Practical Intro to Git(Hub)

and AncientMetagenomeDir

By James A. Fellows Yates

Slides: https://hackmd.io/@jfy133/H19kmDaIw

Scope

- Not technical
  - i.e. not for computer-scientists
- Covers core functionality: max 3-5 ‘concepts’
  - All you need to know!
- Primarily through github website
  - No need for command-line shenanigans!

Concepts

- Navigation
- Branches
- Commit
- Pull Request and Review
- Merging

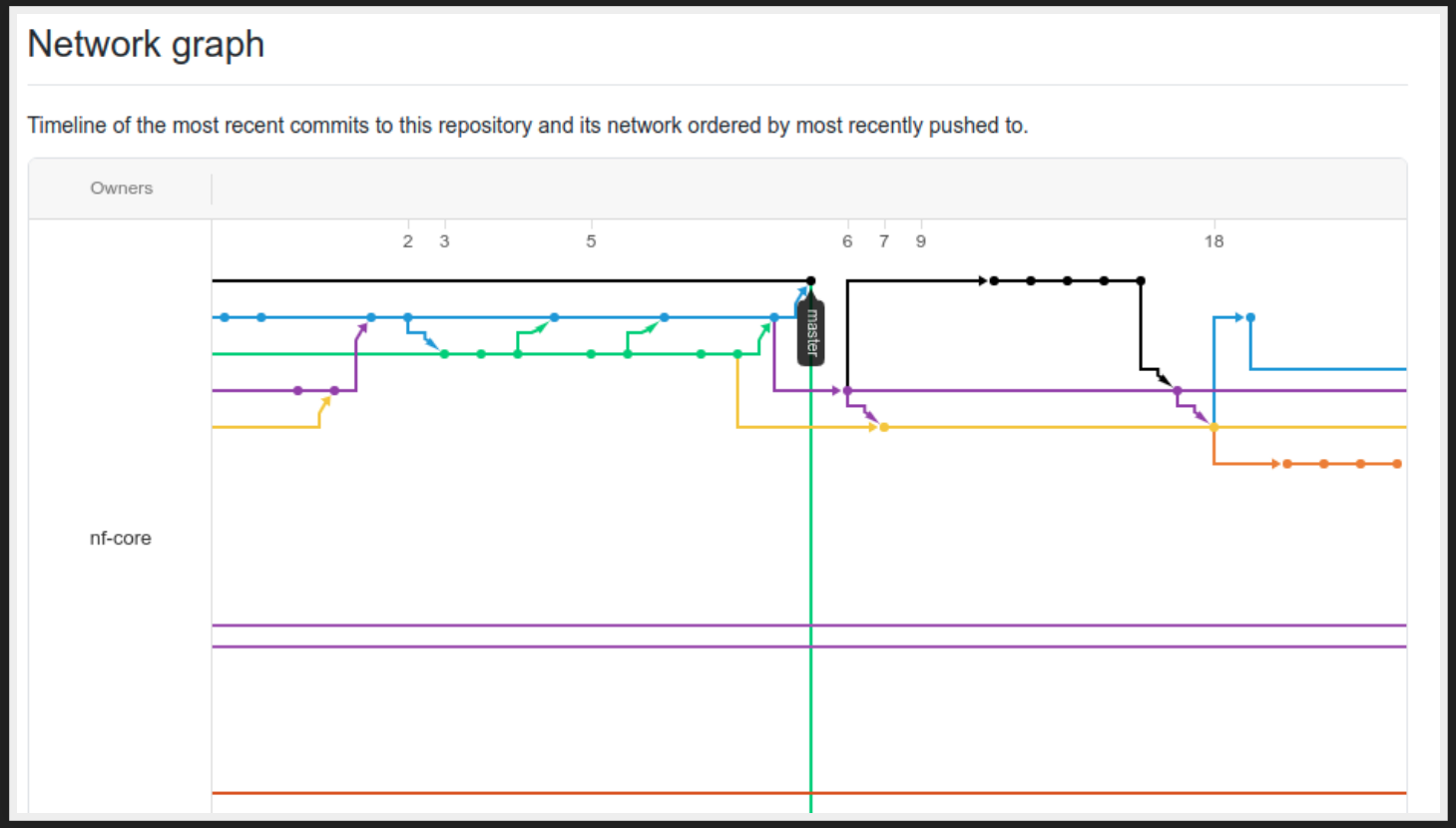
What is Git?

“Git is a free and open source distributed version control system designed to handle everything from small to very large projects with speed and efficiency.” - Git Website

What is Git(Hub)?

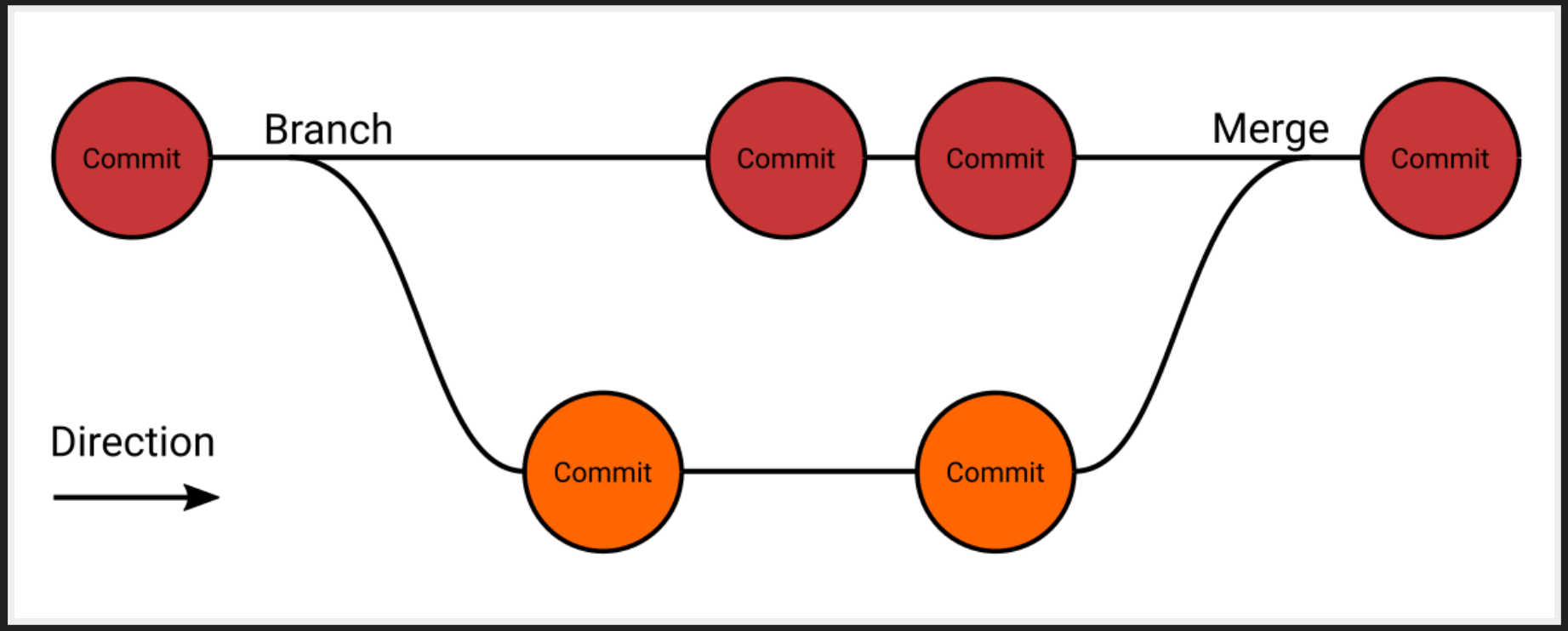
“GitHub, Inc. is a United States-based global company that provides hosting for software development and version control using Git. It has been a subsidiary of Microsoft since 2018. It offers the distributed version control and source code management functionality of Git, plus its own features.” - Wikipedia

What it can look like



From nf-core/eager

What it essentially is



What this means

- **Branch**: make a copy of ‘main’ code
  - work here at your own place in your own ‘playground’
- **Commit**: a ‘saved’ edit
  - Each has unique ‘hash’ for version control
- **Pull Request**:
  - Via a ‘pull request’ on GitHub
  - Peer reviewed, to check edits makes sense!
- **Merge**:
  - Integrating your edits in the ‘main’ code

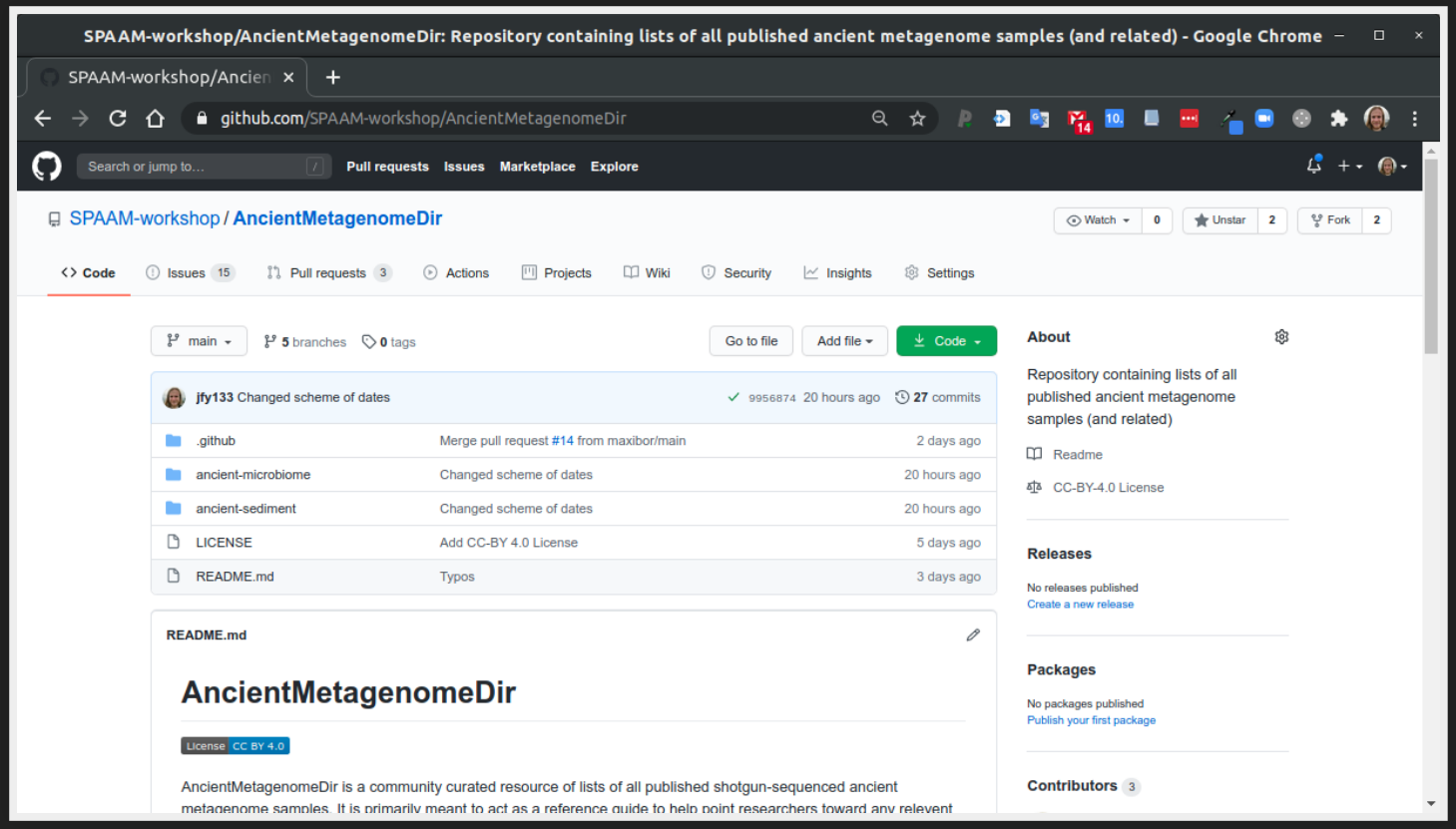
What is AncientMetagenomeDir?

- Community-curated list of all published ancient metagenome samples
- Contains basic metadata for summary statistics and pointers to data
- Standardised, versioned and citable!

What is AncientMetagenomeDir?

