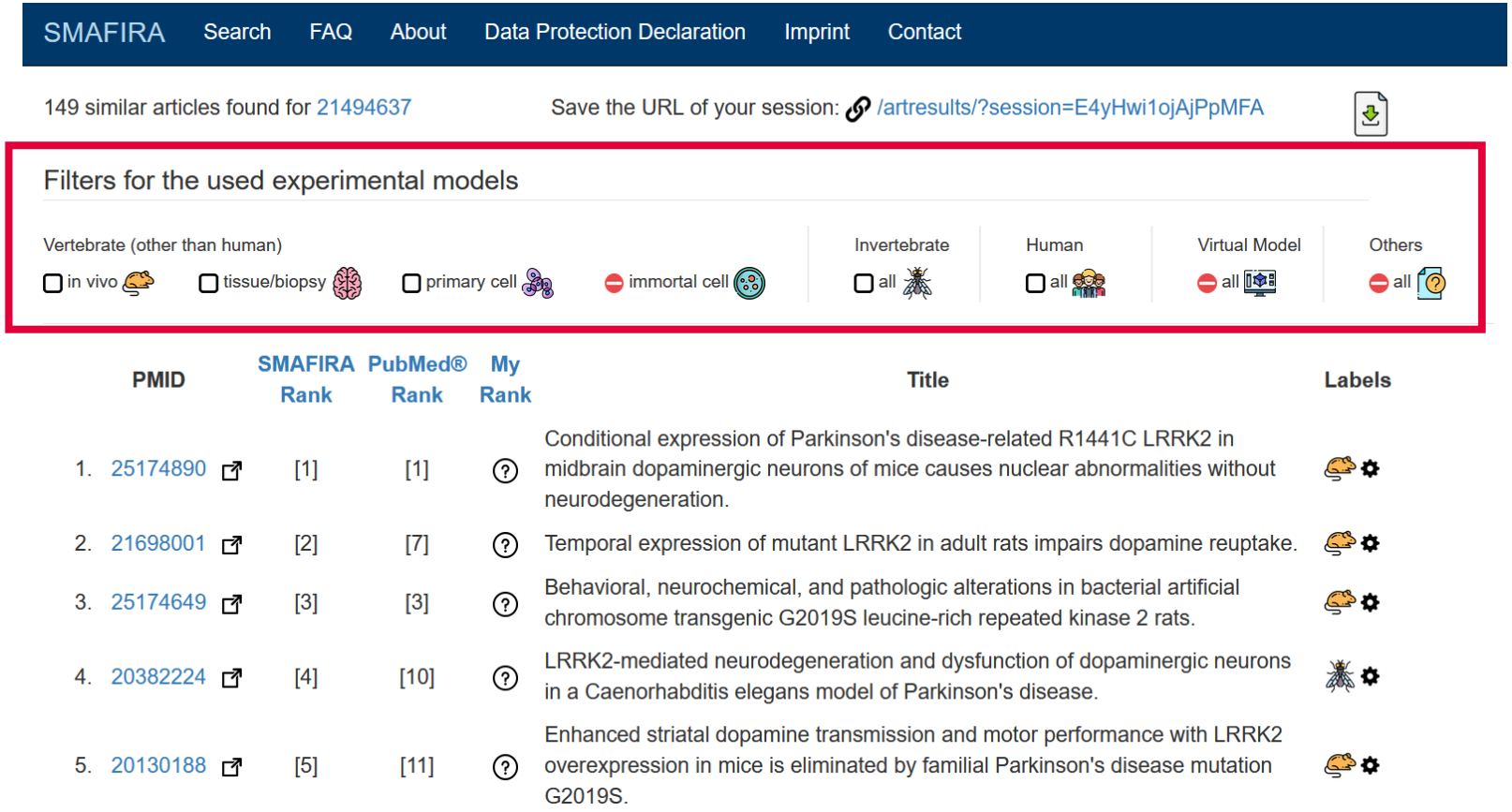
« 1 2 3 4 5 ... 14 15 »

