VEuPathDB User Impact and Sustainability Survey- Report[1]

Contents

[1] Version 1.0, Feb 28, 2025

Executive Summary

*The **VEuPathDB User Impact and Sustainability Survey** was conducted to assess the importance of VEuPathDB resources for biomedical researchers, and explore potential funding strategies following the loss of NIAID support, effective on 15 September 2024. This survey gathered insights into end-users' reliance on VEuPathDB databases, key features accessed, and the feasibility of alternative funding models.*

Methodology

The survey was conducted from **14 September 2024 – 7 January 2025**, using **Qualtrics**. It was disseminated *via* email, social media, and in presentations and discussions at scientific conferences, yielding **1,862 responses**, with **88-98% completion rates** for quantitative data.

Key Findings

- **VEuPathDB is a core resource for mycology, parasitology and vector biology research: 63% of respondents** reported that the loss of VEuPathDB would reduce their productivity by **75-100%**, citing its important role in data centralization, integration, improved research efficiency, and recruiting new investigators to the field.
- **Frequent usage: 80% of respondents** use VEuPathDB resources **daily or weekly**.
- **Global reach:** Users from **81 countries** participated, with the highest representation from the **USA (33%)**, followed by the **UK, China, India, Germany, Brazil,** and **France**.
- **User base: Students/postdocs accounted for 50%** of respondents; **35% of respondents were Principal Investigators/lab heads**.

Funding Model Considerations

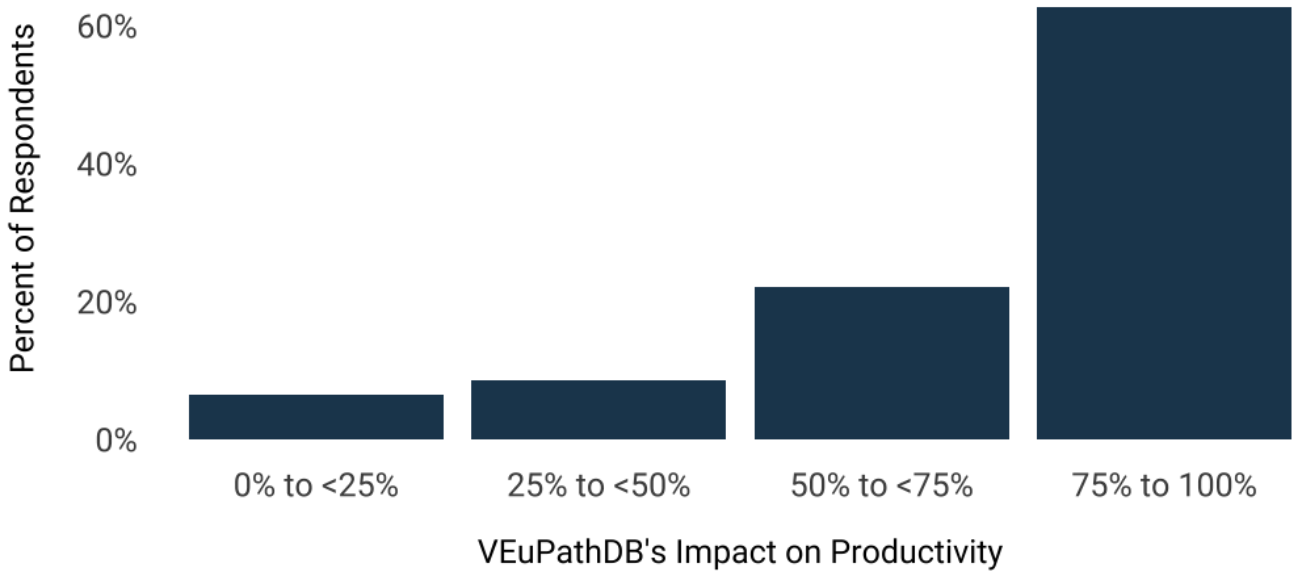
- **Feasibility of a fee-for-service model:** 75% of respondents support or are open to considering a paid model.
- **Concerns about equitable access:** Respondents stressed the need for **regional and institutional pricing** to ensure equitable access by underfunded investigators, including new investigators, small labs, LMIC researchers, educational applications, *etc*.
- **Open Science**: While acknowledging financial constraints and trade-offs, there is a strong community preference for **FAIR, open-access solutions.**
- **Need for Clear Communication:** If fees are introduced, researchers will require **sufficient lead time** to integrate these costs into grant applications.
- **Funding Challenges:** Users cite **funding constraints, administrative hurdles, and philosophical opposition** to monetizing scientific resources as barriers to a paid model.

Conclusion

This survey underscores **the irreplaceable, important role played by VEuPathDB** in biomedical research on fungi, protozoan parasites, and arthropod disease vectors, including the globally important (but often neglected) human/animal and agricultural diseases they cause. The survey also highlights widespread concerns over the possible loss of VEuPathDB, and bewilderment over how the global funding community might allow this to occur. While users are **open to alternative funding models,** the community strongly favors **broad accessibility**. A sustainable future for VEuPathDB may require **a combination of equitable fee structures, diversified funding sources, and strategic cost-saving measures** to preserve its mission.

Q1. What would be the impact on your lab's productivity if VEuPathDB resources were to disappear altogether? (N=1811)

On a scale from 0-100, please estimate the impact on your lab's productivity if VEuPathDB resources were to disappear altogether (numerical value will appear at right of slider). How to define impact is up to you: increased time required to do your work, lab budget, number of publications, etc. - 0% = no impact, or I don't use these databases at all; 100% = would shut down my work completely



The data indicates that a significant majority of respondents (63%) reported that their productivity would be 75-100% affected by the loss of VEuPathDB resources.

Q2. *Free text:* Please describe how the loss of VEuPathDB databases would impact your work (N=1456)

Representative free text responses describing key themes and impacts from the loss of VEuPathDB databases are shown in italics. A full list of all responses is available upon request (88 pages).

1. **Disruption of Scientific Research:**
    - VEuPathDB is integral to many researchers' workflows, serving as a central hub for genomic, transcriptomic, and proteomic data. Its loss would significantly hamper hypothesis generation, experimental design, and data analysis.
    - Many respondents mention having to revert to manual, fragmented, or less accurate methods to gather data, leading to inefficiencies and delays.

        "We use the database daily, frequently searching for sequences and becoming highly familiar with its interface. Transitioning to a new database and potentially replacing the current one will significantly impact our daily workflow."

2. **Resource Centralization (One-stop Shop):**
    - The database aggregates diverse, curated datasets in one user-friendly platform, making it easier to access information & conduct comparative analyses. Without it, scientists fear loss of integration, increased error rates & slower learning curves for new students and researchers.

        "It is an irreplaceable resource for our work. It bundles many resources that would otherwise be difficult to access in a very convenient and user-friendly way. This saves an enormous amount of time when doing database work."

3. **Loss of Specialized Data:**
    - Researchers studying niche organisms like *Aspergillus, Candida,* and other fungi; *Toxoplasma, Plasmodium,* and other parasites; and arthropod disease vectors emphasize that the specificity, breadth, and depth of VEuPathDB's data cannot be easily replicated elsewhere, such as NCBI, Ensembl, UniProt, or Mycocosm resources.

        "There is no other resource I can use with comparable quality and tools to analyse RNASeq data in P. falciparum."

4. **Barriers to Education and Training:**
    - VEuPathDB serves as a valuable teaching tool, offering straightforward access to complex datasets for students and early-career researchers. Its absence would impede training, learning, and broad access to advanced bioinformatics tools, particularly in low and middle-income countries.

        "I train a wide range of students and junior scientists, many with very mixed computational backgrounds and skills. VEUPathDB is a very easy entry point for all trainees to start to explore a very wide range of data types and sets."