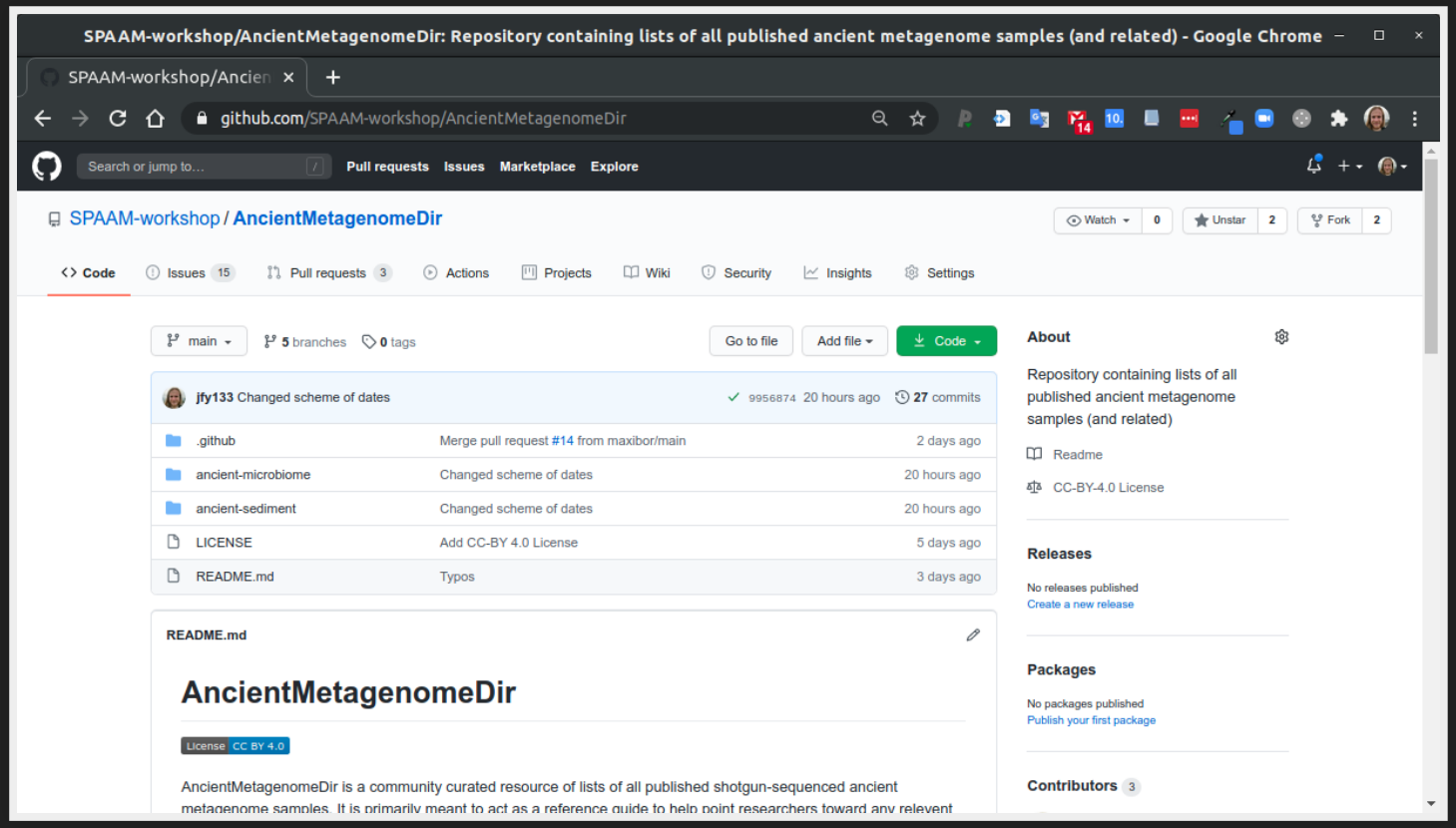
- A set of TSV files with semi-standardised column fields
- Covers:
  - Publication
  - Sample Name
  - Sample Type
  - Age
  - Geographic Location
  - Archive Accessions IDs (not data!)

Where is AncientMetagenomeDir?