BfR

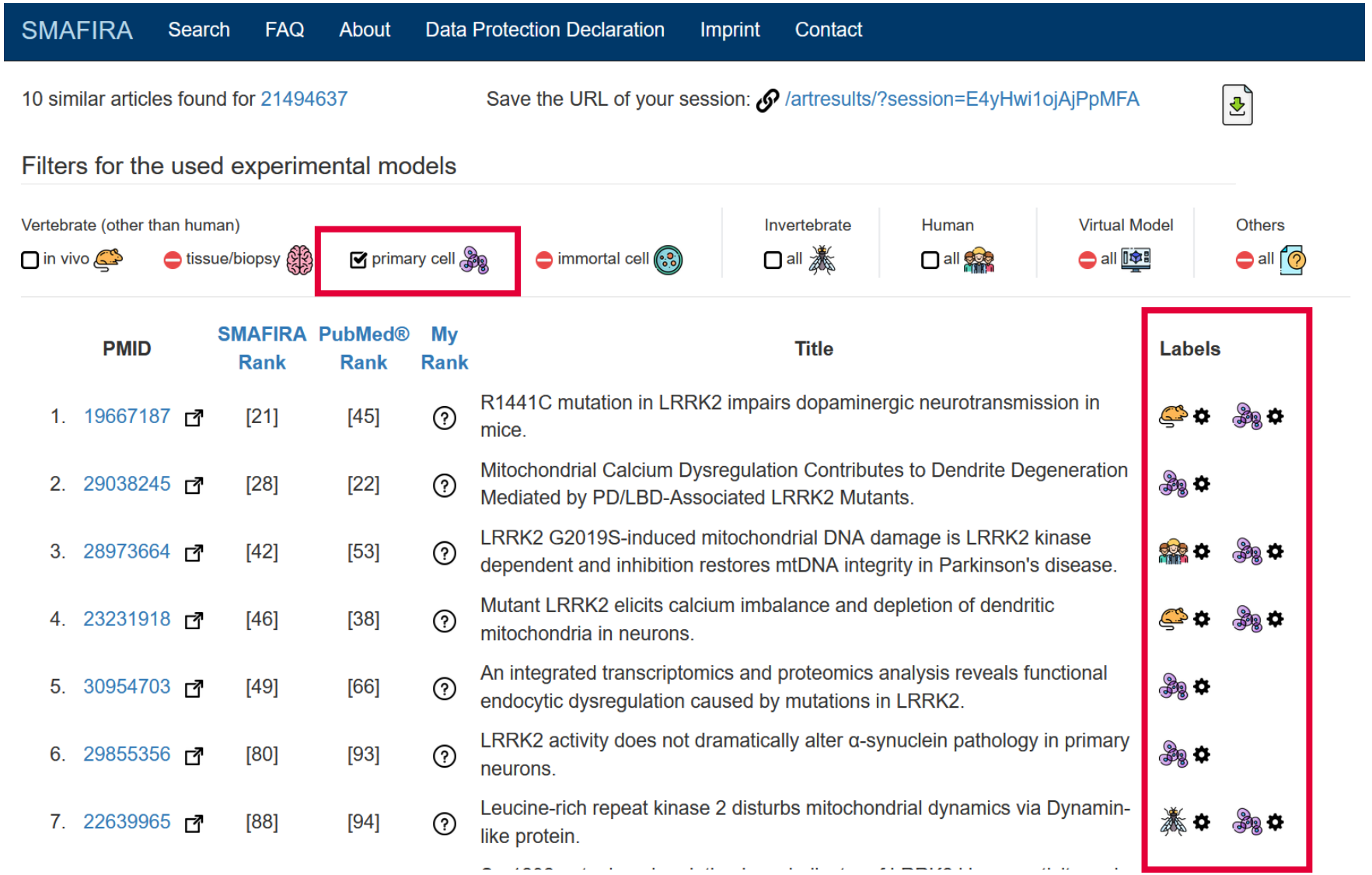
List of results



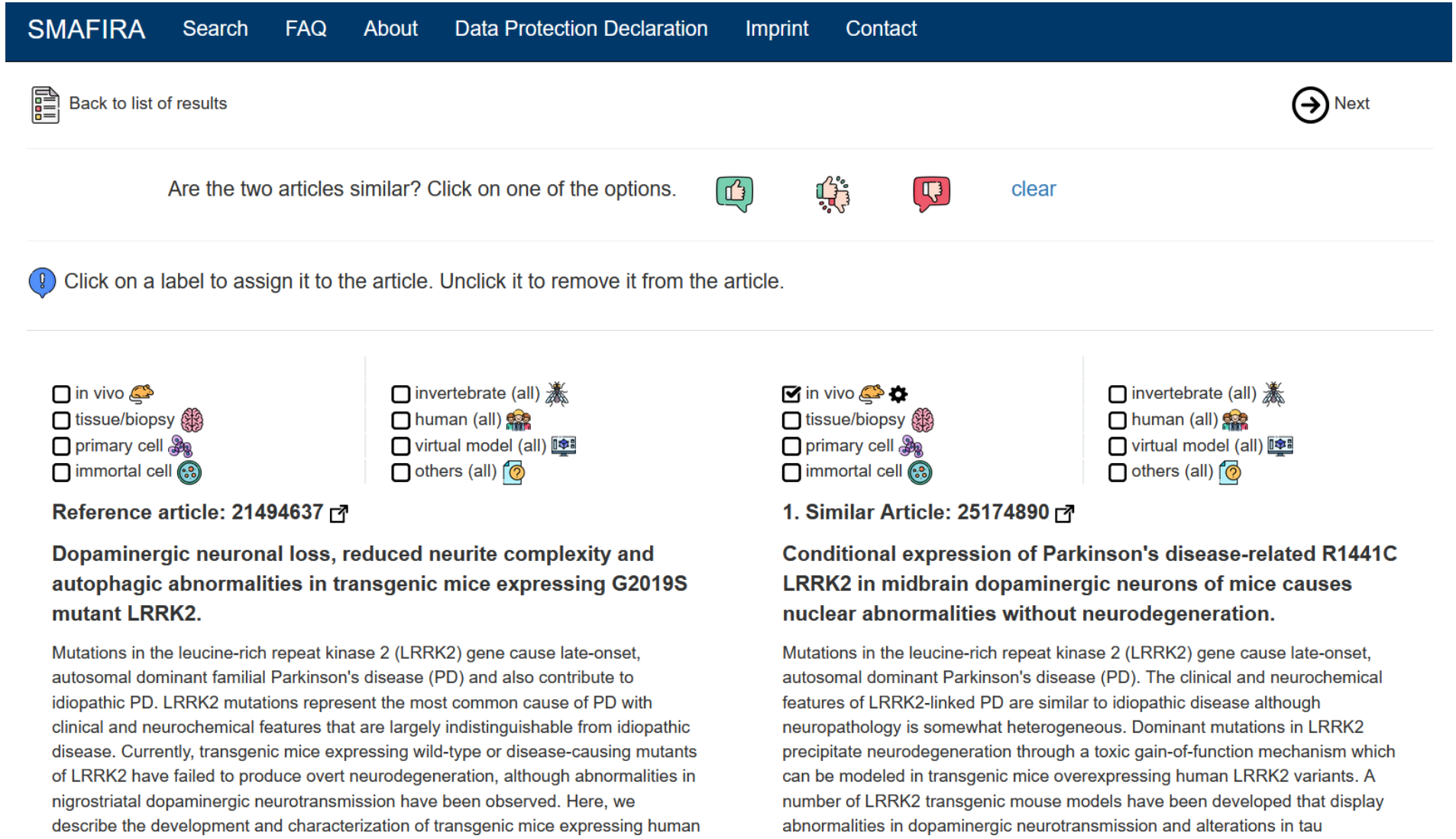
List of results



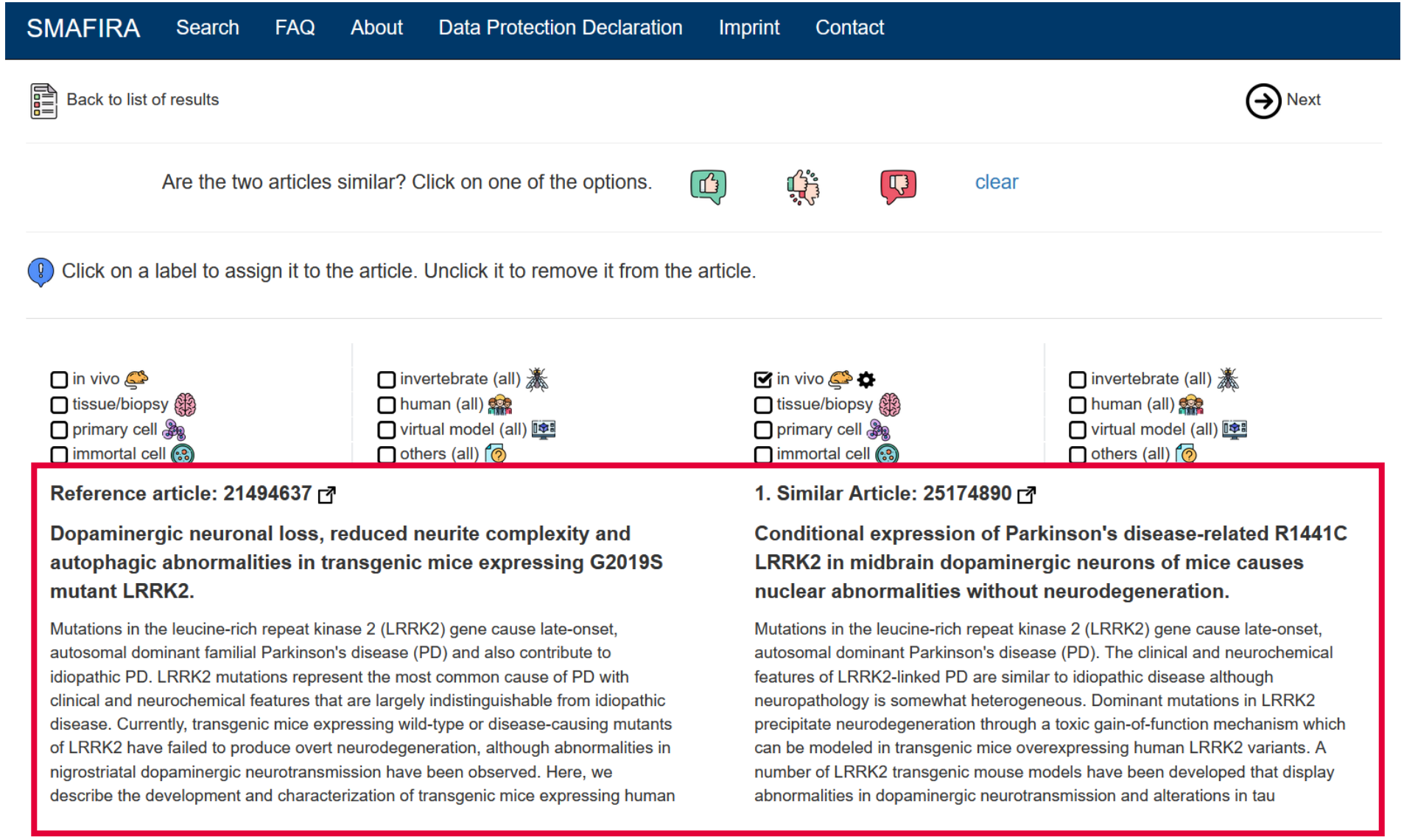
List of results: filters



User feedback

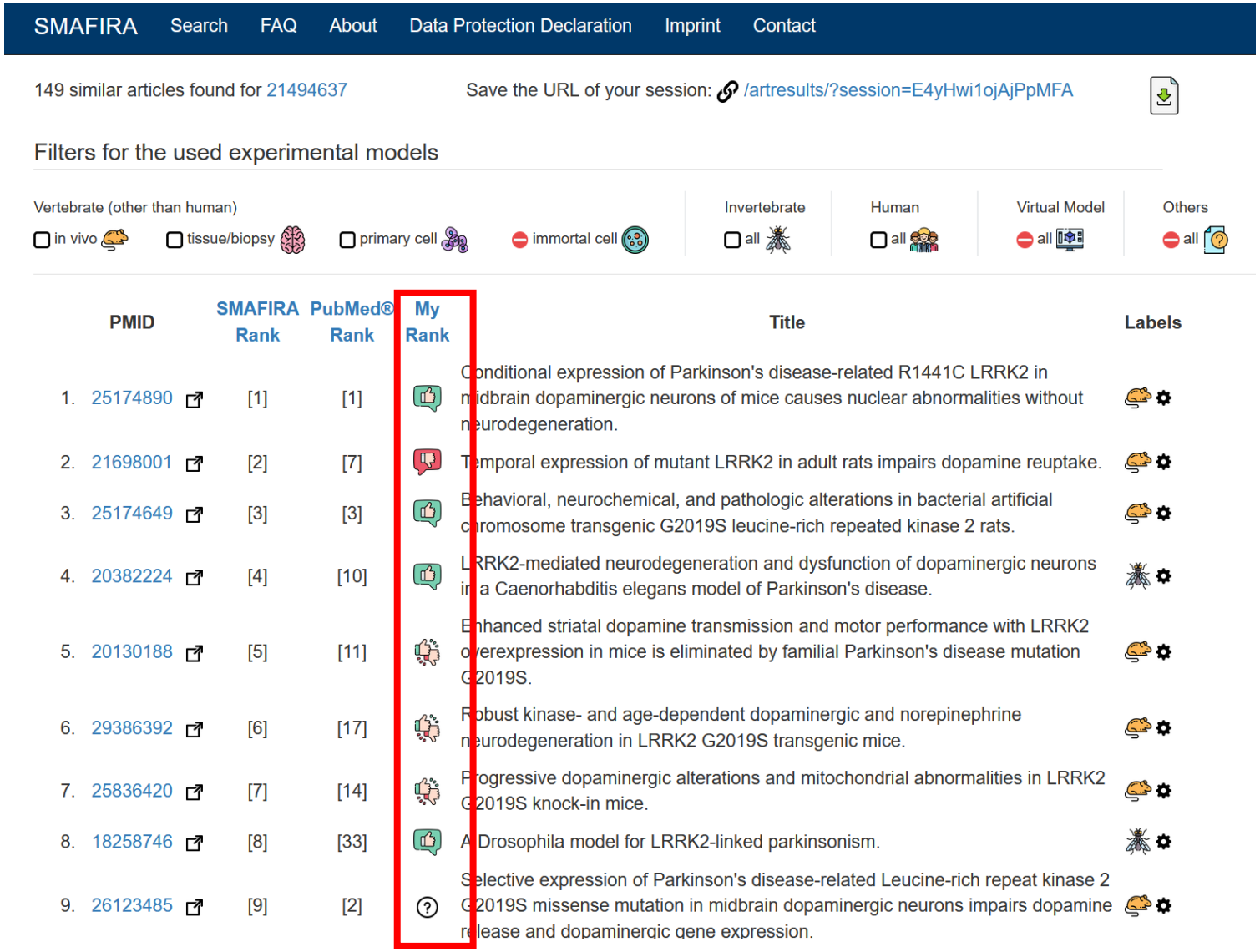


User feedback



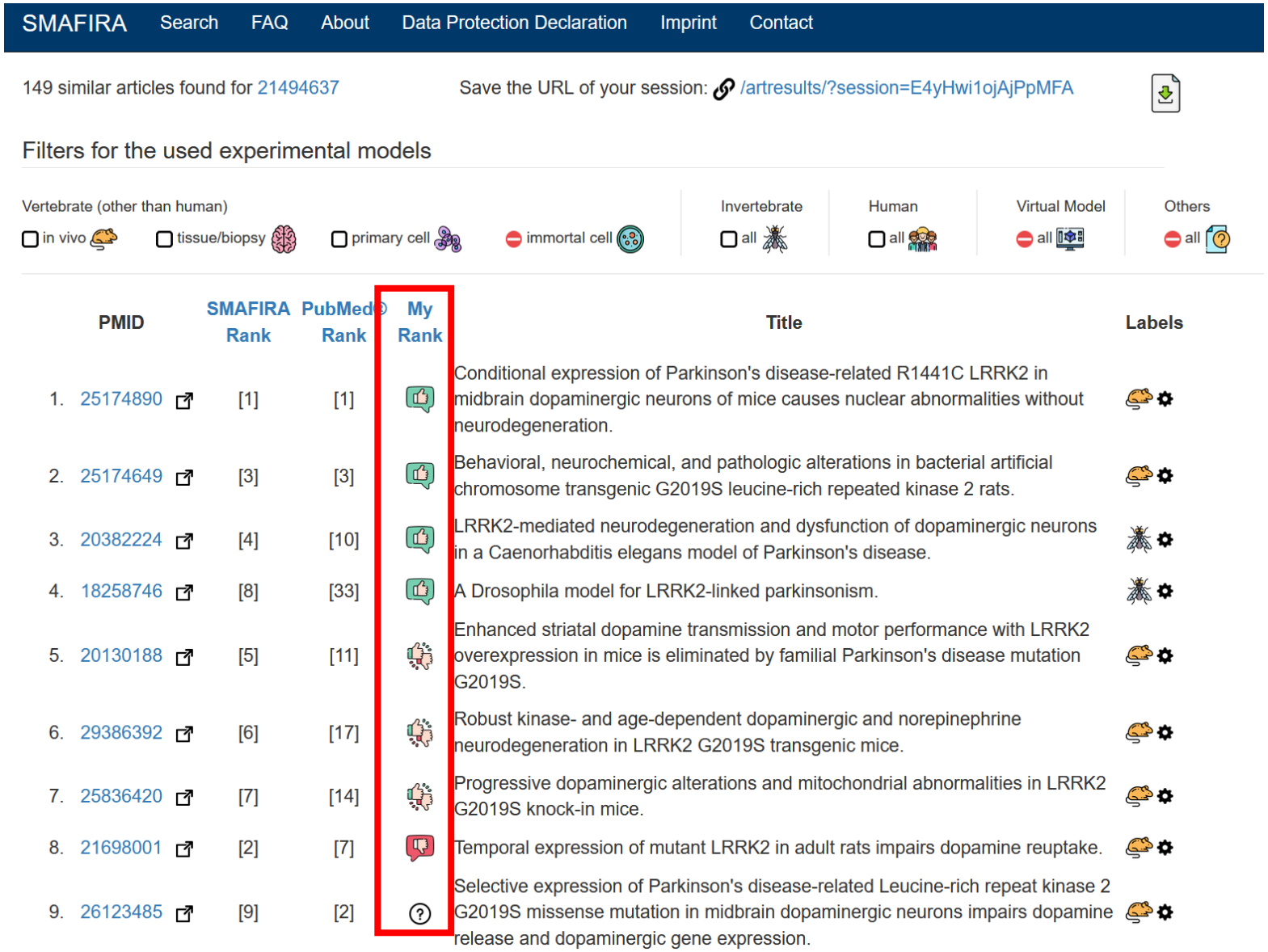
User feedback

My Rank



	PMID	SMAFIRA Rank	PubMed® Rank	My Rank	Title	Labels
1.	25174890	[1]	[1]		Conditional expression of Parkinson's disease-related R1441C LRRK2 in midbrain dopaminergic neurons of mice causes nuclear abnormalities without neurodegeneration.	
2.	21698001	[2]	[7]		Temporal expression of mutant LRRK2 in adult rats impairs dopamine reuptake.	