5. **Data-Sharing, Collaboration and Community Impact:**
    - The platform's standardization promotes collaboration and reproducibility in research. Without it, sharing data and insights across labs will become much more challenging.

"The databases allow our lab to benefit from investments made by other labs and vice versa, increasing output from every investment made into research on eukaryotic pathogens."

6. **Strain on Economic and Informatics Resources:**
    - Many researchers would need to invest significant time and/or resources into rebuilding local versions of the database or adopting alternative workflows, increasing costs and lowering productivity.

        "The data compilation and curation provided a first step and often repeated step in hypothesis generation and experimentation. Without this resource, members of our lab would spend 2-5 times more time investigating feasibility and gathering rationale for ideas."

7. **Impact on Future Research:**
    - The absence of VEuPathDB may lead to gaps in research continuity, hinder cross-species analyses, limit the scope of novel discoveries in fields like infectious diseases, genomics, and molecular biology, and discourage new investigators from working in these fields.

        "Halting the maintenance of such resources would not only jeopardize ongoing research but also impede future scientific advances that depend on access to new experimental data."

Overall analysis of key themes and impacts from the loss of VEuPathDB databases:

1,456 responses from this survey indicate that VEuPathDB's role in centralizing and simplifying data access, curation, and analysis is unparalleled for many researchers, particularly in parasitology, fungal genomics, vector biology, and molecular genetics. Its absence would create an academic and logistical vacuum, significantly increasing the time and effort required for individual and collaborative efforts, and discouraging research in these important scientific areas. Ensuring the continuity of functionalities currently provided by VEuPathDB is a high priority for maintaining productivity in these fields.

"Losing VEuPathDB would set research back by 20-30 years."

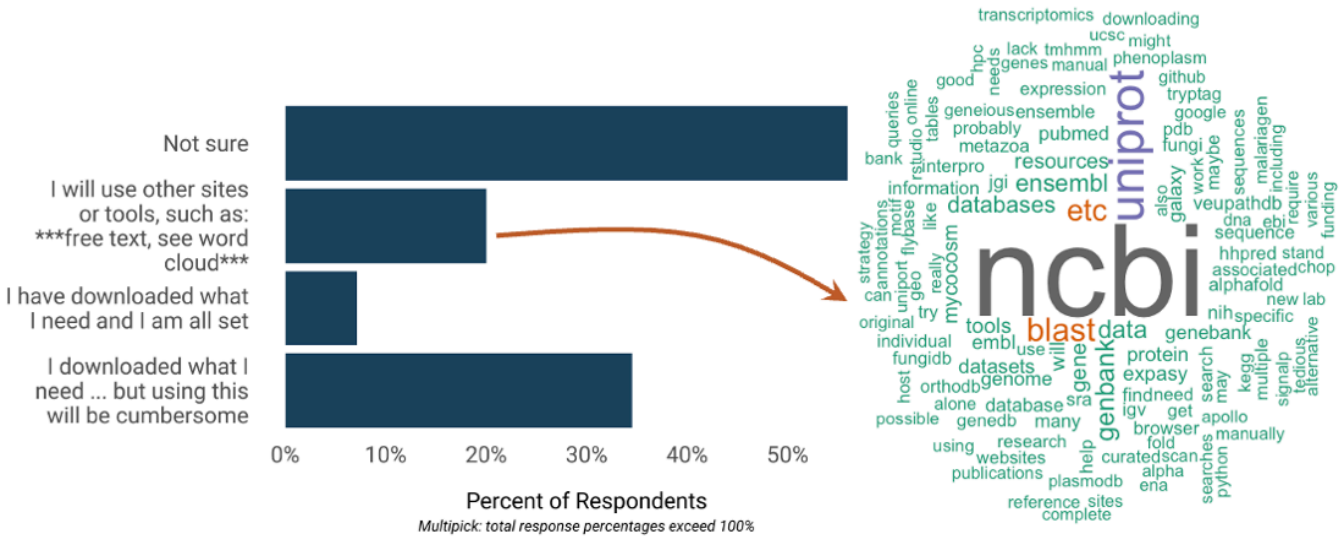
"It's not that it's a great resource for Eukaryotic pathogens, it's a great resource. Other fields would do well to set up their own."

"Virtually on an hourly basis some question comes up that can be easily answered by working with the VEuPathDB interfaces. This includes my personal curiosity as well as teaching students, discussing projects, working with collaborators, and the list goes on."

"In general I like that VEuPathDB is very usable by biologists who aren't bioinformatics experts. Nobody goes into infectious disease research because they love computers"

"The smooth progress of my academic project is really inseparable from this website!!!"

Q3a. How else would you accomplish the tasks for which you use VEuPathDB? (N=1821)



The data indicates that a majority of respondents (85%) indicated that they were either not sure of how they would accomplish the tasks for which they currently use VEuPathDB or that accomplishing tasks would be cumbersome without VEuPathDB.

Q3b. *Free text:* I will use other sites or tools, such as: (N=332)

332 respondents indicated specific resources that they will use (word cloud and summarized below).

A full list of all free text responses is available upon request (8 pages).

Key Takeaways

- **NCBI Dominance**: NCBI was by far the most dominant tool mentioned, referenced 239 times in various contexts (e.g., NCBI BLAST, GenBank, GEO, SRA). This dominance suggests it is either the easiest fallback option or the most universally recognized. However, usability issues were repeatedly mentioned, with respondents finding NCBI cumbersome or incomplete.

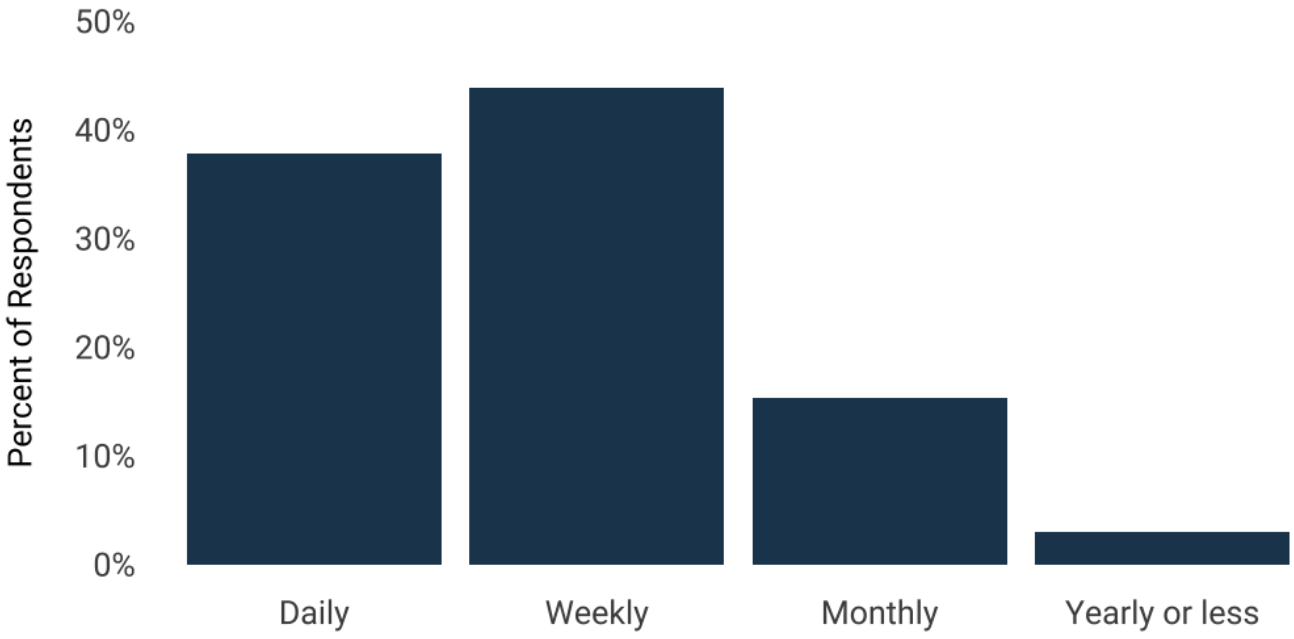
- **Fragmentation Issues**: Several respondents highlighted the difficulty of having to use multiple tools/databases, which is time-consuming and inefficient. Many respondents underscored the need for platforms like VEuPathDB, which provide curated, integrated data tailored to specific research fields. Its absence forces users to rely on fragmented and less user-friendly tools.