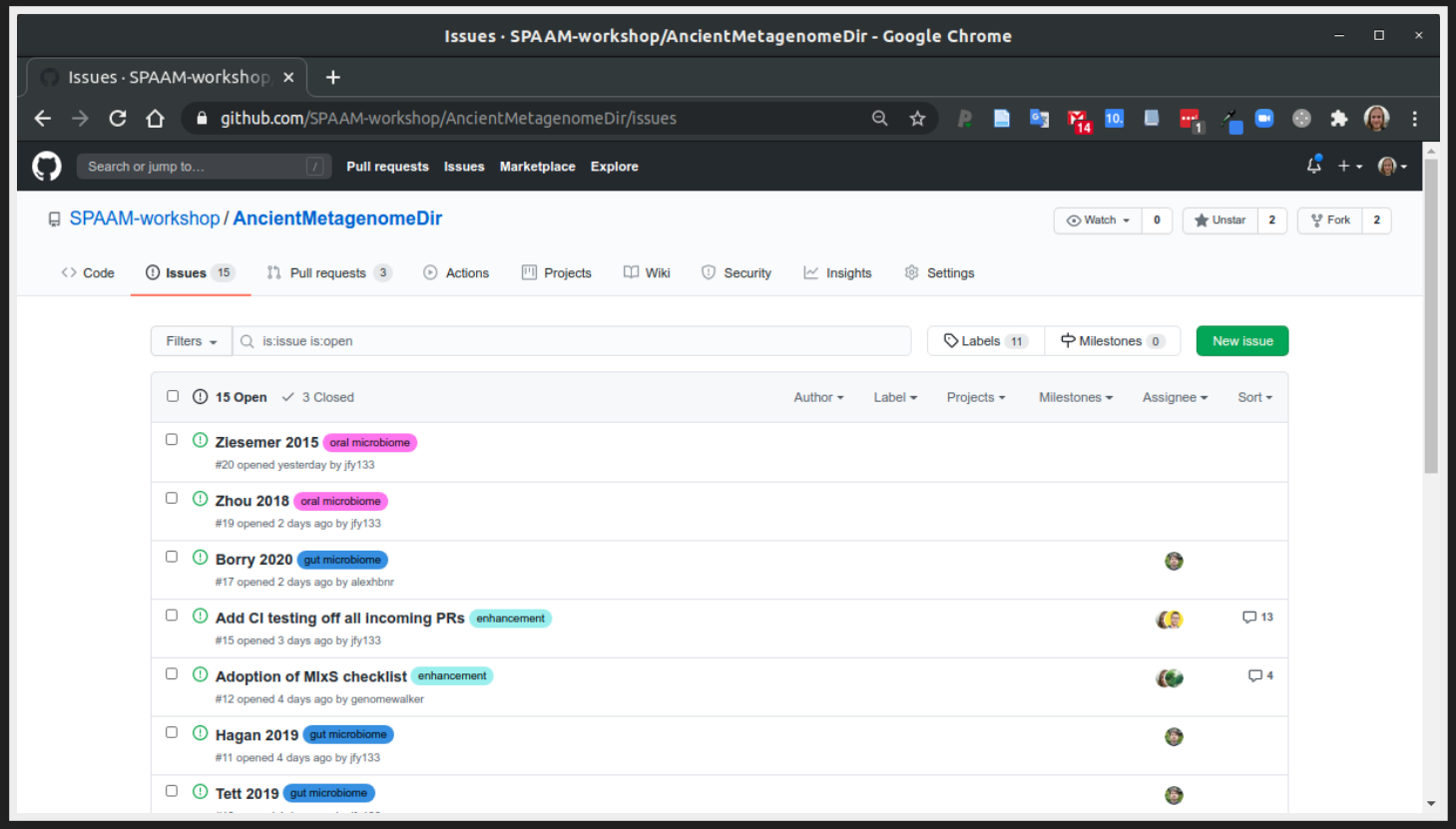


Click ☝️

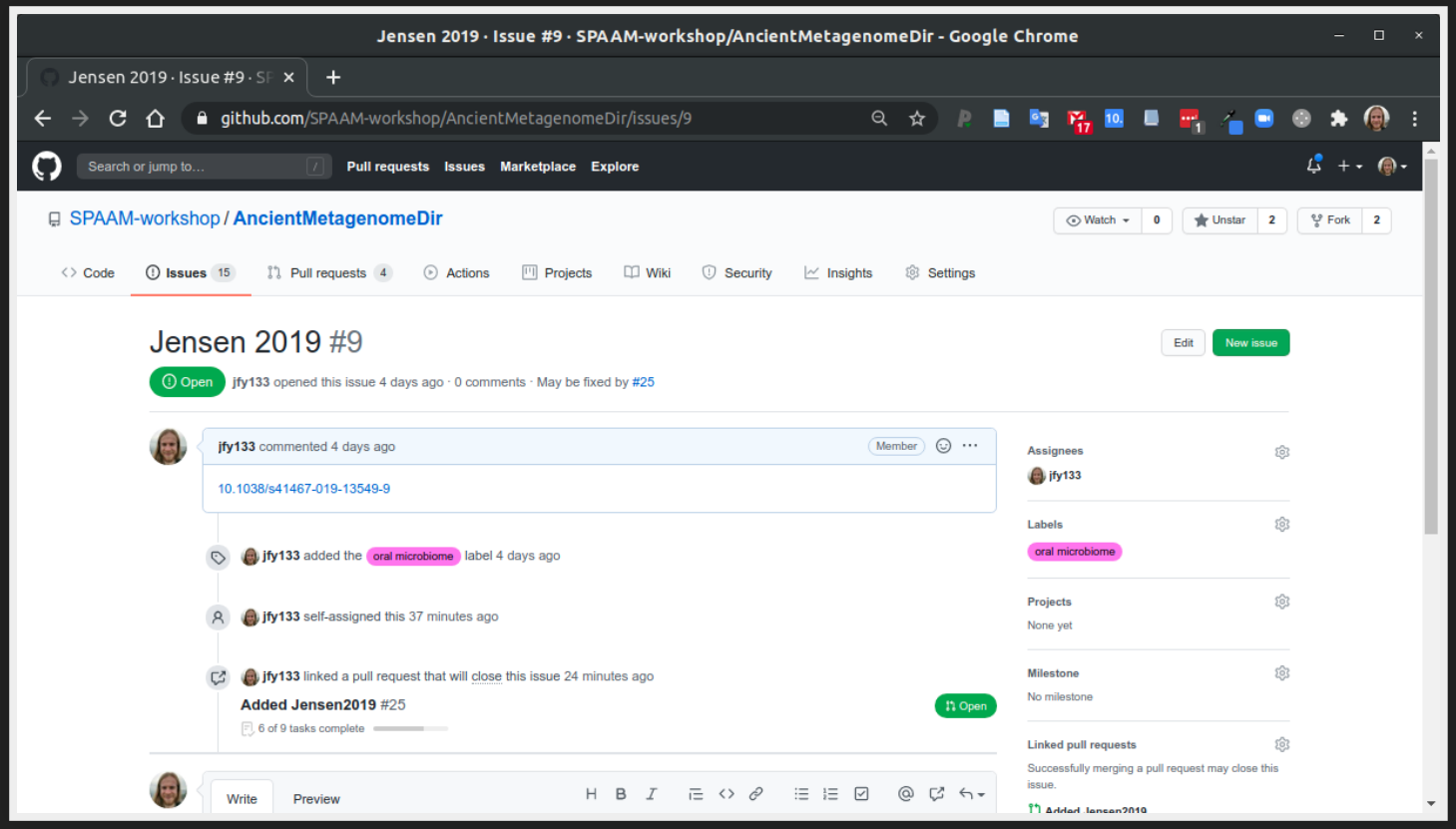
Navigation: Code Tab and README



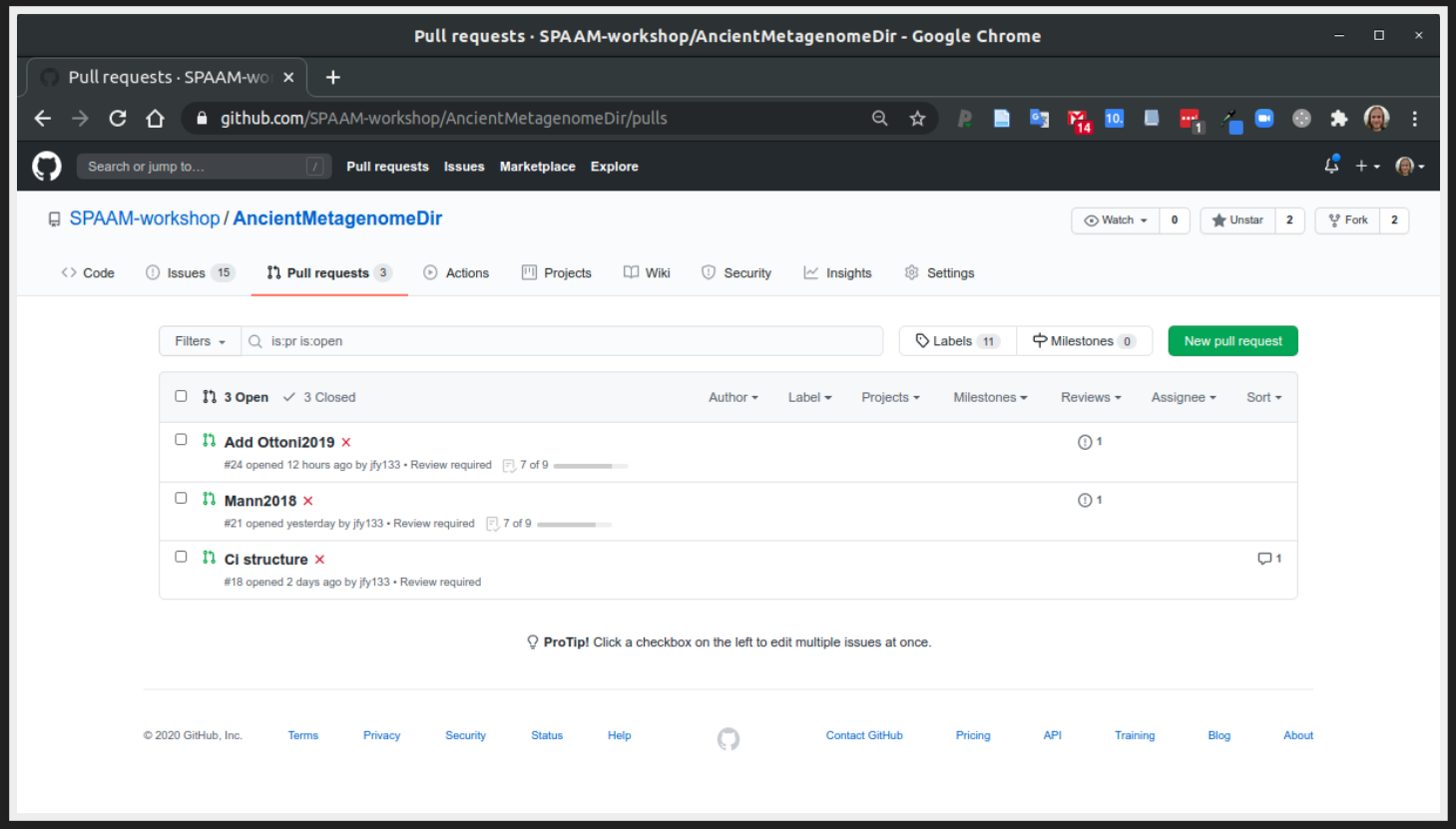
Issue: Issues Tab



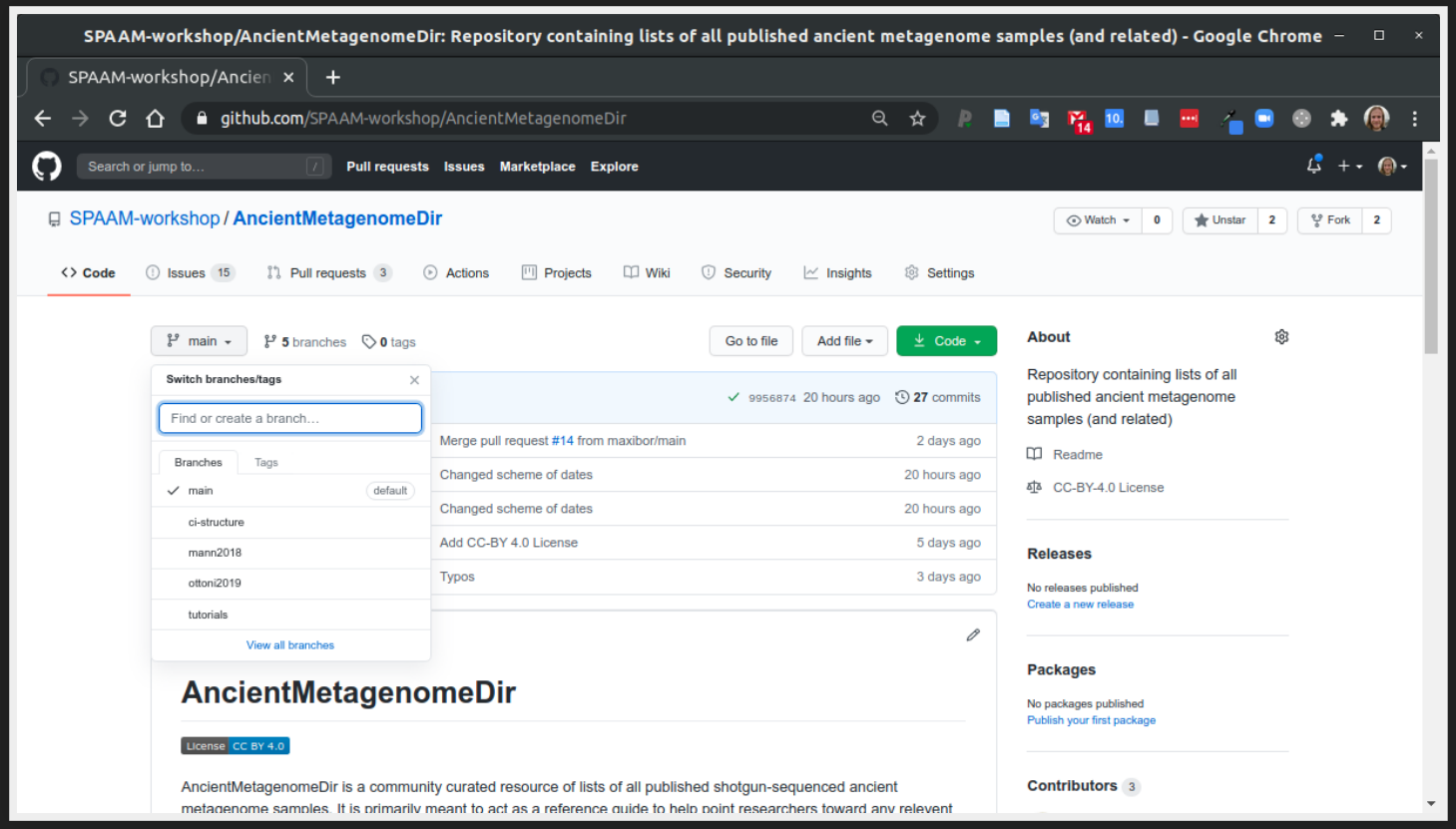
Issue: Inside an Issue



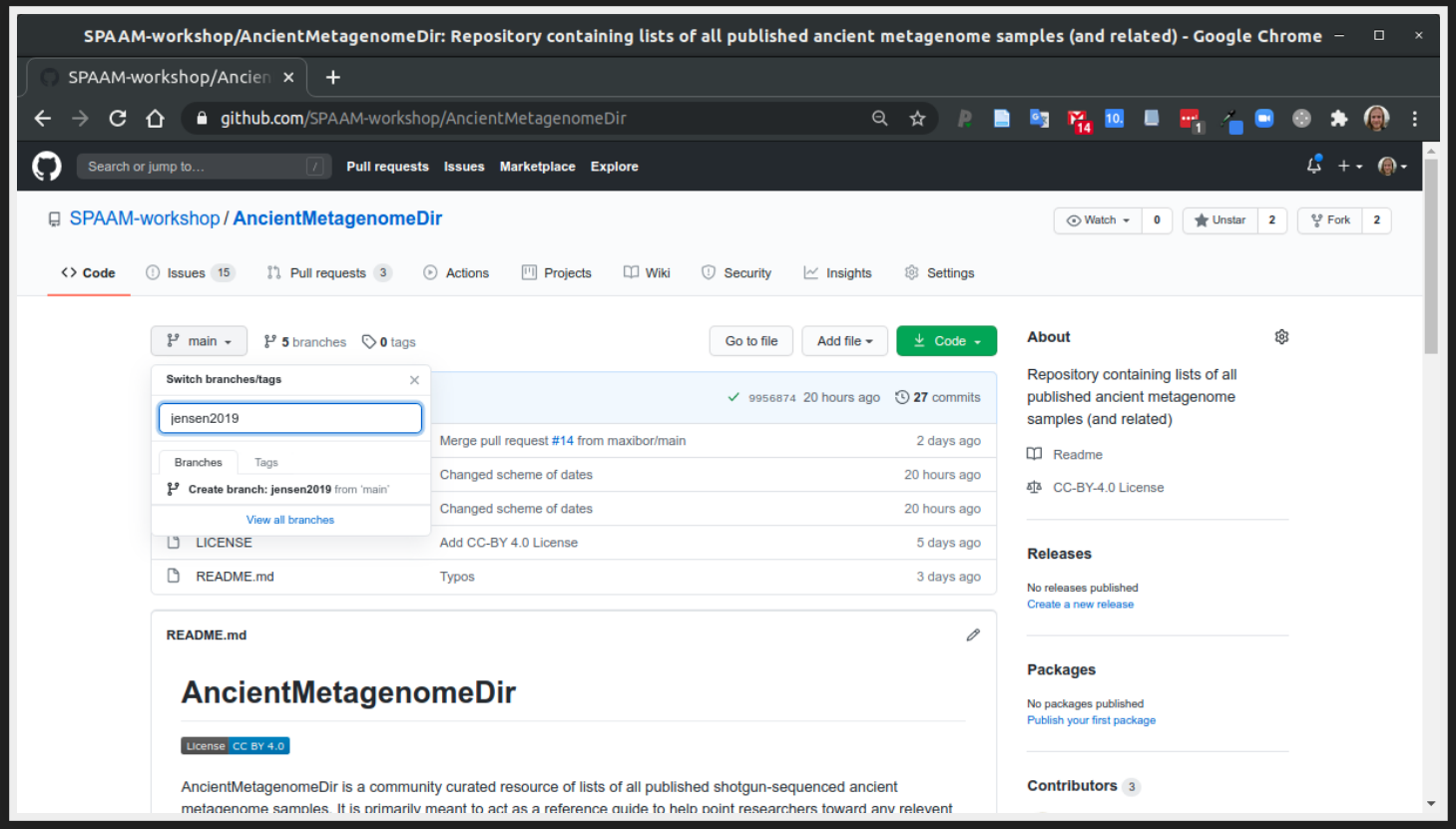
Pull Request: Pull Request Tab



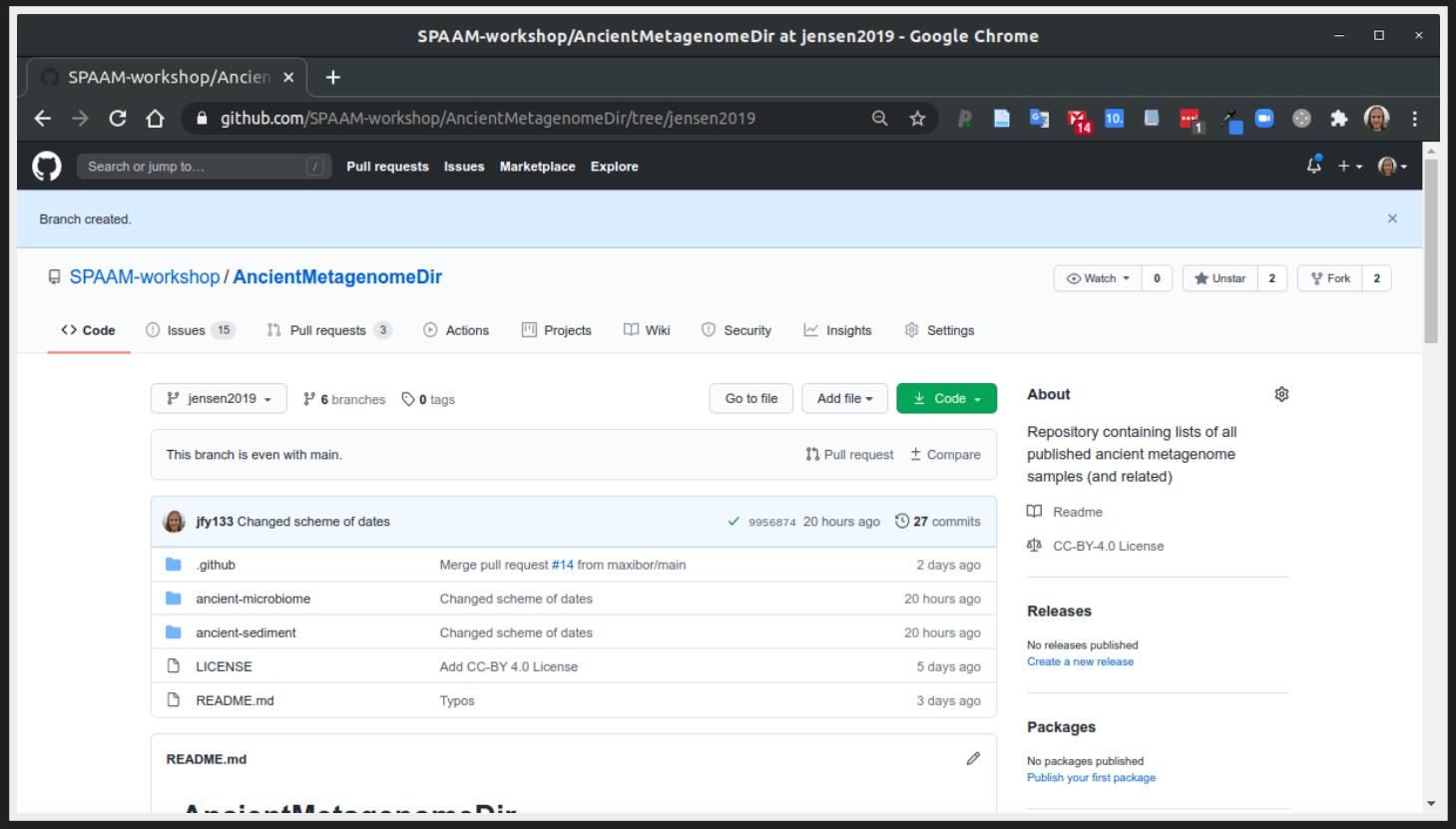
Branch: Accessing Branches



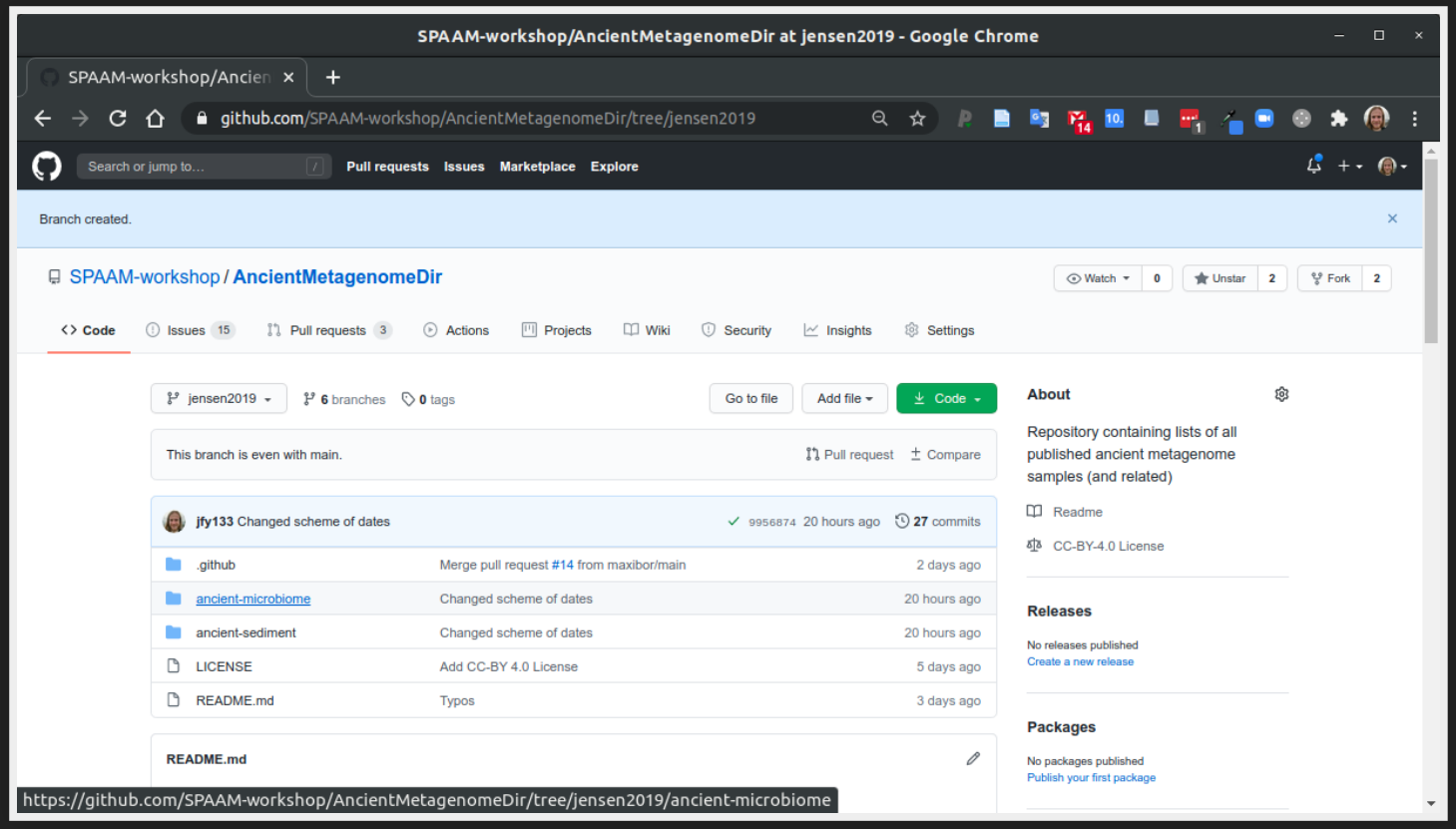
Branch: Making New Branch



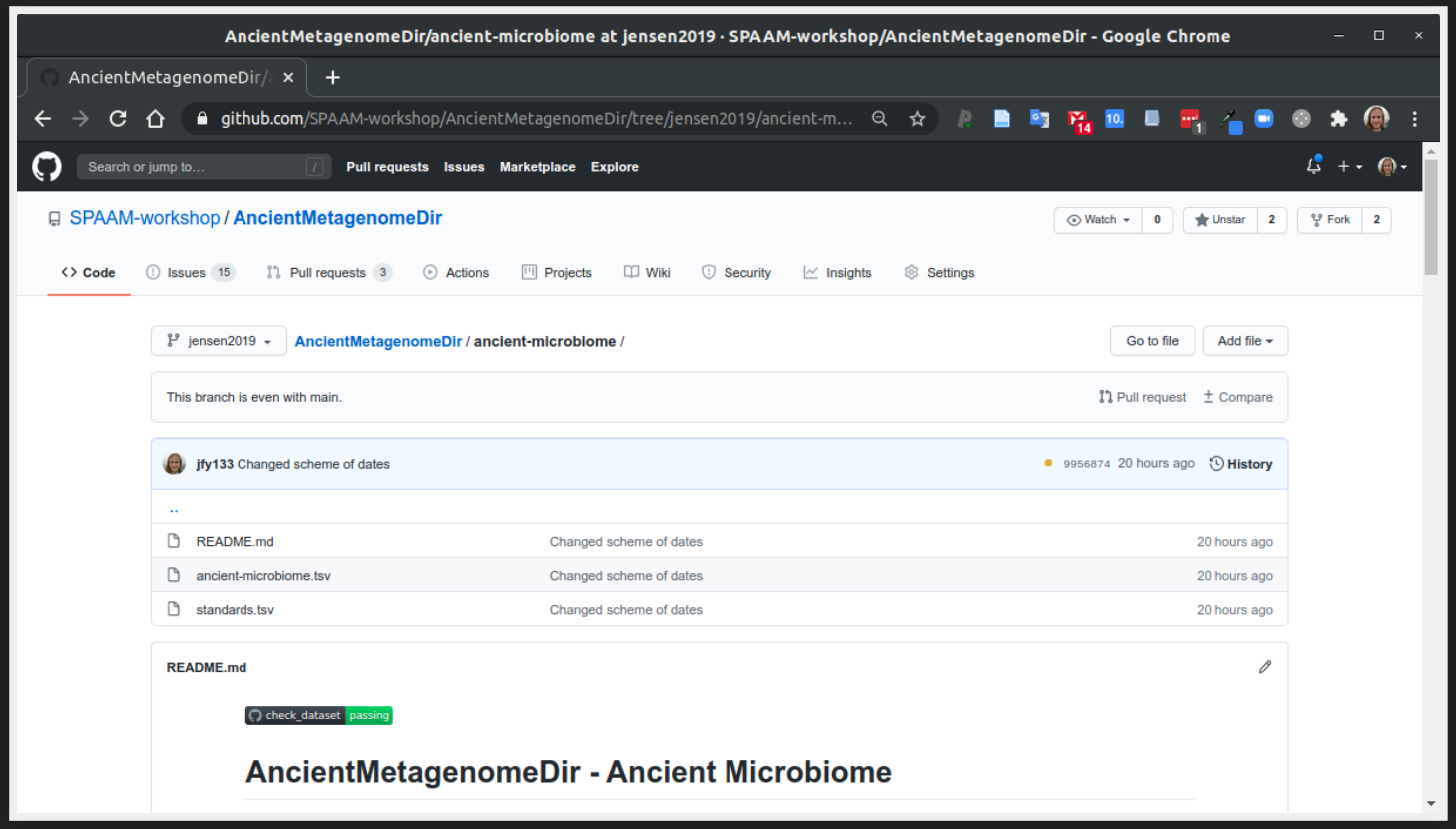
Branch: New Branch Success



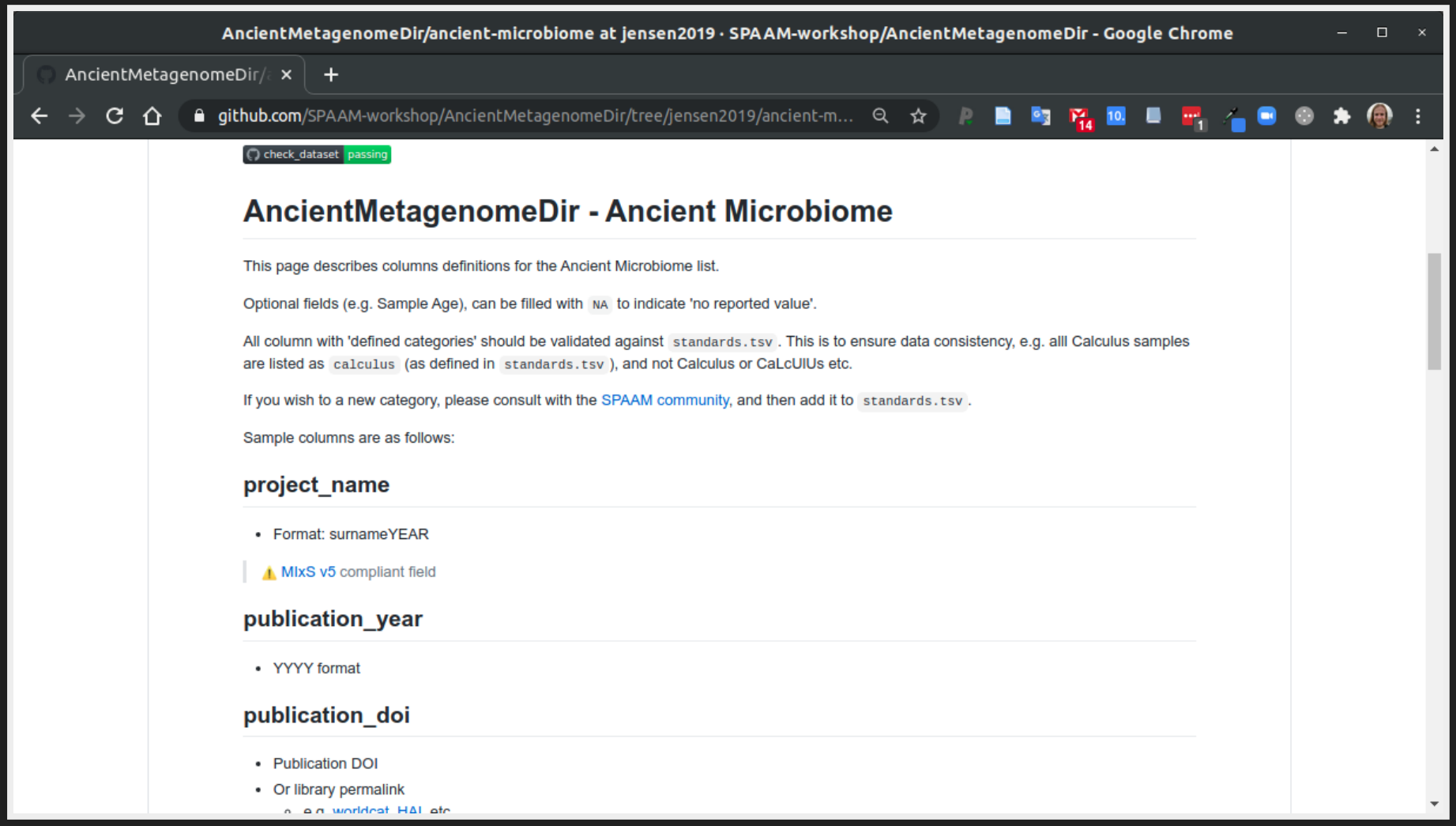
Navigation: Accessing Directory