3.	25174649	[3]	[3]		Behavioral, neurochemical, and pathologic alterations in bacterial artificial chromosome transgenic G2019S leucine-rich repeated kinase 2 rats.	
4.	20382224	[4]	[10]		LRRK2-mediated neurodegeneration and dysfunction of dopaminergic neurons in a Caenorhabditis elegans model of Parkinson's disease.	
5.	20130188	[5]	[11]		Enhanced striatal dopamine transmission and motor performance with LRRK2 overexpression in mice is eliminated by familial Parkinson's disease mutation G2019S.	
6.	29386392	[6]	[17]		Robust kinase- and age-dependent dopaminergic and norepinephrine neurodegeneration in LRRK2 G2019S transgenic mice.	
7.	25836420	[7]	[14]		Progressive dopaminergic alterations and mitochondrial abnormalities in LRRK2 G2019S knock-in mice.	
8.	18258746	[8]	[33]		A Drosophila model for LRRK2-linked parkinsonism.	
9.	26123485	[9]	[2]		Selective expression of Parkinson's disease-related Leucine-rich repeat kinase 2 G2019S missense mutation in midbrain dopaminergic neurons impairs dopamine release and dopaminergic gene expression.	

My Rank



	PMID	SMAFIRA Rank	PubMed® Rank	My Rank	Title	Labels