    - *"It would be very tedious and would require using many, many other resources."*

- **Lack of alternatives**: Responses like *"Pentagrams and animal sacrifice"* may be facetious but underscore frustrations in finding equally functional alternatives.

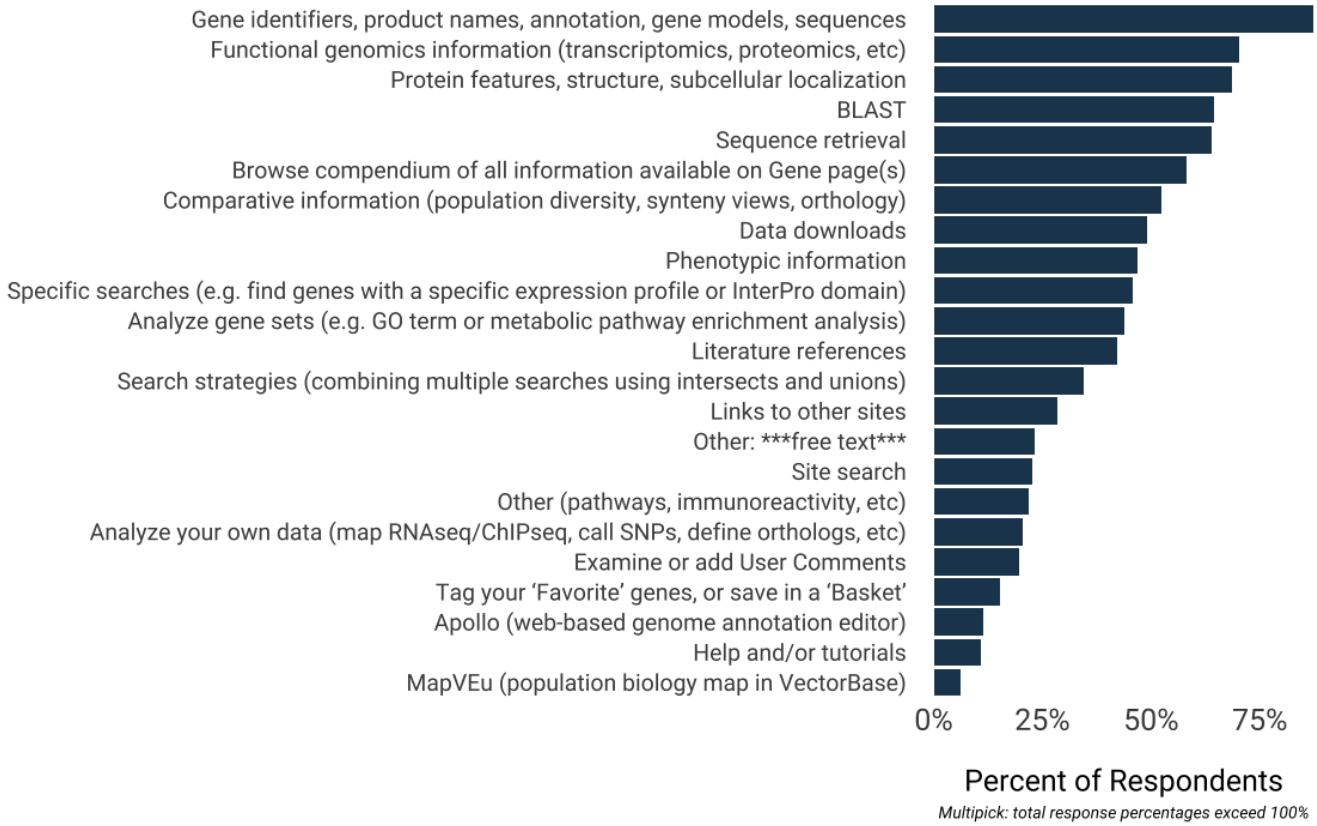
Overall Analysis: Researchers are frustrated with the lack of centralized, curated resources, which increases their workload and introduces inefficiencies. The absence of VEuPathDB creates a gap that existing resources cannot easily fill, leading to fragmented workflows.

Q4. How often do you typically access VEuPathDB sites? (N=1830)



The data indicates that a large majority of respondents (82%), reported that they use VEuPathDB on a frequent basis, daily or weekly.

Q5a. What features have you used most often (check all that apply)? (N=1830)



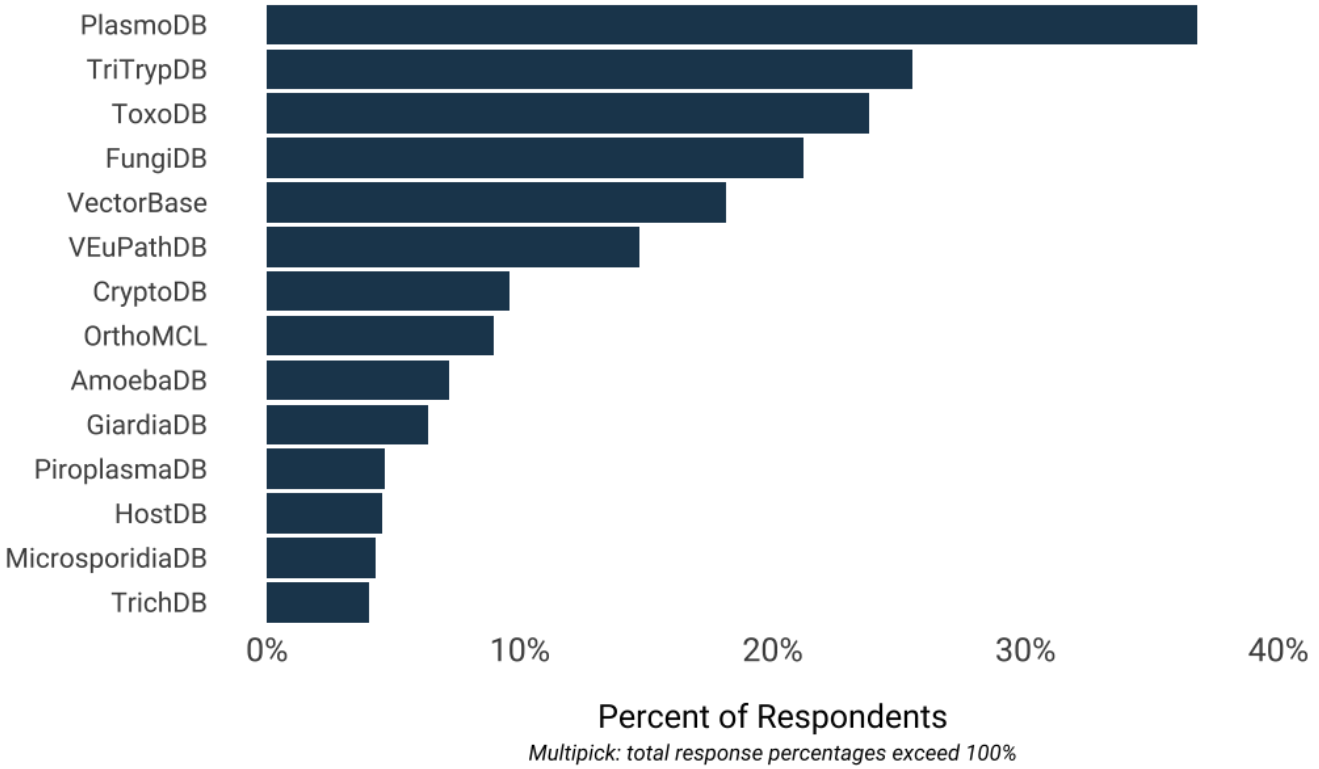
VEuPathDB features used most frequently include gene features (gene identifiers, product names, annotation, gene models, sequence), functional genomics, protein features, BLAST, and sequence retrieval.