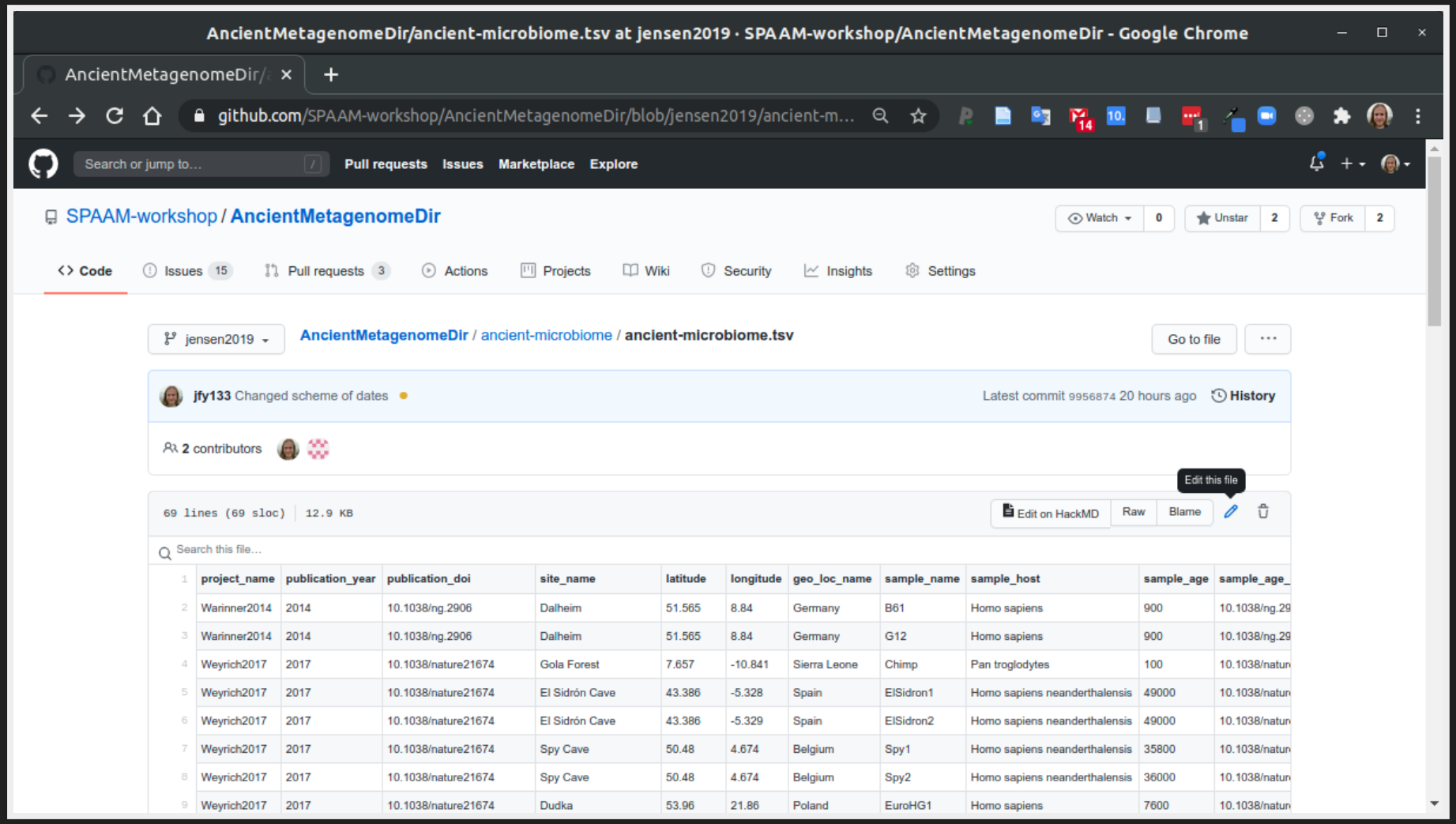
Navigation: Inside Directory



Navigation: Internal README

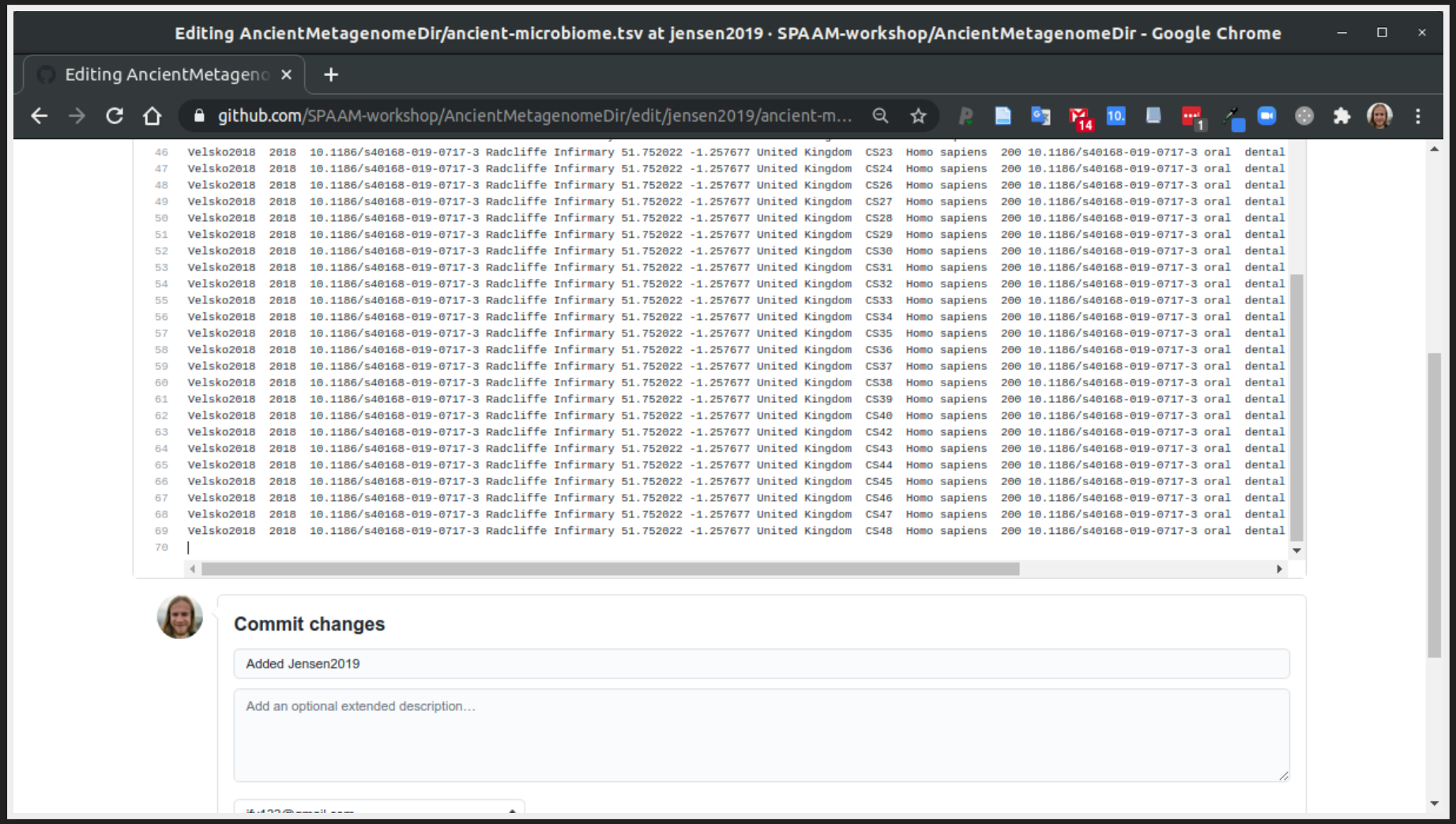


Navigation: Opening a File

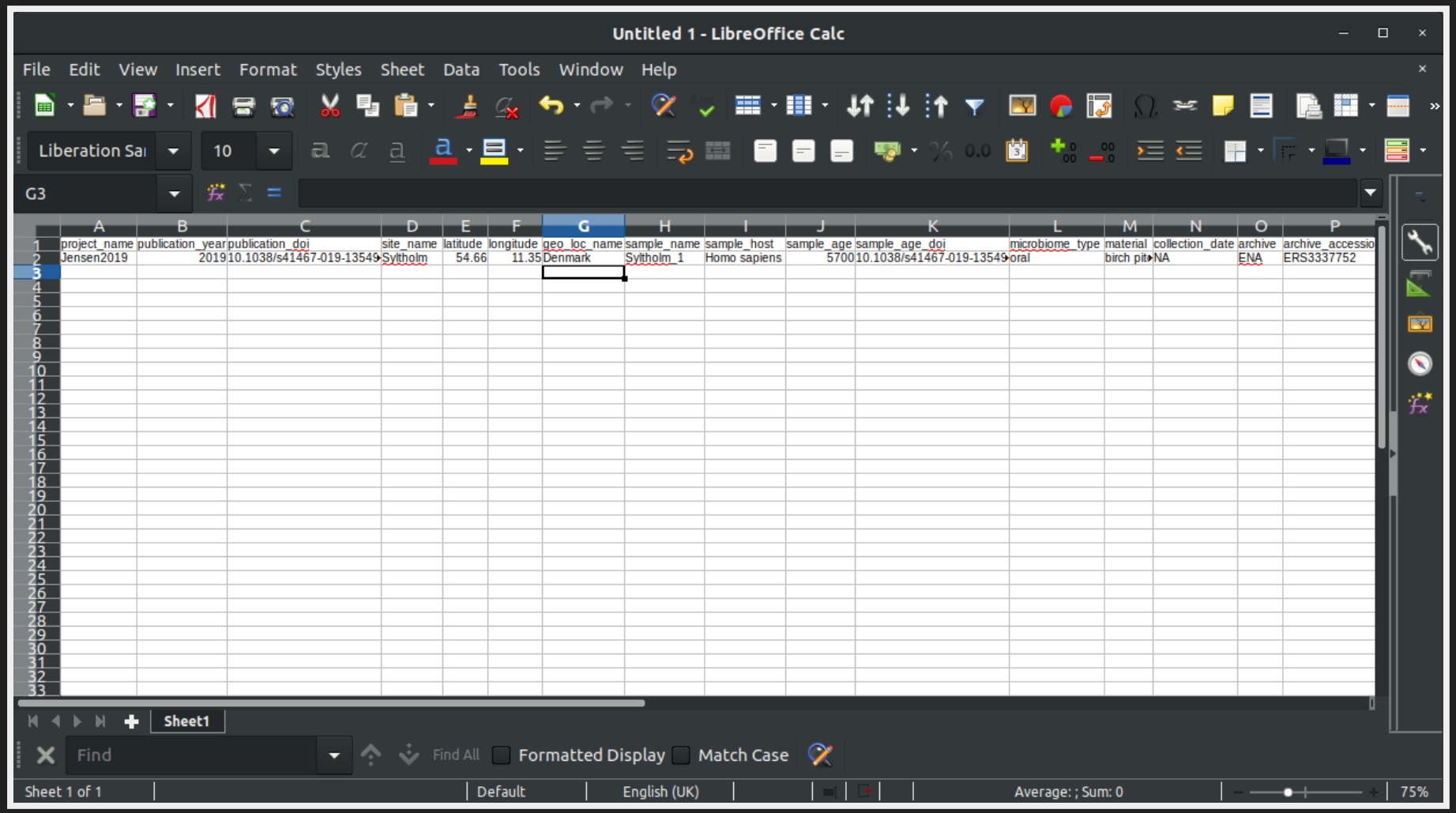


	project_name	publication_year	publication_doi	site_name	latitude	longitude	geo_loc_name	sample_name	sample_host	sample_age	sample_age_
2	Warinner2014	2014	10.1038/ng.2906	Dalheim	51.565	8.84	Germany	B61	Homo sapiens	900	10.1038/ng.29
3	Warinner2014	2014	10.1038/ng.2906	Dalheim	51.565	8.84	Germany	G12	Homo sapiens	900	10.1038/ng.29
4	Weyrich2017	2017	10.1038/nature21674	Gola Forest	7.657	-10.841	Sierra Leone	Chimp	Pan troglodytes	100	10.1038/natur
5	Weyrich2017	2017	10.1038/nature21674	El Sidrón Cave	43.386	-5.328	Spain	ElSidron1	Homo sapiens neanderthalensis	49000	10.1038/natur
6	Weyrich2017	2017	10.1038/nature21674	El Sidrón Cave	43.386	-5.329	Spain	ElSidron2	Homo sapiens neanderthalensis	49000	10.1038/natur
7	Weyrich2017	2017	10.1038/nature21674	Spy Cave	50.48	4.674	Belgium	Spy1	Homo sapiens neanderthalensis	35800	10.1038/natur
8	Weyrich2017	2017	10.1038/nature21674	Spy Cave	50.48	4.674	Belgium	Spy2	Homo sapiens neanderthalensis	36000	10.1038/natur
9	Weyrich2017	2017	10.1038/nature21674	Dudka	53.96	21.86	Poland	EuroHG1	Homo sapiens	7600	10.1038/natur