1.	25174890	[1]	[1]		Conditional expression of Parkinson's disease-related R1441C LRRK2 in midbrain dopaminergic neurons of mice causes nuclear abnormalities without neurodegeneration.	
2.	25174649	[3]	[3]		Behavioral, neurochemical, and pathologic alterations in bacterial artificial chromosome transgenic G2019S leucine-rich repeated kinase 2 rats.	
3.	20382224	[4]	[10]		LRRK2-mediated neurodegeneration and dysfunction of dopaminergic neurons in a Caenorhabditis elegans model of Parkinson's disease.	
4.	18258746	[8]	[33]		A Drosophila model for LRRK2-linked parkinsonism.	
5.	20130188	[5]	[11]		Enhanced striatal dopamine transmission and motor performance with LRRK2 overexpression in mice is eliminated by familial Parkinson's disease mutation G2019S.	
6.	29386392	[6]	[17]		Robust kinase- and age-dependent dopaminergic and norepinephrine neurodegeneration in LRRK2 G2019S transgenic mice.	
7.	25836420	[7]	[14]		Progressive dopaminergic alterations and mitochondrial abnormalities in LRRK2 G2019S knock-in mice.	
8.	21698001	[2]	[7]		Temporal expression of mutant LRRK2 in adult rats impairs dopamine reuptake.	
9.	26123485	[9]	[2]		Selective expression of Parkinson's disease-related Leucine-rich repeat kinase 2 G2019S missense mutation in midbrain dopaminergic neurons impairs dopamine release and dopaminergic gene expression.	