Q5b. *Free text:* Other features used (N=27)

A full list of all free text responses is available upon request (1 page).

27 Survey respondents provided additional information on other tools they use most often. 26% of these free text responses indicated that they use CRISPR/CAS9 tools (including gRNA design) frequently.

Q6. Which VEuPathDB websites do you regularly use? (N=1827)

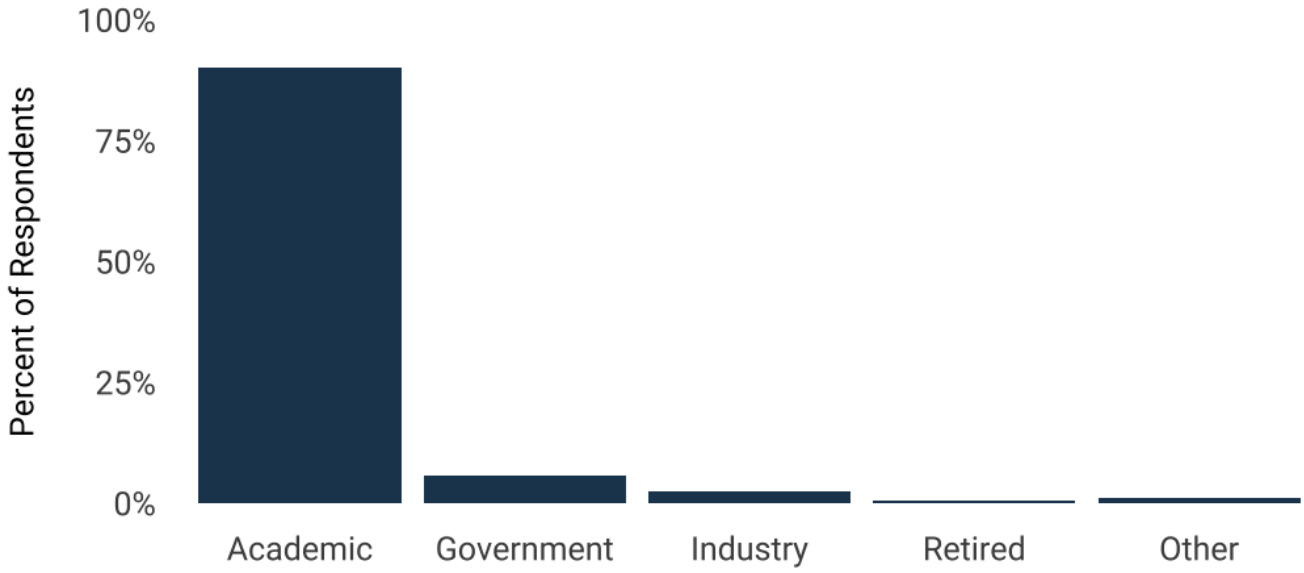


The proportion of respondents who are regular users of various VEuPathDB platforms generally reflects usage of our sites. The higher proportion of respondents who are parasitology site users could be attributed to the fact that the survey was launched during major in–person parasitology meetings.

Q7. In what country do you work? (N=1790)

Respondents represented 81 countries, highlighting the global reach of VEuPathDB. The highest proportion of respondents were from the USA (33%), followed by the UK, China, India, Germany, Brazil, and France.

Q8. What type of institution do you work in? (N=1813)



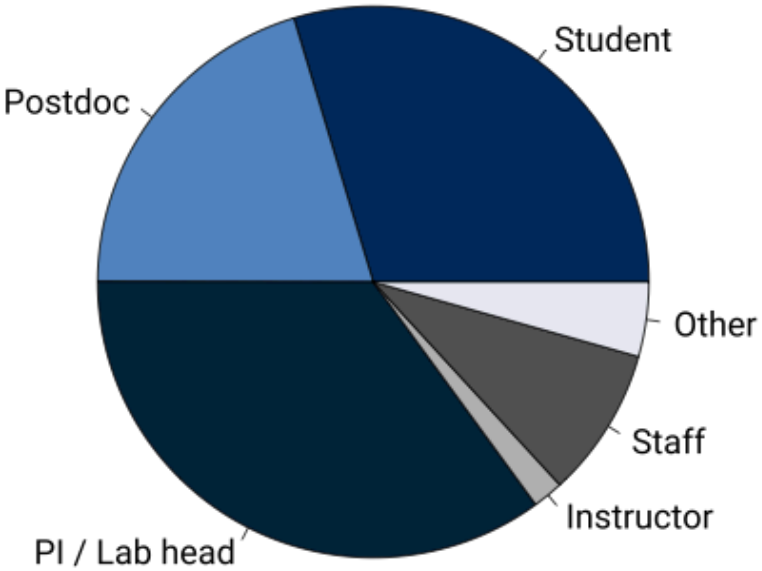
90% of respondents were from academic institutions, 6% from government institutions, and 3% from industry.

Free text: Other institution type (N=19)

A full list of all free text responses is available upon request (1 page).

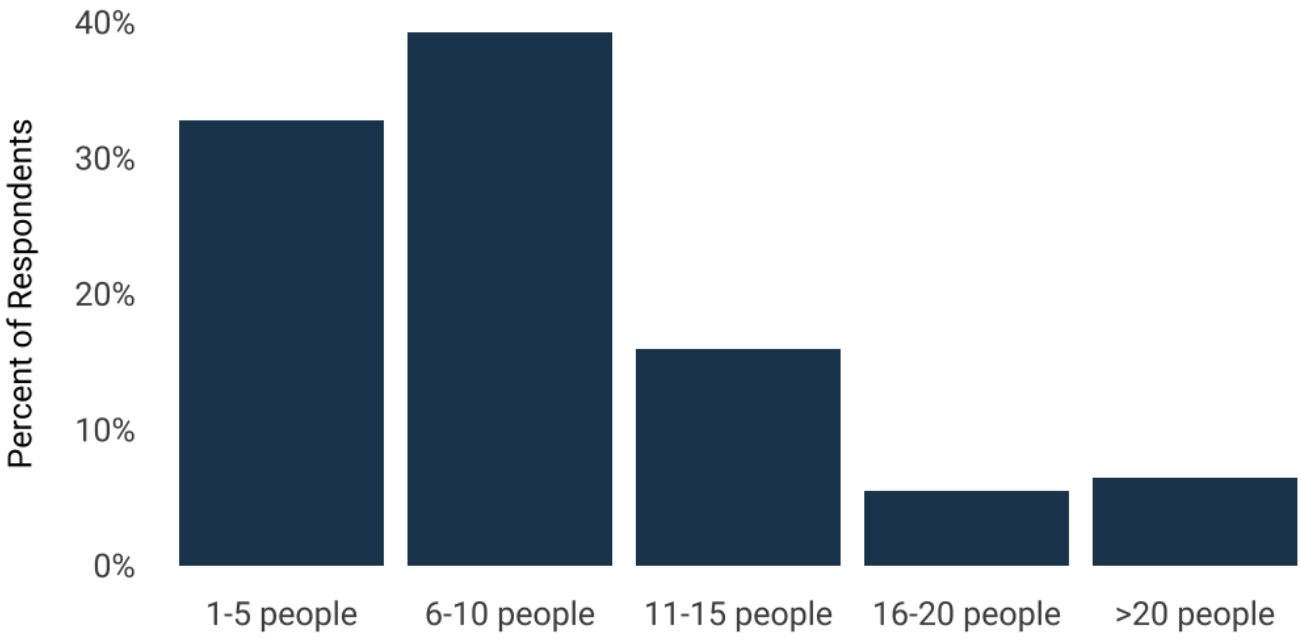
32% of free text responses mentioned research institutions, 16% mentioned hospitals and healthcare settings, and 16% mentioned foundations and NGOs.