Commit: Editing a File

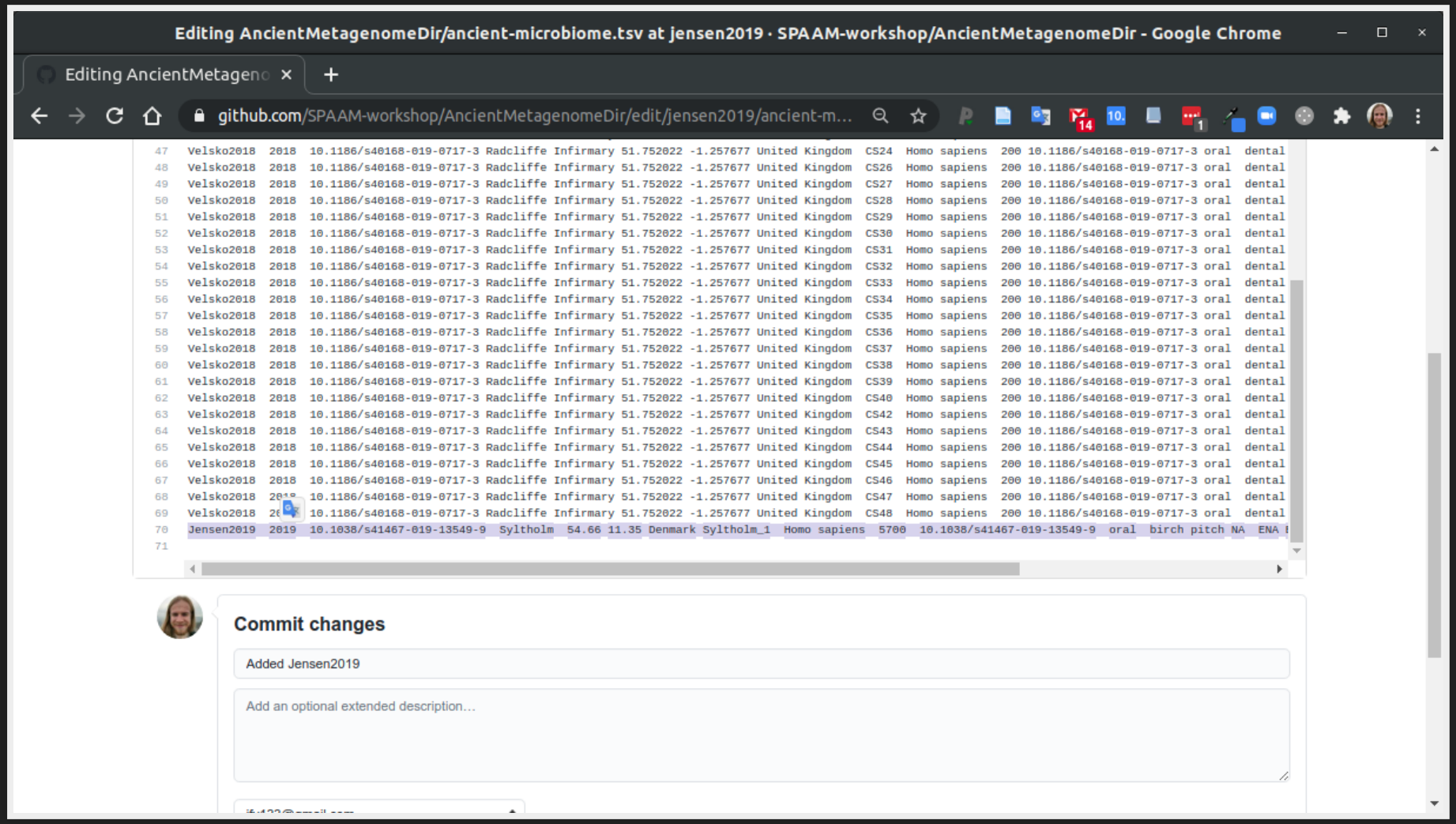


Commit: Prepare Data

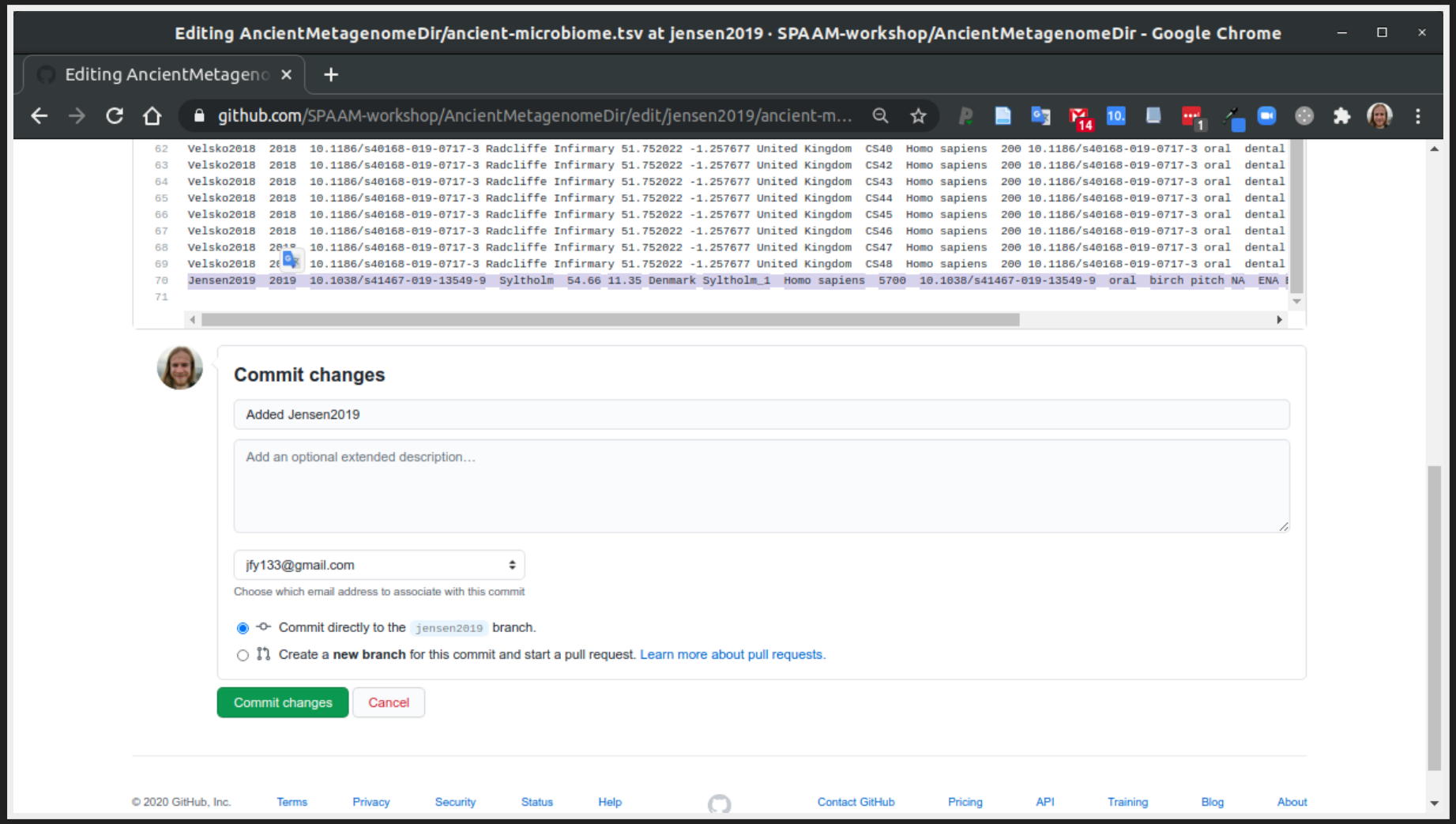


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	project_name	publication_year	publication_doi	site_name	latitude	longitude	geo_loc_name	sample_name	sample_host	sample_age	sample_age_doi	microbiome_type	material	collection_date	archive	archive_accessio
2	Jensen2019	2019	10.1038/s41467-019-13549	Sylthoim	54.66	11.35	Denmark	Sylthoim_1	Homo sapiens	5700	10.1038/s41467-019-13549	oral	birch pit	NA	ENA	ERS3337752

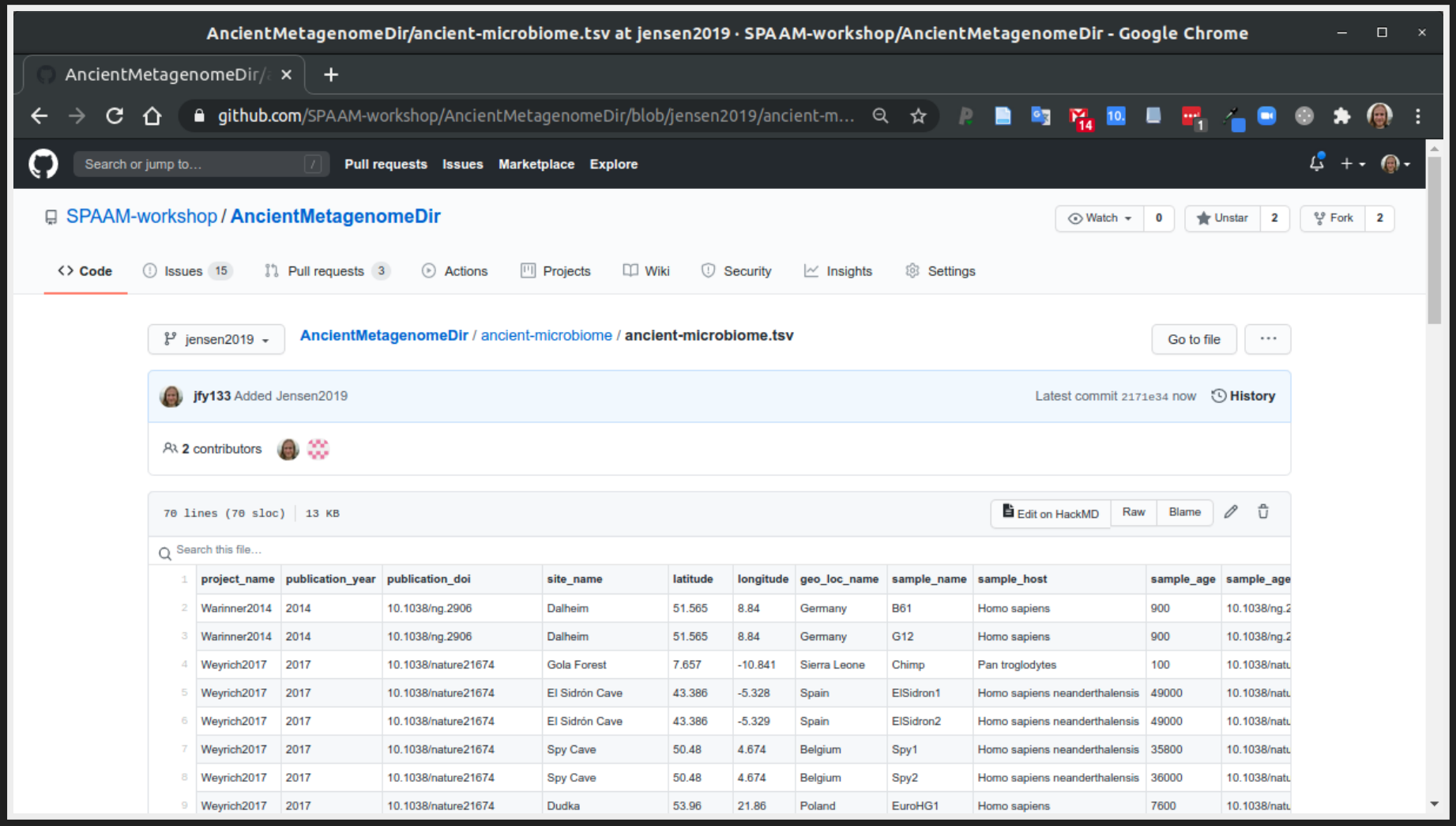
Commit: Update File



Commit: Commit Commit

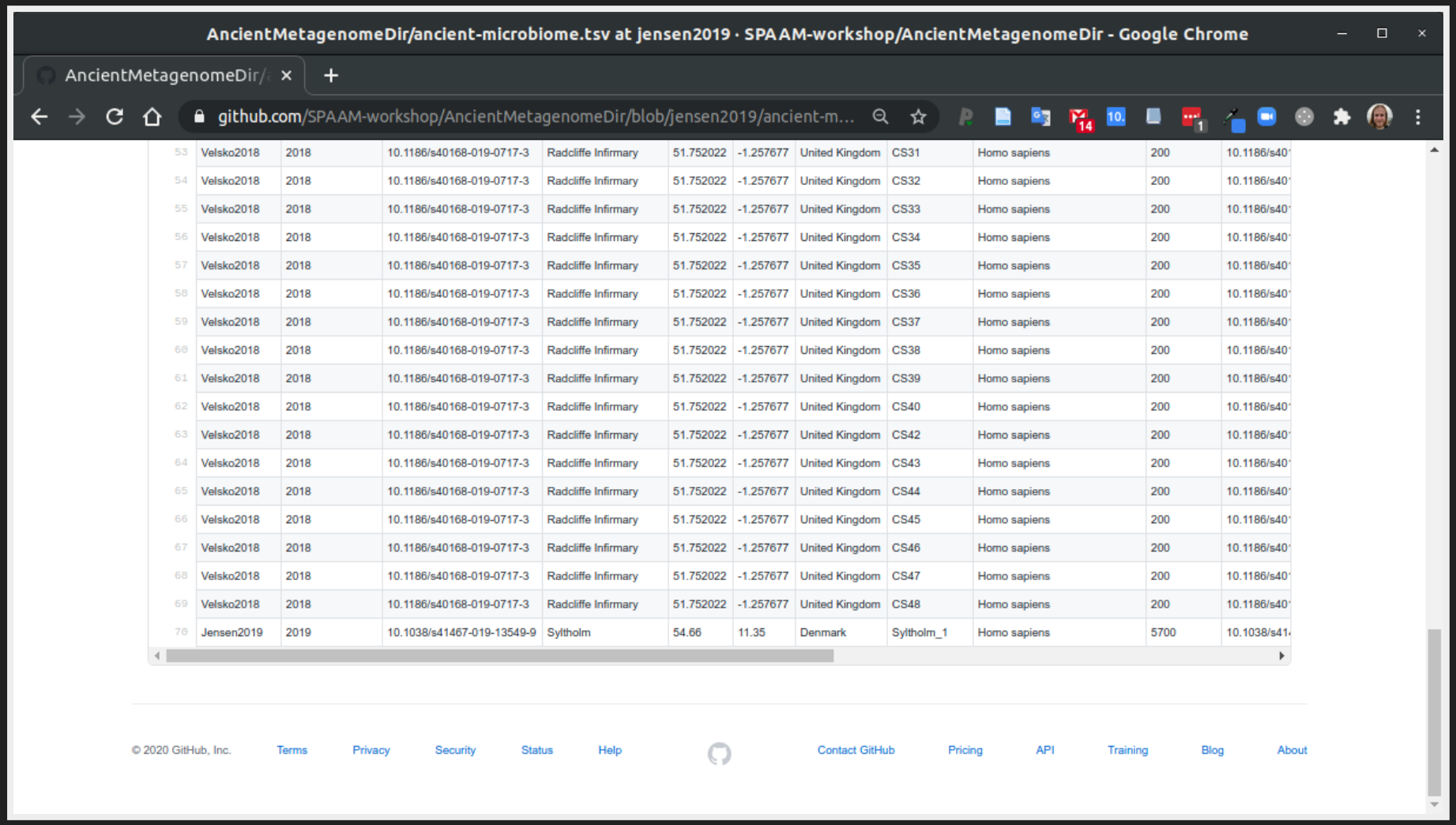


Commit: Commit Committed



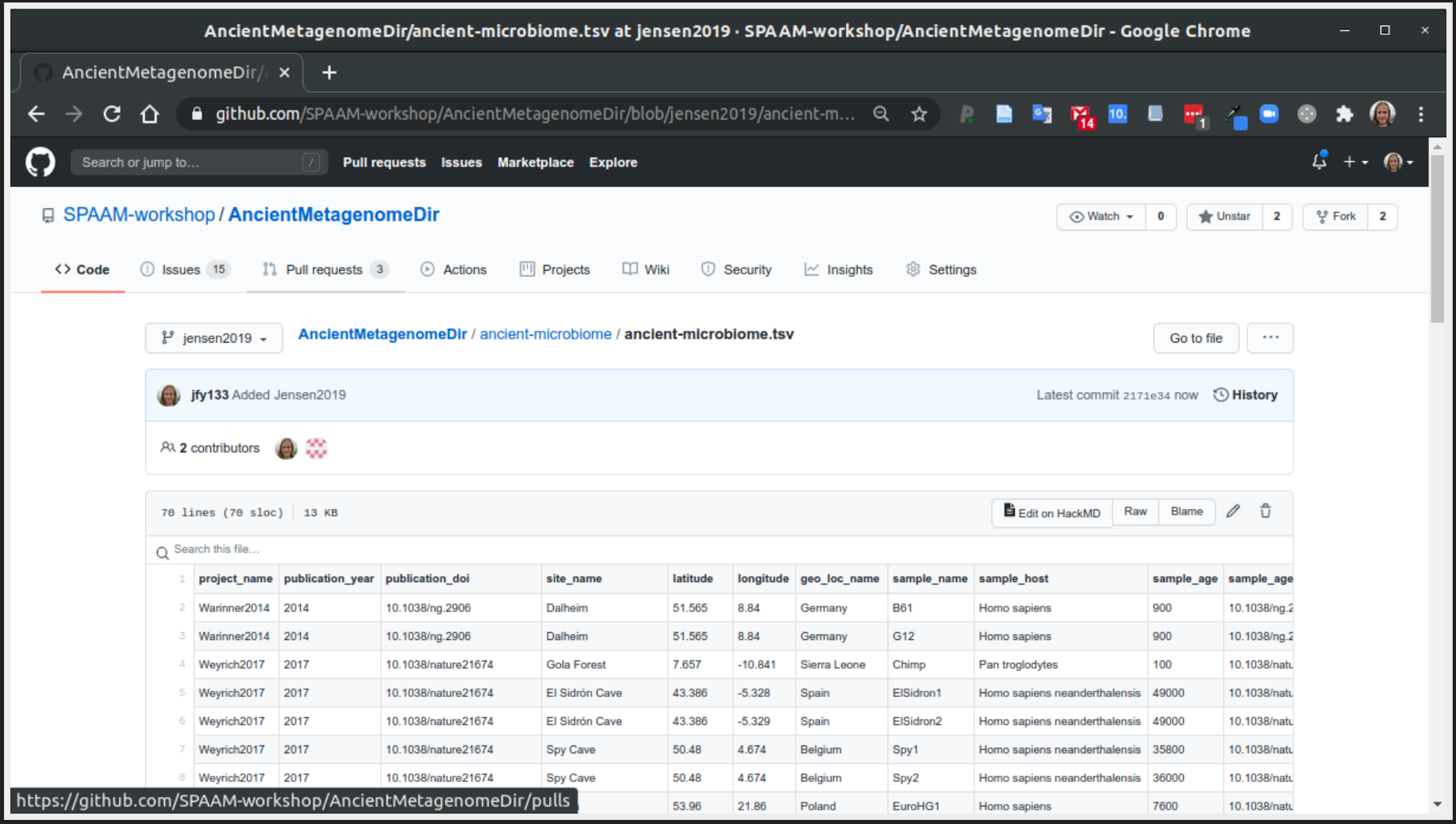
	project_name	publication_year	publication_doi	site_name	latitude	longitude	geo_loc_name	sample_name	sample_host	sample_age	sample_age
2	Warinner2014	2014	10.1038/ng.2906	Dalheim	51.565	8.84	Germany	B61	Homo sapiens	900	10.1038/ng.2
3	Warinner2014	2014	10.1038/ng.2906	Dalheim	51.565	8.84	Germany	G12	Homo sapiens	900	10.1038/ng.2
4	Weyrich2017	2017	10.1038/nature21674	Gola Forest	7.657	-10.841	Sierra Leone	Chimp	Pan troglodytes	100	10.1038/natu
5	Weyrich2017	2017	10.1038/nature21674	El Sidrón Cave	43.386	-5.328	Spain	ElSidron1	Homo sapiens neanderthalensis	49000	10.1038/natu
6	Weyrich2017	2017	10.1038/nature21674	El Sidrón Cave	43.386	-5.329	Spain	ElSidron2	Homo sapiens neanderthalensis	49000	10.1038/natu
7	Weyrich2017	2017	10.1038/nature21674	Spy Cave	50.48	4.674	Belgium	Spy1	Homo sapiens neanderthalensis	35800	10.1038/natu
8	Weyrich2017	2017	10.1038/nature21674	Spy Cave	50.48	4.674	Belgium	Spy2	Homo sapiens neanderthalensis	36000	10.1038/natu
9	Weyrich2017	2017	10.1038/nature21674	Dudka	53.96	21.86	Poland	EuroHG1	Homo sapiens	7600	10.1038/natu