Classification of experimental models

BfR

GoldHamster corpus

- **1,600 annotated abstracts**

Mariana Neves, Antonina Klippert, Fanny Knöspel, Juliane Rudeck, Ailine Stolz, Zsofia Ban, Markus Becker, Kai Diederich, Barbara Grune, Pia Kahnau, Nils Ohnesorge, Johannes Pucher, Gilbert Schönfelder, Bettina Bert & Daniel Butzke

	Total	No. Docs	Labels
1 label	862	213	"in vivo"
		156	"invertebrates"
		136	"human"
		97	"others"
		92	"organs"
		85	"in silico"
		30	"immortal cell lines"
		28	"primary cell lines"
		25	none
2 labels	31	6	"in silico+in vivo"
		5	"in vivo+primary cell lines"
		4	"in silico+invertebrates", or "immortal cell lines+in vivo", or "in vivo+organs"
		2	"organs+primary cell lines", or "in silico+primary cell lines", or "human+invertebrates", or "in silico+others"
3 labels	6	1	"human+in vivo+primary cell lines", or "in vivo+invertebrates", or "immortal cell lines+primary cell lines", or "in vivo+others", or "in silico+organs", or "immortal cell lines+organs"

Labels	Kappa	Level Of agreement
invertebrates	0.82	almost perfect
in vivo	0.78	substantial
in silico	0.72	substantial
human	0.63	substantial
organs	0.62	substantial
immortal cell lines	0.49	moderate
primary cell lines	0.45	moderate
others	0.42	moderate
overall	0.62	substantial

Performance of the model

- **Transformers pre-trained on PubMedBERT (BiomedBERT) and fine-tuned on the GoldHamster corpus**

Labels	BioBERT	PubMedBERT	SciBERT	BLUE-BERT
invertebrates	0.89	**0.95**	**0.95**	**0.95**
in_vivo	0.88	0.88	**0.89**	0.83
human	0.82	**0.86**	0.82	0.73
organs	0.80	**0.82**	**0.82**	0.71
primary_cells	0.67	**0.75**	0.73	0.67
immortal_cell_lines	**0.91**	0.83	0.80	0.89
in_silico	0.67	0.75	0.75	**0.86**
others	0.70	**0.78**	0.67	0.76
All (average)	0.79	**0.83**	0.80	0.80

Hugging Face

microsoft / **BiomedNLP-BiomedBERT-base-uncased-abstract-fulltext**

Current work

“biomed-approach” corpus

- **> 3000 annotated abstracts**
- **23 disease categories (MeSH terms)**
- **Additional labels: human immortal cell lines, human primary cells, human organs/tissues, clinical studies**
- **Promising preliminary results**

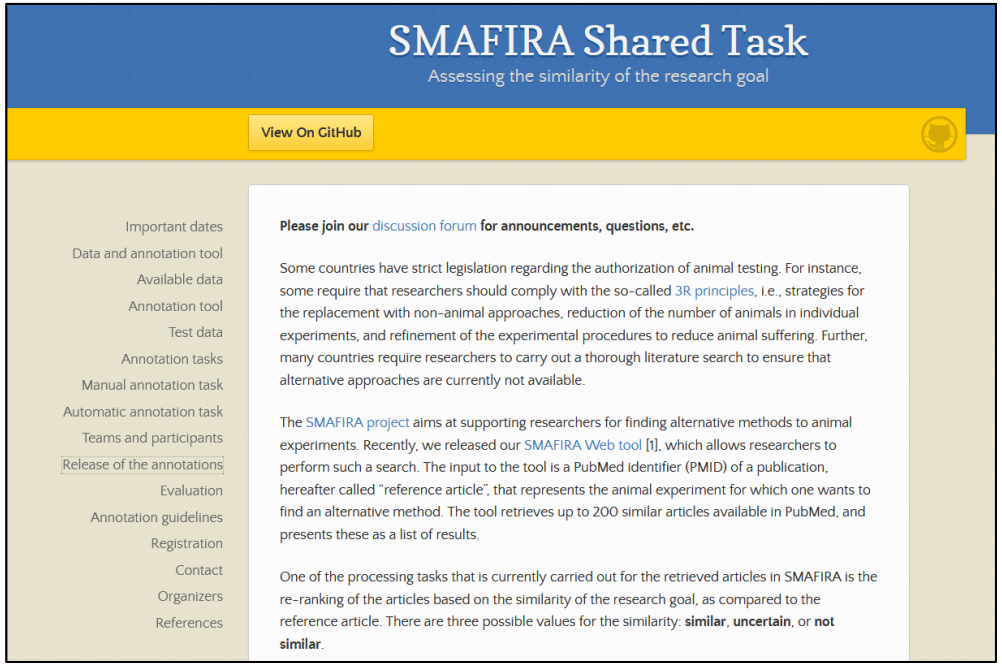
SMAFIRA Shared task

https://smafira-bf3r.github.io/smafira-st/

- **Call for participants!**

- **Annotation for selected reference articles (top 20)**
- **Training/evaluation set for re-ranking**