Q9. What is your primary position? (N=1831)



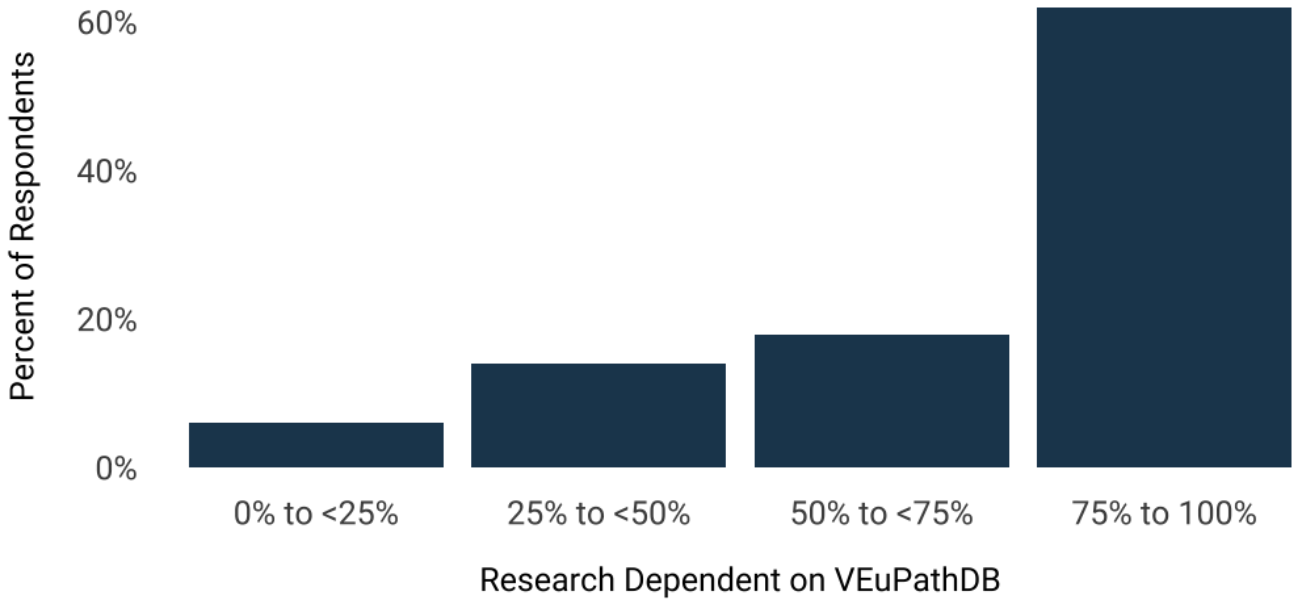
Trainees (postdocs and students) represent 50% of the respondents while Principal investigators/ lab heads make up 35%.

Q10. How many people work in your group? (N=1799)



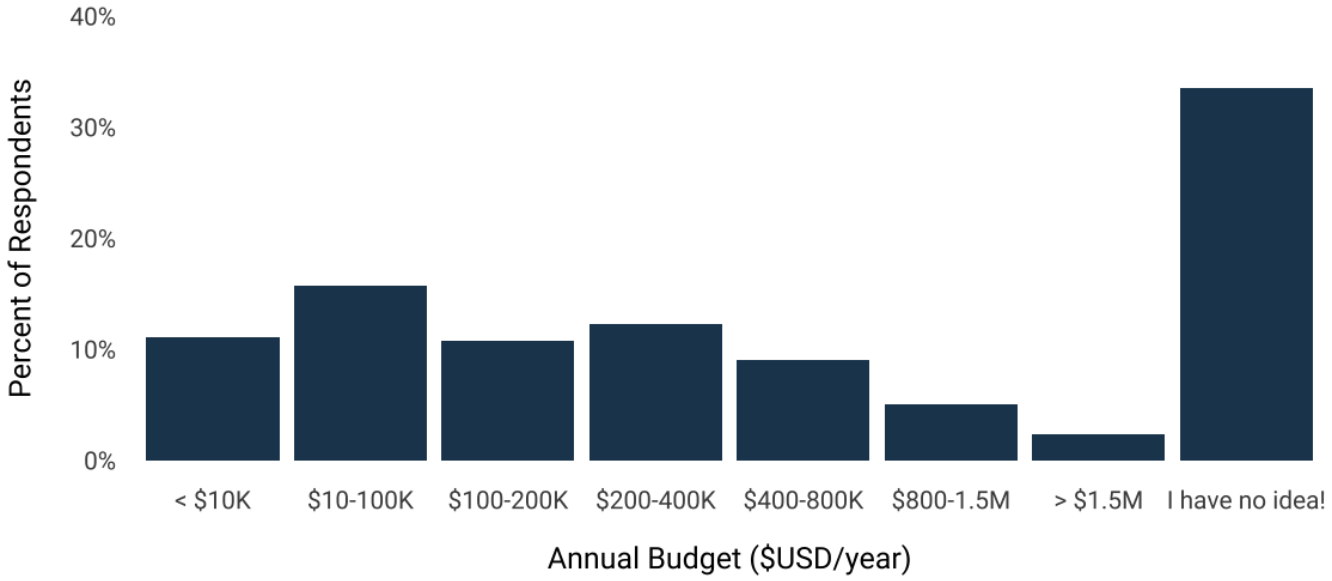
72% of respondents work in research groups of 1-10 people.

Q11. What percentage of the research in your laboratory depends to some extent on VEuPathDB databases? (N=1635)