Commit: Check Commit



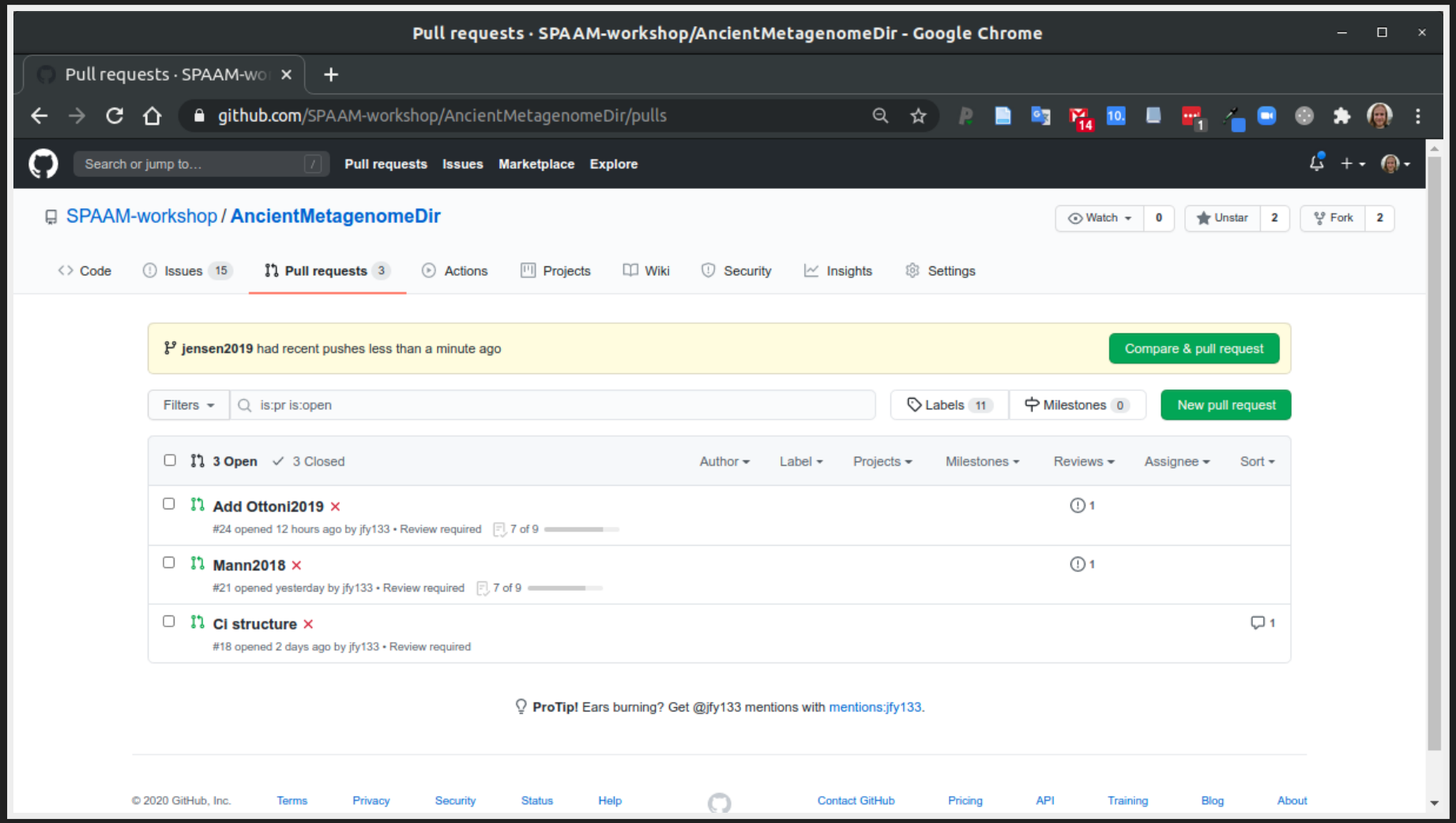
53	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS31	Homo sapiens	200	10.1186/s40
54	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS32	Homo sapiens	200	10.1186/s40
55	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS33	Homo sapiens	200	10.1186/s40
56	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS34	Homo sapiens	200	10.1186/s40
57	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS35	Homo sapiens	200	10.1186/s40
58	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS36	Homo sapiens	200	10.1186/s40
59	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS37	Homo sapiens	200	10.1186/s40
60	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS38	Homo sapiens	200	10.1186/s40
61	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS39	Homo sapiens	200	10.1186/s40
62	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS40	Homo sapiens	200	10.1186/s40
63	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS42	Homo sapiens	200	10.1186/s40
64	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS43	Homo sapiens	200	10.1186/s40
65	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS44	Homo sapiens	200	10.1186/s40
66	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS45	Homo sapiens	200	10.1186/s40
67	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS46	Homo sapiens	200	10.1186/s40
68	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS47	Homo sapiens	200	10.1186/s40
69	Velsko2018	2018	10.1186/s40168-019-0717-3	Radcliffe Infirmary	51.752022	-1.257677	United Kingdom	CS48	Homo sapiens	200	10.1186/s40
70	Jensen2019	2019	10.1038/s41467-019-13549-9	Sylthol m	54.66	11.35	Denmark	Sylthol m_1	Homo sapiens	5700	10.1038/s41

Pull Request: Pull Request Tab

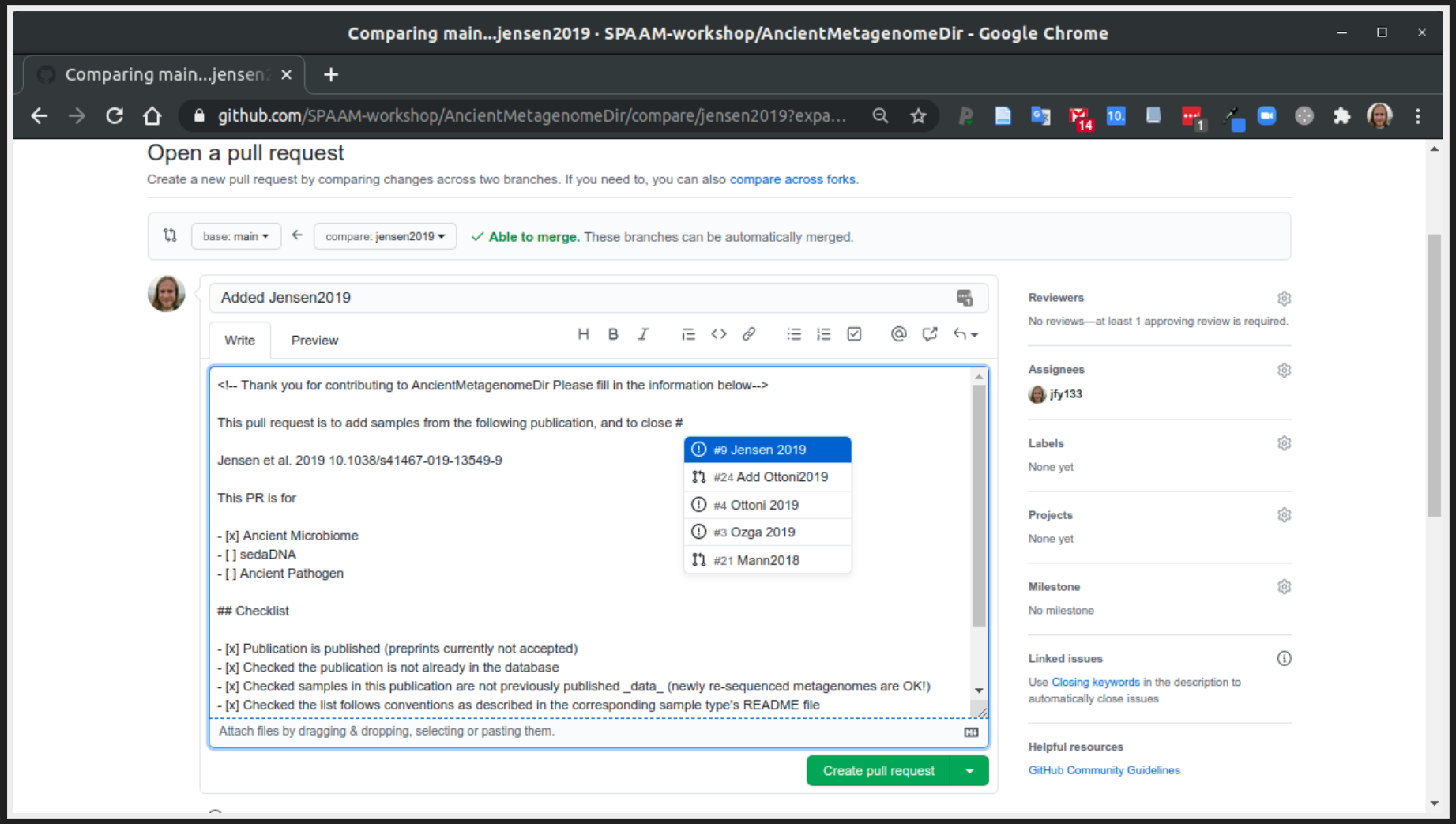


	project_name	publication_year	publication_doi	site_name	latitude	longitude	geo_loc_name	sample_name	sample_host	sample_age	sample_age
2	Warinner2014	2014	10.1038/ng.2906	Dalheim	51.565	8.84	Germany	B61	Homo sapiens	900	10.1038/ng.2
3	Warinner2014	2014	10.1038/ng.2906	Dalheim	51.565	8.84	Germany	G12	Homo sapiens	900	10.1038/ng.2
4	Weyrich2017	2017	10.1038/nature21674	Gola Forest	7.657	-10.841	Sierra Leone	Chimp	Pan troglodytes	100	10.1038/natu
5	Weyrich2017	2017	10.1038/nature21674	El Sidrón Cave	43.386	-5.328	Spain	ElSidron1	Homo sapiens neanderthalensis	49000	10.1038/natu
6	Weyrich2017	2017	10.1038/nature21674	El Sidrón Cave	43.386	-5.329	Spain	ElSidron2	Homo sapiens neanderthalensis	49000	10.1038/natu
7	Weyrich2017	2017	10.1038/nature21674	Spy Cave	50.48	4.674	Belgium	Spy1	Homo sapiens neanderthalensis	35800	10.1038/natu
8	Weyrich2017	2017	10.1038/nature21674	Spy Cave	50.48	4.674	Belgium	Spy2	Homo sapiens neanderthalensis	36000	10.1038/natu
					53.96	21.86	Poland	EuroHG1	Homo sapiens	7600	10.1038/natu