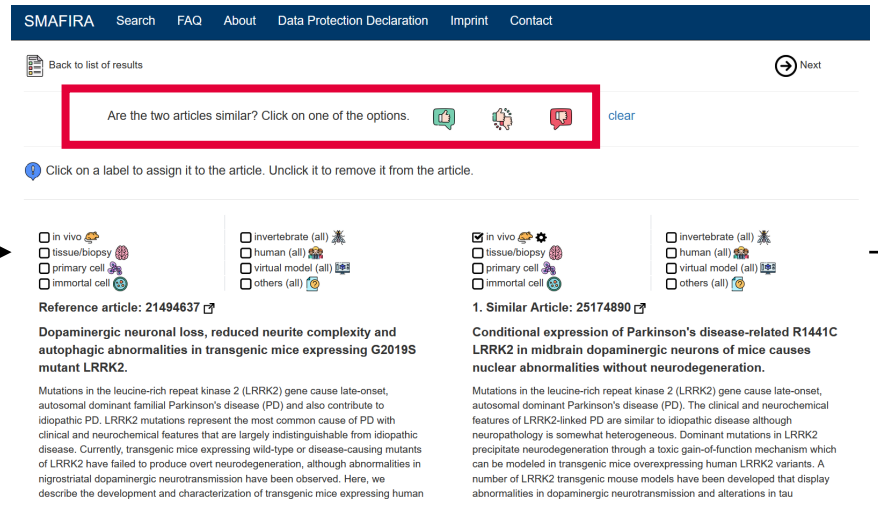
- **Manual (SMAFIRA Tool)**
- **Automatic (systems)**

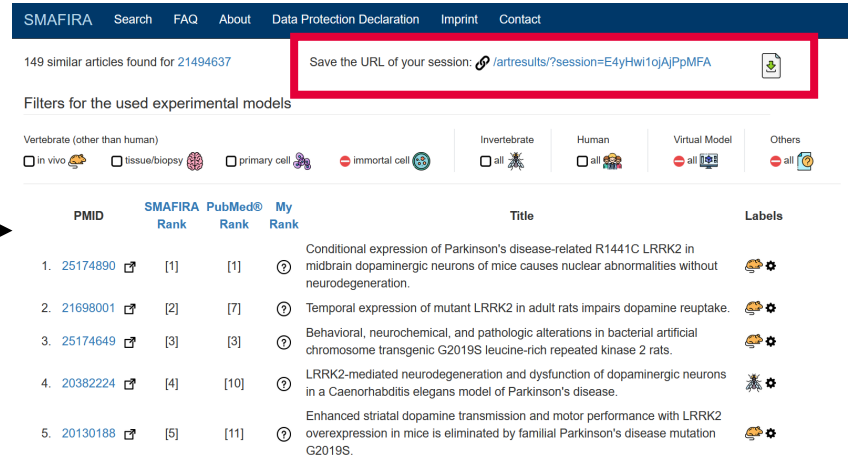


SMAFIRA Shared task

https://smafira-bf3r.github.io/smafira-st/

Infections	Neoplasms	Nervous System Diseases	Cardiovascular Diseases	Immune System Diseases
36159784	34233949	35709748	33635944	34503569
36577999	33320838	37084732	37010266	36179018
32485164	36311701	37339207	37380648	37079985
37071015	37429473	37749256	37268711	37256935
31689515	35623658	37126714	35917178	37168850





https://smafira-bf3r.github.io/smafira-st/

SMAFIRA Shared task

- **Publications**
  - BioNLP Workshop (63rd Annual Meeting of the Association for Computational Linguistics , Vienna, Aug 1st 2025)
  - Joint paper in a journal (end of 2025)

- Contact: **mariana.lara-neves@bfr.bund.de**

BIONLP 2025 and Shared Tasks @ ACL 2025

German Federal Institute for Risk Assessment

Mariana Neves
T +49 30 18412-29411
mariana.lara-neves@bfr.bund.de

German Federal Institute for Risk Assessment
bfr.bund.de/en

BfR | Identifying Risks – Protecting Health

Consumer health protection to go

BfR2GO – the BfR Science Magazine

bfr.bund.de/en/science_magazine_bfr2go.html

Follow us

- @bfrde | @bfren | @Bf3R_centre
- @bfrde
- youtube.com/@bfr_bund
- social.bund.de/@bfr
- linkedin.com/company/bundesinstitut-f-r-risikobewertung
- soundcloud.com/risikobewertung
- threads.net/@bfrde
- bsky.app/profile/bfrde.bsky.social