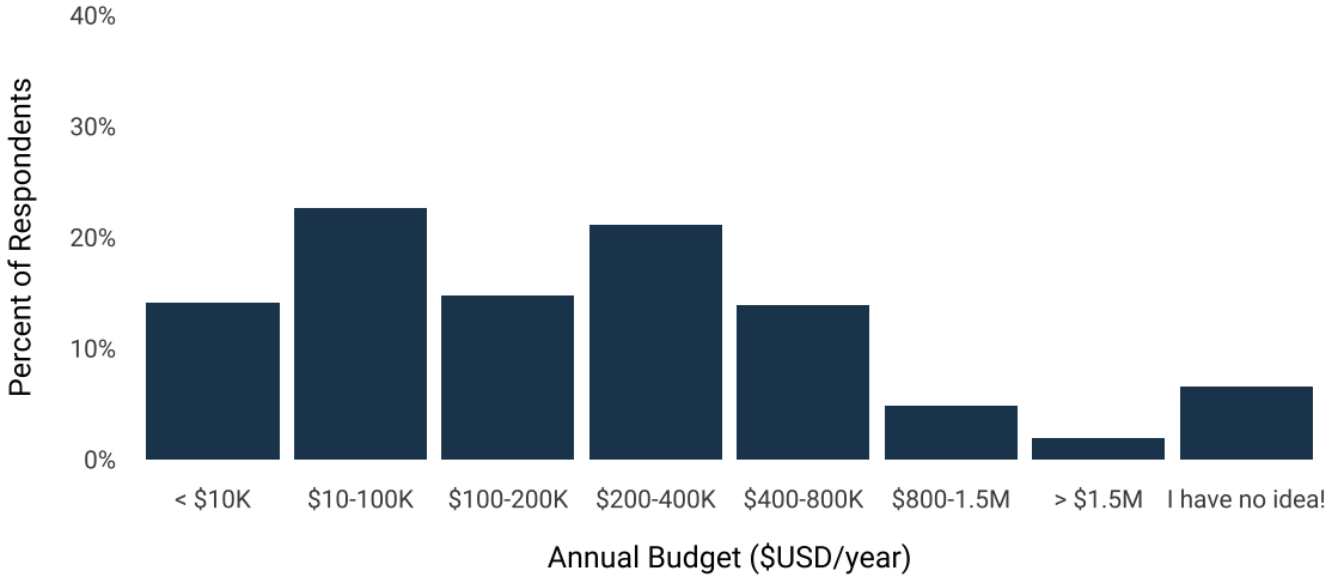


A large majority of respondents (62%) reported that more than 75% of the research in their lab depends to some extent on VEuPathDB databases.

12. What is the approximate total annual budget of your lab? Please include the value of all activities: grants, salaries, studentships, facilities, services, etc. (N=1788)



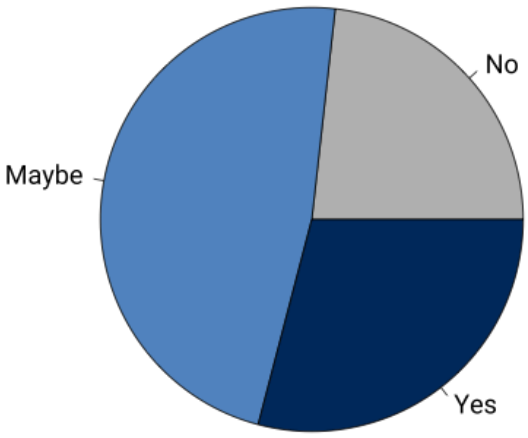
What is the approximate total annual budget of your lab? ***Data restricted to PIs/lab heads.*** **(N=635)**



Lab budgets of most respondents are <USD 800, 000 per year.

Q13a. Would you consider it appropriate to institute a fee-for-service model, with charges for access to VEuPathDB paid from users' grants or by their institutions? (N=1795)

Note: Charges would be based on lab budgets in order to ensure accessibility for those without funding, researchers in low-income countries, teaching institutions, etc.



A clear majority of respondents (79%) indicated full support (*i.e.,* answered "Yes") or partial support (*i.e.,* answered "Maybe") for the fee-for-service model.

Q13b. *Free text:* Comments on fee model (N=485)

Representative free text responses describing perspectives on the fee-for-service model are shown in italics. A full list of all responses is available upon request (22 pages).

There is a wide spectrum of perspectives on a fee-for-service model for VEuPathDB, categorized primarily into:

Supportive with Conditions

1. **Affordability and Fairness**:
    - Many respondents support the fee-for-service model if it is reasonably priced and scaled to reflect lab budgets or regional income disparities (e.g., reduced fees for Low- and Middle-Income Countries).
    - Suggestions included subscription models, institutional licensing, or group fees rather than per-individual charges.

        "My main worry would be how this would be implemented and to ensure researchers from LMICs are able to use VEuPathDB"

2. **Incorporating Costs into Grants**:
    - A significant number of participants proposed including database fees as part of grant budgets, though they acknowledge challenges due to funding cycles and grant constraints.

        "It would be hard to implement at first, because we do not have specific budget items on our grants for this service. In the future, we could enter those items into grant budgets."

3. **Premium Models**:

- Several responses suggest a hybrid model offering free basic access with additional paid premium features or services (e.g., advanced tools or personalized support).

Concerns and Opposition

1. **Impact on Equity**:
   - A recurring concern is that fees will disproportionately affect researchers from underfunded labs, low-income countries, or early-career researchers, deepening existing inequities in research opportunities.
2. **Challenges in Implementation**:
   - Concerns were raised about the administrative burdens of enforcing a fee-based model, compatibility with institutional payment systems, and issues like budget unpredictability.
3. **Moral and Ethical Reservations**:
   - Several respondents emphasized that public research and taxpayer-funded projects should remain freely accessible. They argue that privatizing access to publicly funded resources is counterproductive to the spirit of open science.

     "Privatization of resources that are used for public research will end science."
4. **Research Implications**:
   - Some noted that introducing fees could reduce user engagement, hinder casual or exploratory use, and discourage new users or researchers from less-resourced settings.

Alternatives Suggested

1. Government and Institutional Funding:
   - A large number of responses advocate for stable funding from agencies like NIH or support from international organizations (e.g., WHO, Gates Foundation) to maintain free access.
2. Voluntary Contributions:
   - Voluntary donation models, akin to Wikipedia or Flybase, were suggested as less restrictive funding mechanisms.
3. Institutional Partnerships:
   - Leveraging institutional-level funding or partnerships with academic institutions and industry was proposed to reduce individual user costs.

Key Themes in Recommendations

1. **Accessibility for Low-Income Regions and Small Labs**: Any fee structure should account for regional income disparities and lab size to ensure equitable access.
2. **Sustainability vs. Openness**: While there's acknowledgment of the financial realities, there's a strong preference for balancing sustainability with open access principles.
3. **Importance of Communication**: Respondents stress the need for clear communication about costs and sufficient lead time to integrate these into grant applications.

Q14. Technical difficulties

If you were expected to pay for the services provided by VEuPathDB, are there specific technical difficulties that might have to be addressed? *A full list of all free text responses is available upon request (21 pages).*

Key Themes and Observations from 682 responses:

1. **Funding Challenges**:
   - Budget Constraints: Many respondents indicated that limited lab budgets, particularly in developing countries, make it difficult to pay for such services. Students, early-career researchers, and those between grants are especially vulnerable.
   - Grant Limitations: Researchers highlighted that many grants don't allow funding to be allocated for database subscriptions or require explicit budgeting during the grant proposal phase, which might not have been considered.
   - Institutional Support: Some users noted challenges in convincing institutional management to support these costs, especially given other academic financial demands (e.g., publication fees).

2. **Administrative and Bureaucratic Hurdles**:
   - Complex Payment Processes: Researchers pointed out administrative difficulties, such as delays in payment approvals, complicated procurement systems, and institutional requirements for vendor registration.
   - Regional Constraints: Issues with international payments, restrictions in countries like Brazil, Argentina, and India, and limited options like lack of credit cards or payment flexibility were common concerns.
   - Need for Invoices and Purchase Orders: Many respondents emphasized the necessity for clear billing structures, such as invoices and quotes, to facilitate institutional payment processes.

3. **Access and Licensing Concerns**:
   - Multi-User and Institutional Access: Respondents suggested creating institutional or lab-based licenses rather than individual ones. They emphasized the need for flexible access for multiple users, including temporary access for project students or visitors.
   - Scalability: Some worried about subscription models creating inequitable access, with smaller labs or institutions being disproportionately affected.