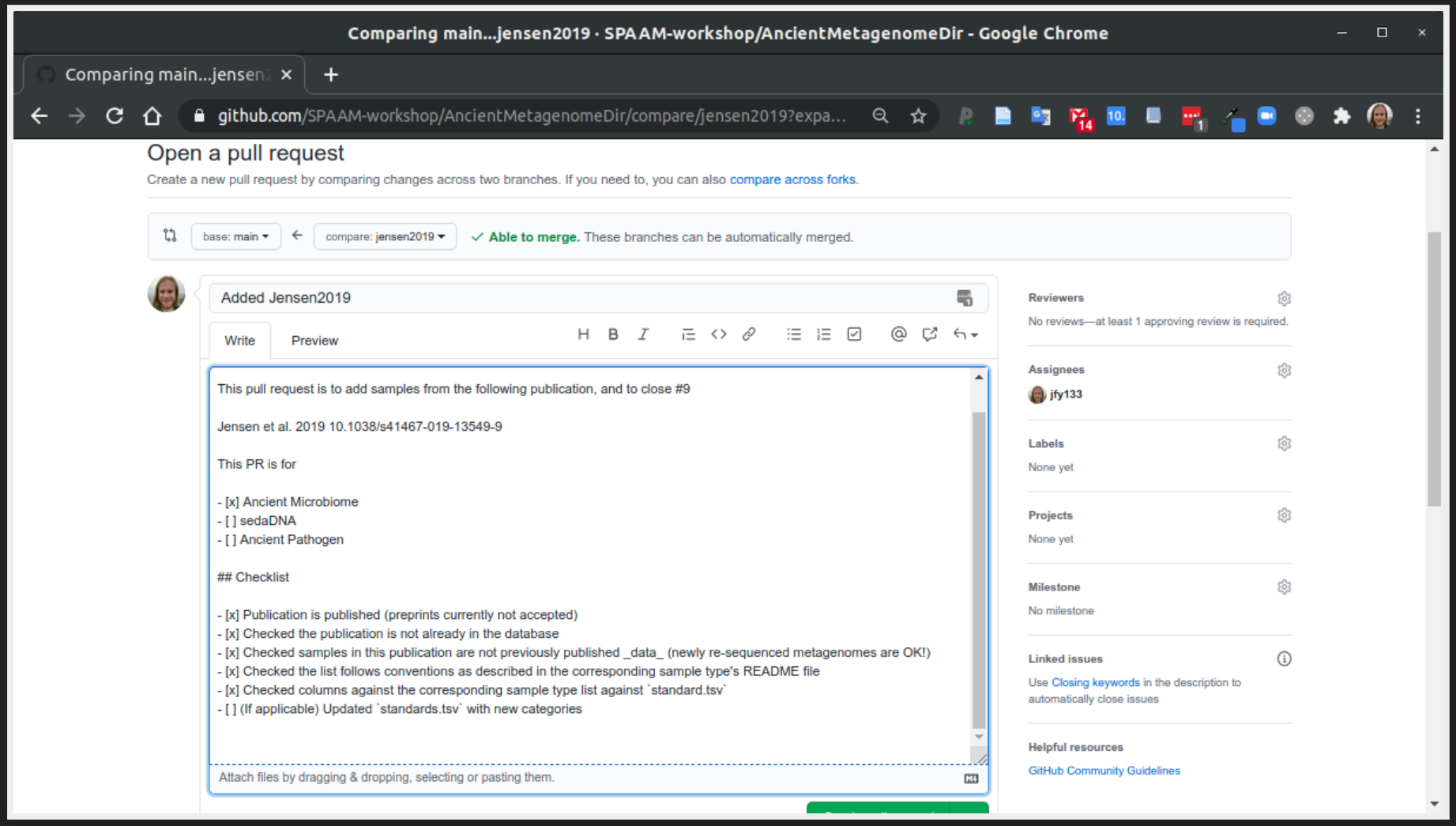
Pull Request: Make PR



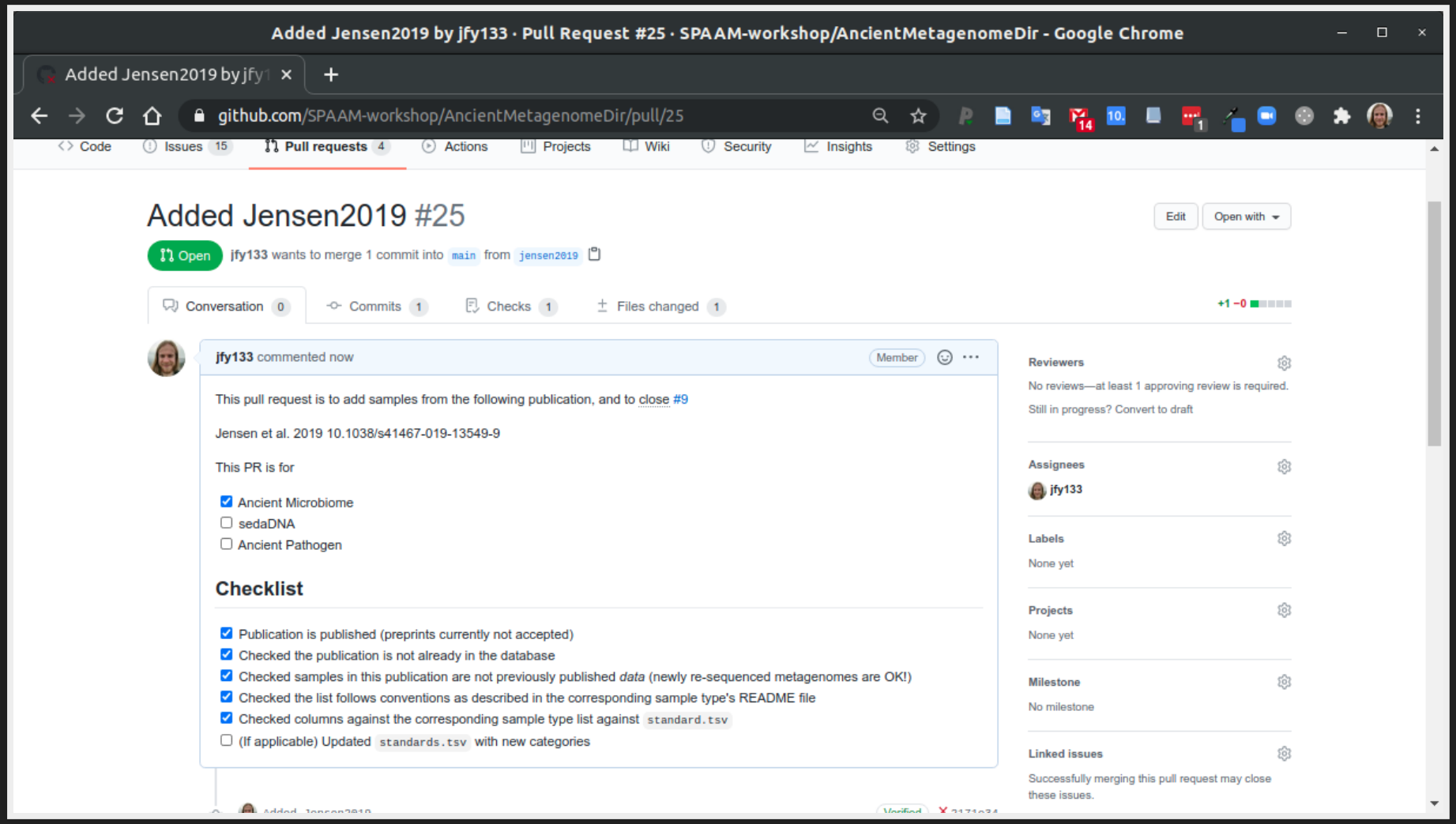
Pull Request: PR Information



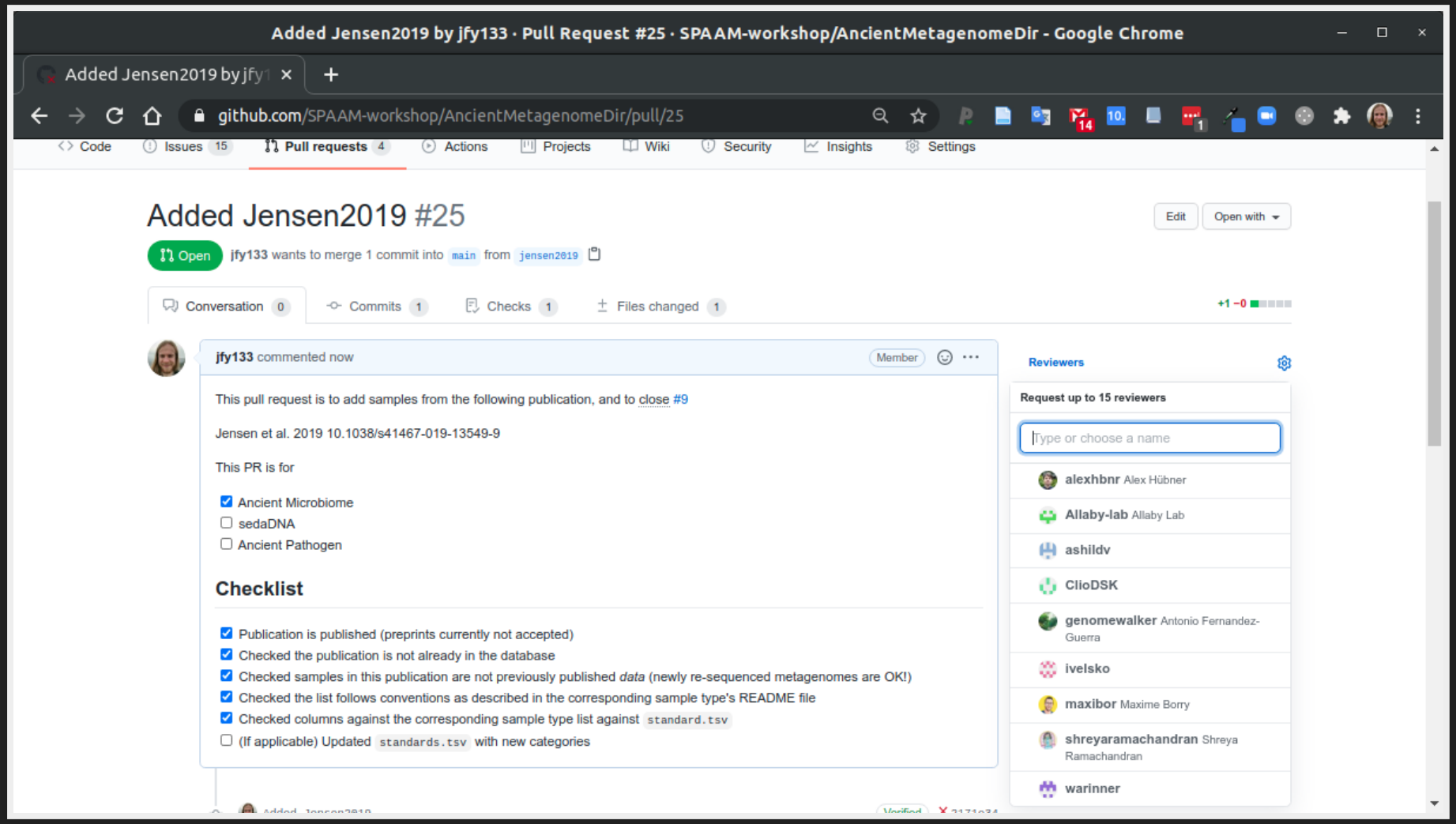
Pull Request: PR Checklist



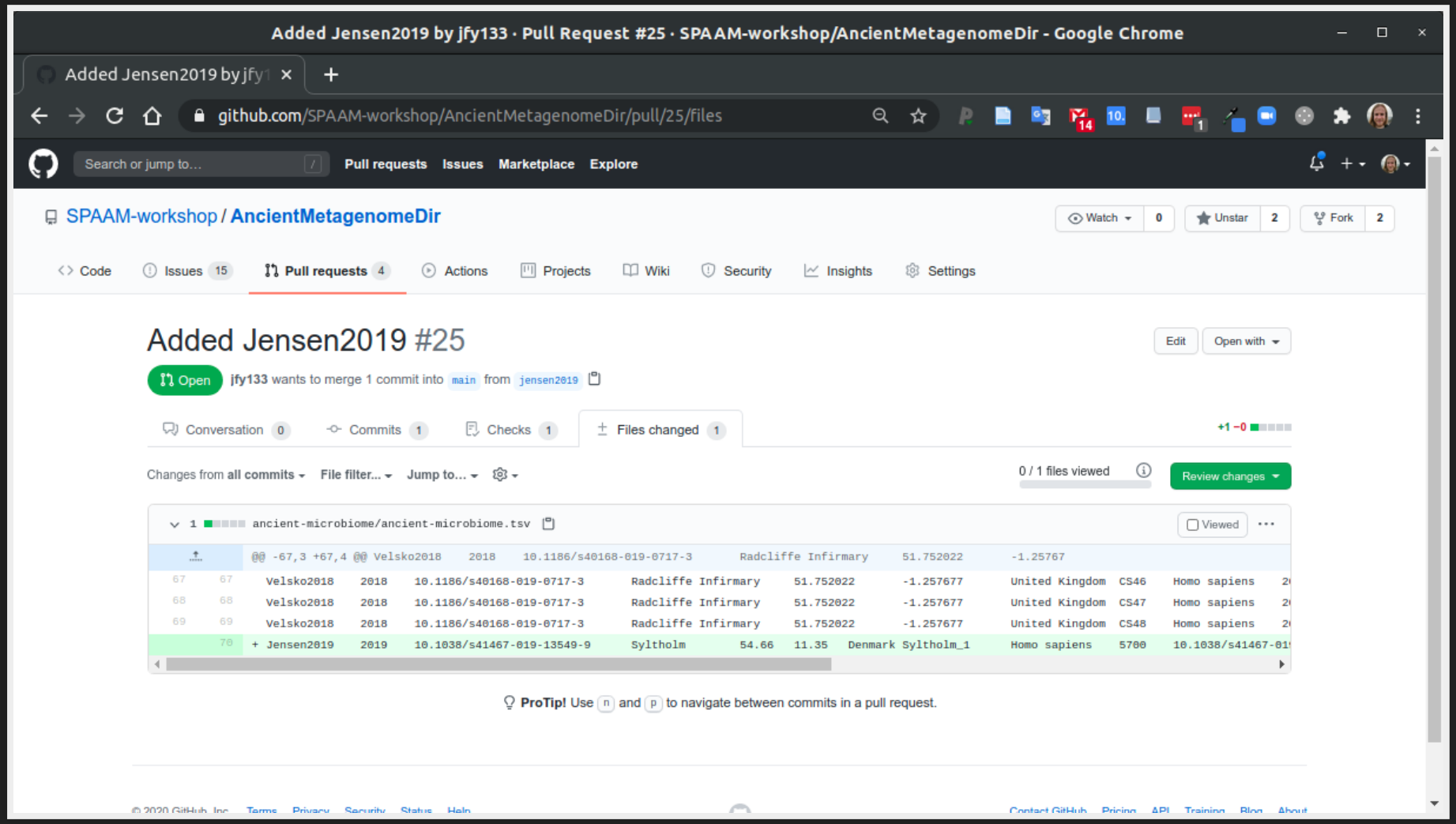
Pull Request: Submit PR



Pull Request: Request Review



Pull Request: View PR



You’re done!

Just rinse and repeat!

Everything else: Just Ask (Google first 😉)

- Atlassian git tutorials
- GitHub tutorials

Questions?

Notes on AncientMetagenomeDir

- Still ‘under dev’ - expect changes
  - Keep eye on slack channel and github issues
- Help requested - definitions of fields for:
  - Pathogens
  - sedaDNA
- Future
  - Add automatic checking of PRs
  - Creating review ‘team’
  - Expansion to data?

Remaining time

- Assign your self an issue
- Make a branch and start!

Thank you!

You can find me on

- GitHub (jfy133)
- Slack
- or email me