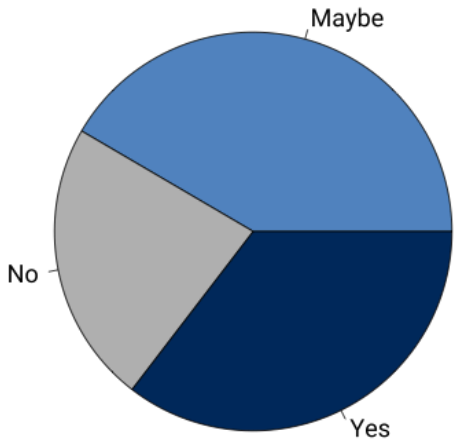
4. **Technical Expectations**:
   - Improved Support: If the service becomes paid, users expect enhanced technical support, faster access, user-friendly interfaces, and high-quality data integration.

5. **Ethical and Philosophical Concerns**:
   - Impact on Open Science: Respondents raised concerns about restricting access to essential research tools, especially in fields that benefit global health, which could hinder collaboration and scientific progress.
   - Equity Issues: Many feared that placing these resources behind a paywall would disproportionately disadvantage researchers in low-income countries or smaller institutions.

6. **Alternative Solutions Suggested**:
    - Funding by Agencies: Some proposed that national or international funding agencies directly support VEuPathDB to keep it free or subsidized for users.
    - Flexible Pricing Models: Suggestions included usage-based fees, discounted rates for low-income countries, and scalable fees based on lab or institutional budgets.
    - Trial Periods: A trial model was recommended to allow users to assess the value before committing to a subscription.

General Sentiment: The overall sentiment among respondents is that while some might be able to adapt to a paid model, a significant proportion of users would face barriers due to funding constraints, administrative challenges, and philosophical opposition to monetizing scientific resources.

Q15a. Would you consider it appropriate to institute a tiered pricing structure, with basic services available for free, but where the most recent data, or high-end services, or customer support is available only to those able to contribute financially? (N=1736)



The majority of respondents (68%) are willing to consider a tiered pricing structure.

Q15b. *Free text:* Comments on tiered pricing structure (N=338)

Representative free text responses describing perspectives on a tiered pricing structure are shown in italics. A full list of all free text responses is available upon request (14 pages).

There is a wide spectrum of perspectives on a tiered pricing structure for VEuPathDB, categorized primarily into:

1. **Equity Concerns (Major Theme)**:
    - Many respondents expressed concerns about the inequities a tiered pricing model would introduce, particularly for researchers in low- and middle-income countries (LMICs) or smaller labs with limited funding.
    - Restricting access to the latest data or critical services would widen the gap between well-funded and underfunded research institutions.
    - A paywall would disadvantage groups most affected by neglected diseases, for whom VEuPathDB is a critical resource.
    - Equity vs. Sustainability: The responses reveal a tension between ensuring equitable access to scientific resources and the need to secure funding for database maintenance. While some see tiered pricing as a pragmatic solution, others fear it will deepen disparities in research opportunities and outcomes.
        > *"I just read a paper from a low income lab that only had the resources to do bioinformatics and relied on VEuPathDB for their work. I'm not sure that paper would have been possible if they had to pay for premium services."*
2. **Support for Tiered Pricing**:
    - Some respondents agreed with the idea of charging for "high-end services," advanced customer support, or computationally expensive tasks, considering it a reasonable way to sustain the database.
    - Several users suggested that basic services, including access to the latest data, should remain free, with fees applied only for advanced tools or premium features.
3. **Opposition to Paywalls for Data**:

- A strong recurring opinion was that scientific data, particularly publicly funded or NIH-funded data, should remain freely accessible.
- Respondents worried that paywalling recent data would contradict the ethos of open science and potentially violate funding requirements.

4. **Alternative Funding Suggestions**:

   - Suggestions included:
     - Charging industry users or well-funded labs more.
     - Implementing a donation model or optional subscription.
     - Exploring government or institutional funding to maintain free access.
     - Offering discounts or exemptions for underfunded labs or researchers from LMICs.

5. **Ethical and Philosophical Concerns**:

   - Some respondents argued that paywalls for data contradict the principles of knowledge sharing, open access, and equitable research progress.
   - They stressed that science should not become a "pay-to-play" system.

6. **Practical Considerations**:

   - Respondents raised concerns about defining "basic" versus "high-end" services and the potential for confusion or mismanagement in implementing a tiered system.
   - A few highlighted fears of a "slippery slope," where free tiers might become less functional over time, as seen in other business models.

7. **Conditional Support**:

   - Several respondents were willing to accept a tiered pricing model if it were the only way to ensure the database's sustainability, provided that mechanisms were in place to ensure equitable access for low-budget users.

8. **Concerns About Impact on Research**:

   - Restricting access to data and tools might stifle innovation, delay research progress, and disproportionately affect researchers working on diseases prevalent in under-resourced regions.

Recommendations for Actionable Steps:

1. Preserve Free Data Access: Ensure that all researchers, regardless of financial capacity, can access core data and basic tools.
2. Charge for Premium Features: Introduce fees for high-end services, advanced analysis tools, or customer support, but offer waivers or discounts for LMICs and underfunded labs.
3. Transparent Communication: Clearly define "basic" and "premium" services, and communicate the rationale for any changes to the community.
4. Pursue Alternative Funding: Consider crowdfunding, government or institutional funding, or a donation-based model as supplementary revenue streams.
5. Equity Safeguards: Develop policies to ensure equitable access, such as tiered pricing based on institutional budgets or providing subsidies for disadvantaged groups.

Additional comments

Key Themes Identified in 254 additional comments:

Representative free text responses describing perspectives on a tiered pricing structure are shown in italics. A full list of all free text responses is available upon request (15 pages).

1. Funding Concerns and Suggestions

- Many respondents expressed frustration and disappointment over the loss of NIH funding, emphasizing the essential role of VEuPathDB in their research.

  "I do think there should be funding for such an essential tool."

  "It seems illogical that they are willing to pay extortionate open access fees to publishers for research outputs but not support the underlying bioinformatic resources."

- Multiple suggestions were proposed for alternative funding models, including:
  - Institutional or tiered user fees.
  - Crowdfunding, micro-donations, or sponsorships.
  - Seeking support from international agencies like WHO, Wellcome Trust, or the European Union.
  - Subscription models similar to journal access.
  - Exploring corporate sponsorship or partnerships (e.g., cloud hosting discounts, scientific advertising).
  - Leveraging social media (e.g., monetized YouTube tutorials).

2. Value of VEuPathDB to the Research Community

- Many researchers highlighted that they have relied on VEuPathDB for decades, both for their work and as a teaching tool.
- The database was described as critical for global research on parasitic and neglected diseases.
- Concerns were raised that losing access would negatively impact research, particularly in low- and middle-income countries (LMICs).

  "The cost of $5M/yr seems to me minimal compared to the cost that parasitic diseases represent in economic terms and in human lives worldwide."

  "The VEuPathDB system has been an outstanding democratizer of access to data and analysis tools for scientists worldwide. It is a necessary, powerful, and perfectly functional tool."

3. Opposition to Paywalls and Subscription Models

- Some respondents were open to minimal user fees, but many strongly opposed paywalls, fearing they would hinder research equity, especially for underfunded labs.
- A tiered pricing structure based on institutional resources was suggested to ensure continued access for researchers in developing countries.

4. **Frustration with NIH's Decision**

- Several researchers criticized NIH for defunding VEuPathDB while continuing to fund less widely used resources.
- Some called for advocacy efforts to pressure NIH or other funding bodies to reconsider.

> *"NIH wants data to be accessible, shared, and managed -- VEuPathDB does all of that!!"*
>
> *"5M/year seems like exceptional value for the VEuPathDB services. I am at a loss at the short-sightedness of this decision."*

5. **Potential Downsizing Strategies**

- Suggestions included reducing data update frequency, removing underutilized features, and simplifying the user interface.
- Some proposed focusing on core databases while allowing downloadable static datasets.

6. **Broader Implications for Research**

- Concerns were raised about long-term impacts on neglected disease research.
- Some respondents feared that without VEuPathDB, accessing and analyzing data would become significantly more challenging, slowing down scientific progress.

> *"The shutdown of the service has been a devastating blow to all researchers in this era of omics"*

7. **General Support and Encouragement**

- Many messages expressed gratitude for the efforts of the VEuPathDB team and offered moral and financial support.
- Researchers from various disciplines acknowledged the database's role in enabling groundbreaking work.

> *"Thank you for providing this high-quality data served in such a fantastic queryable way for all these years."*
>
> *"Thank you for providing this service and for fighting to keep it alive."*
>
> *"VEuPathDB is the rare biological database that got it right in terms of UI, content and community involvement."*

Analysis and Recommendations

1. Funding Strategy

- A diversified funding model appears necessary, combining:
  - Institutional contributions (e.g., university subscriptions).
  - Tiered or optional user fees.

    - International funding (e.g., WHO, EU, Wellcome Trust).
    - Crowdfunding and sponsorships.
- Transparency regarding operational costs and required funding targets could help justify financial contributions.

2. Balancing Accessibility and Sustainability

- If fees are introduced, they should be structured to minimize barriers for LMIC researchers.
- Free access to core functionalities, with optional paid advanced tools, may be a compromise.

3. Advocacy and Awareness

- Engaging the global scientific community to advocate for reinstated funding.
- Demonstrating the impact of VEuPathDB through case studies and publications to highlight its importance.

4. Operational Adjustments

- Exploring cost-cutting measures without compromising usability.
- Partnering with cloud service providers for discounted hosting.
- Streamlining data management to prioritize essential updates.

Conclusion: The survey responses highlight a strong consensus that VEuPathDB is an irreplaceable resource. While there is some openness to alternative funding models, the community strongly favors solutions that maintain broad accessibility. In addition to an equitable fee-based model, advocacy, diversification of funding sources, and strategic cost reductions may be necessary to ensure the database's survival.

VEuPathDB
Eukaryotic Pathogen, Vector & Host Informatics Resources

Your answers to this anonymous survey will help us to improve VEuPathDB functionality, and develop innovative mechanisms for funding FungiDB, VectorBase, HostDB, OrthoMCL, and various parasitology resources (AmoebaDB, CryptoDB, GiardiaDB, MicrosporidiaDB, PiroplasmaDB, PlasmoDB, ToxoDB, TrichDB, TriTrypDB).

When you are done, please remember to click on the submit button at the bottom of the survey!

1. On a scale from 0-100, please estimate the impact on your lab's productivity if VEuPathDB resources were to disappear altogether (numerical value will appear at right of slider). *How to define impact is up to you: increased time required to do your work, lab budget, number of publications, etc.*

0% = no impact, or I don't use these databases at all; 100% = would shut down my work completely

0 100

2. Please describe how the loss of VEuPathDB databases would impact your work:

3. How else would you accomplish the tasks for which you use VEuPathDB?

- [ ] Not sure
- [ ] I have downloaded what I need and I am all set
- [ ] I downloaded what I need ... but using this will be cumbersome
- [ ] I will use other sites or tools, such as: _____

4. How often do you typically access VEuPathDB sites?

- ○ Daily
- ○ Weekly
- ○ Monthly
- ○ Yearly or less

5. What features have you used most often (check all that apply)?

- [ ] Browse compendium of all information available on Gene page(s)
- [ ] > Gene identifiers, product names, annotation, gene models, sequences
- [ ] > Literature references
- [ ] > Links to other sites

- [ ] > Comparative information (population diversity, synteny views, orthology)
- [ ] > Functional genomics information (transcriptomics, proteomics, etc)
- [ ] > Protein features, structure, subcellular localization
- [ ] > Phenotypic information
- [ ] > Other (pathways, immunoreactivity, etc)
- [ ] Site search
- [ ] Specific searches (e.g. find genes with a specific expression profile or InterPro domain)
- [ ] Search strategies (combining multiple searches using intersects and unions)
- [ ] Analyze gene sets (e.g. GO term or metabolic pathway enrichment analysis)
- [ ] Sequence retrieval
- [ ] Data downloads
- [ ] BLAST
- [ ] Tag your ‘Favorite’ genes, or save in a ‘Basket’
- [ ] Examine or add User Comments
- [ ] Apollo (web-based genome annotation editor)
- [ ] Analyze your own data (map RNAseq/ChIPseq, call SNPs, define orthologs, etc)

- [ ] MapVEu (population biology map in VectorBase)
- [ ] Help and/or tutorials
- [ ] Other

6. Which VEuPathDB websites do you regularly use?

- [ ] AmoebaDB
- [ ] CryptoDB
- [ ] FungiDB
- [ ] GiardiaDB
- [ ] HostDB
- [ ] MicrosporidiaDB
- [ ] OrthoMCL
- [ ] PiroplasmaDB
- [ ] PlasmoDB
- [ ] ToxoDB
- [ ] TrichDB
- [ ] TriTrypDB
- [ ] VectorBase
- [ ] VEuPathDB

7. In what country do you work?

Select one

8. What type of institution do you work in?

- Academic
- Industry
- Government
- Retired
- Other

9. What is your primary position?

- Lab head/Principal Investigator
- PostDoc
- Masters or PhD student
- Instructor
- Staff (e.g., lab manager, technician, etc.)
- Other

10. How many people work in your group?

Select one

11. What percentage of the research in your laboratory depends to some extent on VEuPathDB databases? *For example, enter 33 if 1/3 of your lab works on malaria and 2/3 on cancer.*

12. What is the approximate total annual budget of your lab? Please include the value of all activities: grants, salaries, studentships, facilities, services, etc.

- <$10K per year
- $10-100K per year
- $100-200K per year
- $200-400K per year
- $400-800K per year
- $800-1.5M per year
- >$1.5M per year
- I have no idea!

13. Would you consider it appropriate to institute a fee-for-service model, with charges for access to VEuPathDB paid from users' grants or by their institutions? *Note: Charges would be based on lab budgets in order to ensure accessibility for those without funding, researchers in low-income countries, teaching institutions, etc.*

- ◯ Yes
- ◯ Maybe
- ◯ No

Comments:

14. If you were expected to pay for the services provided by VEuPathDB, are there specific technical difficulties that might have to be addressed?

15. Would you consider it appropriate to institute a tiered pricing structure, with basic services available for free, but where the most recent data, or high-end services, or customer support is available only to those able to contribute financially?

- ◯ Yes
- ◯ Maybe
- ◯ No

Comments:

Optional details so that we can keep providing you with periodic updates

Name

Email

Lab

Institution

Additional information of possible interest:

- The direct cost of maintaining access to existing VEuPathDB databases, even in static form, is ~$500K/yr … although without the ability to load new data, improve interfaces, or develop tools accommodating new datatypes, the useful lifetime of such resources is of course limited.
- In recent years, the direct cost of the entire VEuPathDB project has been ~$5M/yr, including (in addition to web site maintenance) the identification, processing, loading, and QC of new datasets; development and deployment of new data analysis pipelines, tools, and interfaces; and outreach activities to ensure effective utilization by our diverse end-user communities.
- We recognize that this is a considerable sum, although in our opinion quite cost effective, given the scale of database utilization and research impact: 2023 statistics show an average of 46K unique users per month, from >150 countries, with the typical user returning approximately weekly to access >100 web pages per visit, yielding 11M page hits/mo, 7 Tb downloads/yr, and a cumulative total of >35K literature citations to date.

Feel free to add additional comments in the box below:

Submit your Answers